

LA TECHNOLOGY TRANSFER PRESENTA

# MIKE FERGUSON

## LA MODERNIZZAZIONE DEL DATA WAREHOUSE

Nuove tecniche di Data Modeling, Discovery automatizzato dei dati, Il Catalogo Dati,  
Pipeline di Data Ops, Sviluppo e Migrazione del DW su Cloud,  
Data Warehouse Automation, Data Lakehouse, Virtual Data Mart,  
Data Warehouse real-time, Machine Learning Logical DW e Tabelle esterne

ONLINE LIVE STREAMING

4-5 DICEMBRE 2023



info@technologytransfer.it  
www.technologytransfer.it

## DESCRIZIONE

Nell'economia digitale di oggi, il potere del cliente è quasi senza limiti. Possono visitare i siti Web della concorrenza, comparare i prezzi, verificare il sentiment sul vostro business e con un click in qualsiasi momento e in qualsiasi luogo da un device mobile possono cambiare fornitore. In aggiunta i social media hanno dato una voce ai clienti per esprimere opinioni sui vostri prodotti e sul vostro brand. Per molti CEO, la fidelizzazione del cliente, la lealtà, il servizio e la crescita sono al top della loro agenda. Pertanto, desiderano accedere a nuovi dati per arricchire ciò che già conoscono dei clienti. Inoltre le aziende stanno catturando nuovi dati al fine di ottimizzare le operazioni di business.

Questi nuovi dati stanno facendo capire a molte aziende che operano con Data Warehouse e Data Mart tradizionali che questo non è sufficiente per l'Analitica.

Sono necessari altri sistemi e, con il ritmo delle modifiche che accelera, i dati a latenza inferiore e la Machine Learning questi nuovi sistemi sono richiesti ovunque. Tutto ciò è necessario per rimanere competitivi.

Con tanta pressione per rimanere competitivi, come potete modernizzare il vostro set-up analitico, come potete migliorare governance e agilità, inserire nuovi dati, riutilizzare asset di dati, modernizzare il vostro Data Warehouse per farlo adattare facilmente alle modifiche, ridurre la latenza dei dati per l'impresa digitale?

Questo seminario di 2 giorni esamina il business case sul perché è necessario farlo, discute gli strumenti e le tecniche necessari per acquisire nuovi tipi di dati, stabilire una pipeline di dati per produrre data assets riutilizzabili, modernizzare il vostro Data Warehouse e mettere insieme i dati e l'analitica necessari per accelerare il time to value, fornire nuovi spunti per favorire la crescita, ridurre i costi, migliorare l'efficacia e consentire un vantaggio competitivo.

In particolare i partecipanti impareranno:

- A capire perché la modernizzazione del Data Warehouse è necessaria per aiutare a migliorare il decision making e la competitività
- Gli ingredienti per sapere come modernizzare il vostro Data Warehouse per migliorare Agilità, ridurre il TCO e facilitare la manutenzione
- A capire le moderne tecniche di Data Modeling e come ridurre il numero di data stores in un Data Warehouse senza perdere informazioni
- A capire come sfruttare il Cloud Computing a un costo più basso e come migrare al Cloud
- A capire come ridurre la latenza dei dati nel vostro Data Warehouse
- A capire come integrare il vostro Data Warehouse con i dati dei Data Lake memorizzati nel Cloud
- A capire come migrare da un Data Warehouse e Data Marts a cascata a un moderno Logical Data Warehouse con Data Marts virtuali che si integrano facilmente con Cloud storage e altri sistemi analitici
- A sapere come utilizzare la virtualizzazione dei dati per semplificare l'accesso a una serie più ampia di informazioni disponibili su più piattaforme analitiche che eseguono analitiche su diversi tipi di dati per un decision making basato su basi più solide
- A capire il ruolo di un moderno Data Warehouse in una Enterprise data-driven

### Partecipanti

CDO, CIO, IT Manager, CTO, Analisti di Business, Data Scientist, Manager di Business Intelligence, Professionisti di Data Warehousing, Enterprise Architect, Data Architect e Solution Architect

# PROGRAMMA

## 1. Il Data Warehouse tradizionale e perché ha bisogno di essere modernizzato

Nella maggior parte delle organizzazioni oggi il Data Warehouse è ancora basato su un'architettura a cascata con i dati che fluiscono dai sistemi di origine agli operational data stores, alle aree di gestione temporanea, quindi ai Data Warehouse sotto la gestione dei lavori batch ETL.

