



DATA MESH (DATI DECENTRALIZZATI) : FINALMENTE GIUNTA LA MATURITÀ?

Il concetto di **Data Mesh** è ormai considerato una soluzione che porta ad una maggiore collaborazione. Tuttavia, l'approccio si è scontrato con la realtà sul terreno di sistemi storicamente isolati e team divisi tra il decentramento richiesto e il modello centralizzato di base su cui tutto è stato costruito. **È ancora possibile trovare un equilibrio tra centralizzazione e autonomia nella trasformazione Data-Driven?**



STEFANO SPONGA

Head of Insights & Data Center of
Excellence Italy



LUCA BERTOZZI

CTO Insights & Data Italy



I dati sono le unità materiali con cui costruiamo prodotti, servizi, automazione, per creare più valore. Allo stesso tempo, la digitalizzazione delle attività e degli scambi, la dematerializzazione e l'*Internet of Things* stanno moltiplicando le fonti di dati, tanto abbondanti quanto varie. In breve, abbiamo molte più informazioni e molti più modi per sfruttarle. Questo porta maggiori quantità di dati spesso raccolti troppo repentinamente.

La tradizionale gestione centralizzata dei dati, adatta alla BI, si sta rivelando poco pratica nell'era dell'AI. I *Datalake* monolitici stanno diventando così vasti e complessi che non è più umanamente possibile sapere esattamente che cosa contengono. Ognuno usa solo ciò che sa, la collaborazione è molto difficile, incrociare i dati - dove risiede il valore - è impossibile. Di conseguenza, la stragrande maggioranza dei dati raccolti rimane incolta, in attesa di un caso d'uso che non arriverà mai. In quest'ottica l'approccio Data Mesh, che **propone di suddividere il monolite dei dati in diversi domini business-oriented**, ha subito suscitato molto interesse perché è sembrato la risposta perfetta a queste sfide. Purtroppo, le prime iniziative sono state spesso deludenti. Riluttanti a smantellare ciò che avevano costruito così pazientemente, i team che lavorano sui dati, a cui di solito è affidato il compito di mantenerli, costruiscono architetture complicate basate su repository, gateway e strumenti intermedi tra fonti di dati e prodotti finali.

Ossia approcci che spesso si sono dimostrati inefficienti e, sicuramente, inadatti a indirizzare l'approccio Data Mesh.

LA SFIDA: DECENTRALIZZARE SENZA CREARE NUOVI SILOS

La sfida è decentralizzare i dati per renderli più gestibili senza creare nuovi silos. In particolare, occorre preservare in primo luogo la collaborazione - in modo che i gruppi di lavoro possano arricchire e sfruttare l'asset dati con un coordinamento minimo successivamente - e l'estensibilità - in modo che questo patrimonio possa continuare a crescere senza pesare sull'agilità necessaria per nuovi usi.

L'IMPERATIVO: RISPETTARE I PRINCIPI FONDAMENTALI DEL DATA MESH

Data Mesh può essere considerato come la declinazione per i dati di *Domain Driven Design* (DDD), ovvero un **approccio alla progettazione di sistema che postula che il valore di un sistema risiede nella riconciliazione trasversale dei suoi componenti**. DDD mira quindi a massimizzare il disaccoppiamento tra i componenti tecnologici del sistema e tra i gruppi che li creano, producono e mantengono. In altre parole, ogni dominio è gestito a modo suo da chi lo conosce meglio e tutti sono a livello di interfaccia. **Data Mesh**, che può essere definito come un **approccio sociotecnico decentralizzato per gestire e accedere ai dati analitici su larga scala**, scompone la filosofia di DDD in quattro principi principali:

1. la divisione per domini governance federata;
2. la presentazione dell'iniziativa supportata da prodotti adeguati;
3. una piattaforma self-service;
4. l'implementazione secondo regole e principi;

Il successo di data mesh richiede l'applicazione rigorosa di questi quattro principi. Qualora mancasse uno di essi, l'applicazione del paradigma risulterebbe incompleto e non porterebbe i benefici attesi.