Tuttavia, il panorama analitico è cambiato. Le nuove fonti di dati continuano a crescere con la raccolta dei dati nei dispositivi edge, nello storage Cloud, nei data stores NoSQL sia locali che sul Cloud, nei sistemi Hadoop e nella staging del Data Warehouse. Cloud Storage, Spark, piattaforme per lo streaming di dati e database Graph sono ora utilizzati anche nell'analisi dei dati. Inoltre, molte unità aziendali utilizzano il Cloud per sfruttare rapidamente queste nuove tecnologie analitiche a costi inferiori.

Questa sessione di apertura esamina queste nuove attività e spiega perché i Data Warehouse devono cambiare non solo per accelerare lo sviluppo, migliorare l'agilità, ridurre i costi ma anche per sfruttare nuovi dati, abilitare la preparazione dei dati self-service, utilizzare analisi avanzate e integrarsi con queste piattaforme analitiche.

- Il Data Warehouse tradizionale
- Data stores multipli, architettura dati a cascata e flussi di dati
- Nuovi dati entrano in azienda
- I cambiamenti dell'Analitica: nuovi data store e piattaforme analitiche
  - o Big Data Analytics su Spark, memoria Cloud e Hadoop
  - o Real-time streaming data analytics
  - o Graph Analysis e Graph Database
- Nuove sfide determinate da:
  - o Complessità dei dati
  - o Silos di gestione dati
  - o Gestire i dati in un ambiente distribuito

e di computing ibrido

- o Dati self-service verso ETL/DQ
- Problemi con l'architettura Data Warehouse esistente e tecniche di sviluppo
- La necessità di evitare i silos, integrare i nuovi dati, applicare velocemente i cambiamenti e integrare workload analitici per ottenere valore

## 2. Requisiti per un moderno Data Warehouse

Questa sessione analizza gli elementi chiave del moderno Data Warehouse che devono essere predisposti per flessibilità e agilità.

- Tecniche moderne di Data Modeling
- Accelerare il processing ETL usando un Data Discovery automatizzato, un catalogo dati con pipeline DataOps e prodotti dati riusabili
- DBMS analitici basati su Cloud
- Tabelle esterne, Lakehouse e analitica nel database
- Accorciare i tempi di sviluppo usando Data Warehouse Automation
- La virtualizzazione dei dati per l'indipendenza dei dati, la flessibilità e l'integrazione di nuovi data store analitici all'interno del Logical Data Warehouse
- Incorporare dati fast streaming, analitica prescrittiva, embedded e operational BI

## 3. Tecniche Moderne di Data Modeling per l'Agile Data Warehousing

Al fine di migliorare l'agilità, sono emerse nuove tecniche di modellazione dei dati che stanno diventando sempre più popolari nella progettazione di moderni Data Warehouse.

Questa sessione esamina la modellazione dei dati e si chiede: lo "Star Schema" è morto? Quale tecnica di modellazione del Data Warehouse è più adatta alla gestione del cambiamento? Dovreste usare Data Vault? La progettazione del Data Warehouse deve

cambiare? Il design del Data Mart deve cambiare? Infine si sofferma anche sugli svantaggi di tali tecniche e su come superarli.

- Approcci di modeling del Data Warehouse: Inmon verso Kimball verso Data Vault
- La necessità di gestire facilmente il cambiamento
- Cosa è il Data Vault?
- Componenti di modeling del Data Vault: hubs, links e satellites
- Pro e contro delle tecniche di Data Modeling
- Usare la virtualizzazione dei dati per migliorare l'Agilità nei Data Mart mentre si riducono i costi

#### 4. Modernizzare il vostro Processing ETL

Questa sessione esamina le sfide poste dai nuovi dati sul processing ETL. Inoltre, quali opzioni sono disponibili per modernizzare il processing ETL, dove dovrebbe essere eseguito e quali sono i pro e i contro di ciascuna opzione? In che modo questo ha un impatto sull'architettura dei dati?

- Nuovi dati e processing ETL
  - o alto volume di dati, dati semi-strutturati, dati non strutturati, dati streaming (ad esempio dati IoT)
- Quali sono le implicazioni e le sfide di questi nuovi dati sul processing ETL?
- Tutti questi dati dovrebbero andare nel Data Warehouse o no?
- Quali opzioni sono disponibili per modernizzare il processing ETL del Data Warehouse?
  - o Data Lake o Lakehouse o Data Mesh
  - o Scaricare i dati di staging su un Data Lake e usare Spark per il processing ETL scalabile
  - o L'opzione dati real-time
  - o Usare il software di Data Warehouse Automation per generare processing ETL
- Pro e contro di queste opzioni
- Le implicazioni sull'Architettura dati sulla modernizzazione e democratizzazione del Data engineering

#### 5. Accelerare il processing ETL usando prodotti dati, un Data Lake multi-purpose, un Lakehouse o un Data Mesh

Questa sessione esamina come utilizzare un Data Lake multiuso, Lakehouse e Data Mesh per accelerare il processing ETL e l'integrazione dei dati per il Data Warehouse.

- Come si può accelerare il processing ETL e la preparazione dei dati self-service?
- De-centralizzare il Data Engineering
- Usare Data Fabric per costruire pipeline DataOps per creare prodotti dati riusabili in un Data Mesh
- Opzioni per l'implementazione di Data Mesh: Data Lake, Lakehouse, piattaforme dati su Cloud, Kafka e virtualizzazione dei dati
- Accedere direttamente alle sorgenti dati invece che immetterli nella staging area pronti per il processing ETL
- Usare un Catalogo dati per il Discovery automatico, classificare, profilare, catalogare e mappare i dati in un glossario di business
- GDPR: individuare dati sensibili durante un Data Discovery automatico
- Mascherare dati sensibili GDPR durante l'immissione o il processing pipeline ETL
- Usare Machine Learning nelle pipeline DataOps per processare dati non strutturati
- Pipeline processing per streaming data in un Data Warehouse real-time
  - o Tipi di streaming data: dati IoT, weblog, OLTP, ecc.
  - o Principali tecnologie per il processing di streaming data: Kafka, streaming analytics e event stores
  - o Portare OLTP change data capture nei data stream di Kafka
  - o Usare Kafka come sorgente data per processare dati in real-time
  - o Il futuro proofing streaming ETL processing usando Apache Beam

o Immettere streaming data nel vostro Lakehouse o Data Warehouse

- Data Warehouse real-time: integrare il vostro Data Warehouse con gli streaming data - tabelle esterne, virtualizzazione dei dati e Data Lake
- Usare data pipelines ETL per produrre prodotti dati riusabili da utilizzare nel vostro Data Warehouse, Data Science o in altri data store analitici
- Pubblicare i dati riusabili in un catalogo pronto per il consumo
- Usare la Data Science per sviluppare nuovi modelli analitici da girare nel vostro Data Warehouse

## 6. Sviluppo Rapido del Data Warehouse usando Data Warehouse Automation

Oltre a modernizzare un processing ETL, questa sessione analizza come utilizzare gli strumenti di Data Warehouse automation basati sui metadati per creare, modificare ed estendere rapidamente i moderni Data Warehouse e Data Mart sia sul Cloud che on premises. Spiega come questi strumenti vi aiutano ad adottare rapidamente nuove moderne tecniche di modellazione dei dati, come generano schemi e lavori di integrazione dei dati e come possono aiutarvi a migrare verso i tuoi nuovi sistemi di Data Warehouse sul Cloud.

- Cosa è Data Warehouse Automation?
- Usare i tools di Data Warehouse Automation per lo sviluppo rapido di Data Warehouse e Data Mart
  - o Generare Data Vault E/R e design Star Schema
  - o Data pipeline generation
  - o Processing di streaming data usando Data Warehouse Automation
  - o Integrare Big Data con il Data Warehouse usando Data Warehouse Automation
  - o Integrare Data Warehouse sul Cloud con Data Warehouse Automation
  - o Usare Data Warehouse Automation per migrare Data Warehouses
  - o Usare la virtualizzazione dei dati per proteggere gli strumenti di BI esistenti dai cambiamenti nel design

- Il mercato dei tools di Data Warehouse Automation: IDERA WhereScape, biGENIUS, Qlik Attunity Compose, Time Extender, BIML Studio, Vault Builder e altri
- Manutenzione del Data Warehouse basato sui metadati

## 7. Costruire un moderno Data Warehouse in un ambiente di Cloud Computing

Una domanda chiave per molte organizzazioni è che cosa fare con il nostro Data Warehouse esistente? Dovreste provare a cambiare il set up esistente per renderlo più moderno o ri-svilupparlo nel Cloud? Questa sessione esamina i vantaggi della creazione di moderni Data Warehouse in un ambiente di Cloud Computing utilizzando un DBMS relazionale analitico basato su Cloud.