I QUATTRO ASSI DI TRASFORMAZIONE

1 EVOLVERE VERSO UN'ARCHITETTURA PIÙ DECENTRALIZZATA, GUIDATA DA CASI D'USO

2 IMPLEMENTARE LA GOVERNANCE FEDERATA

Gli asset di dati sono suddivisi in domini gestiti da gruppi indipendenti, incaricati di mantenere il perimetro, la qualità e l'omogeneità e offrire prodotti di dati chiavi in mano, facilmente comprensibili da parte di non-specialisti. Naturalmente, la condizione di base è preservare l'interoperabilità dei domini. Per questo è possibile utilizzare una metodologia di modellazione comune come Data-Vault, che consente di concordare chiavi incrociate lasciando libertà d'azione. Inizieremo preferibilmente con le aree più mature e il cui ritorno sull'investimento appare il più immediato. Per mantenere la coerenza col patrimonio dati comune, tutti devono poter contribuire. È a questo livello che si deciderà in merito alla creazione e alla divisione dei domini, alle regole di interoperabilità, agli strumenti comuni, ecc.

3 CONFIGURARE UNA PIATTAFORMA SELF-SERVICE

4 SENSIBILIZZARE E FORMARE GLI STAKEHOLDER INTERNI

Data Mesh non è una tecnologia, ma un principio di gestione e sviluppo collettivo, basato su strumenti adeguati e in grado di consentire questa flessibilità preservando la coesione complessiva e la capacità di scalabilità.

Oggi troviamo tali strumenti, robusti e facili da usare, in Snowflake e Google Cloud Platform (BigQuery).

Nella logica del mantenimento della coesione complessiva, un approccio "Infrastructure as Code" (IaC) garantisce la standardizzazione dell'infrastruttura dei nodi Data Mesh e ne facilita la manutenzione.

Questo è senza dubbio il progetto più delicato. In primo luogo, l'organizzazione centrale dei dati vedrà il suo ruolo profondamente messo in discussione, anche se ha generalmente investito molto negli ultimi anni per rafforzarsi. Non deve vivere il Data Mesh come un disconoscimento o un declassamento, ma deve prendere coscienza della sua importanza chiave al centro del sistema, poiché può supportare le linee di business nella propria evoluzione, diffondere le competenze necessarie, animare la governance federata, mantenere il core repository, e perfino affrontare sfide trasversali di *tooling*, interoperabilità e sicurezza.

Dal punto di vista del business, sarà necessario sensibilizzare, riorganizzare, formare e coinvolgere tutti gli attori che hanno appetito per i dati ma che attualmente sono molto poco coinvolti nella sua gestione (ad esempio gli attuari nel settore assicurativo).

Il Data Mesh non è necessario per qualunque tipo di azienda.

Finché il modello centralizzato soddisfa i requisiti aziendali di tempestività e qualità dei dati, non c'è motivo di cambiare. D'altra parte, non appena i prodotti di dati diventano troppo numerosi e servono leve operative basate sui dati, diventa essenziale un approccio decentralizzato. In questo contesto, il Data Mesh è il modo migliore per implementarlo.

Il Data Mesh è l'abilitatore del data-driven. Spetta ai professionisti prendere coscienza della difficoltà di sviluppare i propri progetti di dati con la struttura centralizzata esistente o anche in modo tradizionale, in "shadow IT". Questa trasformazione richiede la consapevolezza e la formazione dei team sulle nuove possibilità offerte dai dati e comporta uno scouting di casi d'uso in collaborazione con l'IT, nonché un graduale trasferimento delle competenze dei dati alle linee di business e ai team di sviluppo dell'IS operativo.

DA RICORDARE

Sta diventando sempre più difficile far fronte all'esplosione delle esigenze e delle fonti di dati con un modello di gestione centralizzato. Occorre un approccio decentralizzato orientato al business: il Data Mesh è la risposta ideale, ma le delusioni iniziali evidenziano l'importanza di rispettarne i principi.





About Capgemini

Capgemini is a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of over 360,000 team members more than 50 countries. With its strong 55-year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fueled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2022 global revenues of €22 billion.

Get The Future You Want | www.capgemini.com



This presentation contains information that may be privileged or confidential and is the property of the Capgemini Group.

Copyright © 2024 Capgemini. All rights reserved.