- Perché usare il Cloud Computing per il vostro Data Warehouse?
- Sviluppo del Data Warehouse basato su Cloud: quali sono le opzioni?
- DBMS relazionali analitici basati su Cloud
  - o Amazon Redshift, Google BigQuery, IBM Db2 Warehouse on Cloud, Microsoft Azure Synapse Analytics, Oracle Autonomous Data Warehouse, Snowflake, SAP Data Warehouse Cloud, Teradata, Kinetica,
- Lakehouses: Apache Iceberg, Databricks, HPE Ezmeral
- Separare memoria dal calcolo per elasticità e scalabilità
- Gestire e integrare dati Cloud e on-premises
- Usare software iPaaS per integrare dati in un ETL processing su Cloud: Informatica IICS, Boomi, Snap Logic, ecc.
- Tools Cloud ETL non-IPaaS: AWS Glue, Azure Data Factory, Google Cloud Dataplex e Data Fusion, IBM Cloud Pak for Data, Talend, Software AG StreamSets
- Gestire streaming data sul Cloud
- Integrare Big Data Analytics in un Data Warehouse basato su Cloud

- Collocare modelli di Machine Learning nel vostro data base analitico per l'analitica in-Warehouse
- Tools e tecniche per migrare sul Cloud un Data Warehouse esistente
- Affrontare i problemi della migrazione al Cloud DW
- Gestire gli accessi ai Data Warehouse basati su Cloud
- Integrare tools BI basati su Cloud con i sistemi on-premises

### **8. Semplificare l'accesso ai dati: creare Data Mart virtuali e una Architettura di Logical Data Warehouse per integrare i Big Data con il vostro Data Warehouse**

Questa sezione esamina come utilizzare il software di virtualizzazione dei dati per modernizzare l'architettura del Data Warehouse, semplificare l'accesso, integrare i dati nel Data Warehouse e i Big Data sottostanti i data store e migliorare l'agilità.

- Cosa è la virtualizzazione dei dati?
- Come lavora la virtualizzazione dei dati?
- Come la virtualizzazione può ridurre il costo di ownership, migliorare l'agilità e modernizzare la vostra architettura di Data Warehouse?
- Semplificare la vostra architettura usando la virtualizzazione dei dati per creare Data Mart virtuali
- Migrare i vostri Data Mart fisici ai Data Mart virtuali per ridurre il costo di ownership
- Layering di tabelle virtuali sopra i Data Mart virtuali per semplificare l'accesso agli utenti di business
- Pubblicare viste e query virtuali come servizi in un catalogo dati per il consumo
- Integrare il vostro Data Warehouse con il vostro Data Lake e i dati a bassa latenza usando tabelle esterne e virtualizzazione dei dati
- Permettere un rapido change management usando la virtualizzazione dei dati
- Creare una Architettura di Logical Data warehouse che integra i dati dalle piattaforme Big Data, graph

- database, piattaforme di streaming data con il Vostro Data Warehouse in un comune livello di accesso per - permettere un facile accesso ai tools e alle applicazioni di BI
- Usare un glossario di business e la virtualizzazione dei dati per creare un comune livello semantico con una comune comprensione coerente da parte dei tools di BI

### **9. Iniziare con la modernizzazione del Data Warehouse**

Questa sessione finale esamina ciò che dovrete fare per iniziare un'iniziativa di modernizzazione del Data Warehouse. In particolare:

- Modernizzazione del Data Warehouse: cambiare o ri-costruire?
- In quale ordine farlo?
- Come minimizzare l'impatto sul business mentre eseguite la modernizzazione?
- Come gestire il backlog dei cambiamenti mentre state cercando di modernizzare?
- Pro e contro sulla costruzione o automatizzare lo sviluppo del Data Warehouse
- Quali nuovi skill sono necessari?
- Fornire nuovo valore di business mentre è in corso la modernizzazione
- Come coinvolgere i professionisti di business nello sforzo di modernizzazione?

## INFORMAZIONI

<b>QUOTA DI PARTECIPAZIONE</b> € 1200 (+iva)	<b>MODALITÀ D'ISCRIZIONE</b> Il pagamento della quota, IVA inclusa, dovrà essere effettuato tramite bonifico, codice IBAN: IT 03 W 06230 03202 000057031348 Banca: Credit Agricole 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  TECHNOLOGY TRANSFER S.r.l. Piazza Cavour, 3 00193 ROMA (Tel. 06-6832227)	<b>CONDIZIONI GENERALI</b> 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.  <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>TUTELA DATI PERSONALI</b> 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 E ORARIO</b> 2 giorni: 9.30-13.00 14.00-17.00			
<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

#### LA MODERNIZZAZIONE DEL DATA WAREHOUSE

4-5 Dicembre 2023

Quota di iscrizione: € 1200 (+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.

È previsto il servizio di traduzione simultanea

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

## ***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.