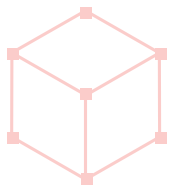


DATA FREEDOM E GOVERNANCE:

Denodo al servizio dello
European Data Act



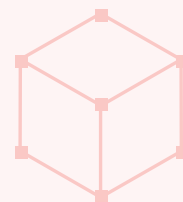
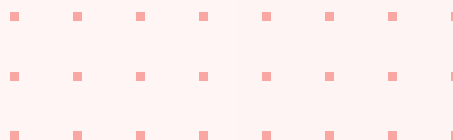


Table of Contents

Premessa	3
Introduzione	5
L'European Data Act, in pillole	6
L'impatto sulle organizzazioni del Data Act	8
Le architetture logiche al servizio dello European Data Act	9
Allegati	12



Lo European Data Act cambia le regole del gioco per l'accesso, la condivisione e la governance dei dati. Denodo offre alle aziende una soluzione agile e sicura per essere compliant, trasformando gli obblighi normativi in un vantaggio competitivo attraverso un'architettura dati logica, flessibile e scalabile

Premessa

Il documento esplora il contributo della tecnologia – e in particolare quello delle architetture logiche per l'integrazione dei dati – nell'ambito del Data Act, un regolamento che ha tuttavia un impatto profondo e diffuso all'interno delle organizzazioni alle quali si riferisce.

Questa pervasività è tale da rendere necessario, prima di proseguire, il chiarimento che ciò di cui si parlerà non è sufficiente a garantire il successo dell'adeguamento a ciò che il regolamento prevede, ma al più ne rappresenta condizione necessaria, sicuramente non sufficiente, visto che l'adeguamento al Data Act – ma in generale a qualsiasi norma o regolamento – investe profondamente l'ambito organizzativo, implicando normalmente una revisione, anche profonda, dei processi interni all'organizzazione, ma anche quello culturale, essendo sovente chiesto alle persone che in tali organizzazioni operano, di pensare in modo diverso e di cambiare il modo in cui hanno fatto ciò che dovevano fare.

Va anche sottolineato, che il Data Act si colloca in un contesto ampiamente regolamentato, dove sono molte le direttive che devono essere rispettate, per cui, anche se nel documento si esplorerà solamente tale regolamento, non va dimenticato che, quando detto, andrà in ogni caso armonizzato con ciò che si è fatto, si sta facendo e si farà in relazione ad altri regolamenti o normative.



EXECUTIVE SUMMARY

I Il problema

Lo European Data Act (EDA) segna un passaggio fondamentale nella costruzione di un'economia europea dei dati equa, interoperabile e sostenibile. Nasce con l'obiettivo di riequilibrare il potere nell'ecosistema digitale, garantendo accesso equo, trasparente e sicuro ai dati generati da dispositivi connessi, promuovendone la condivisione tra aziende, cittadini e istituzioni pubbliche. Tuttavia, la realizzazione di questo obiettivo richiede una profonda revisione dell'approccio alla gestione dei dati, oggi troppo frammentato, poco accessibile e poco governato.

I Conclusioni principali

- Il Data Act impone **nuovi diritti e obblighi** in termini di accesso, interoperabilità, condivisione e portabilità dei dati tra attori pubblici e privati.
- Richiede architetture dati che siano **flessibili, resilienti, scalabili e in grado di preservare la sovranità e la governance dei dati**.
- Le **architetture logiche**, in particolare la **Denodo Platform**, offrono un modello ideale per realizzare quanto previsto dal regolamento: integrazione senza duplicazione, gestione centralizzata dei metadati, semantica condivisa e accesso self-service ai dati.
- Componenti come il **modello semantico**, il **Data Catalog**, e la **governance logica** consentono di creare un ecosistema dati condiviso, modulare, facilmente accessibile e pronto all'uso da parte di persone, applicazioni e sistemi intelligenti.
- Le **tecnologie di intelligenza artificiale** trovano nella Denodo Platform un ambiente pronto per l'adozione, grazie alla disponibilità di dati ben modellati, accessibili e governati.

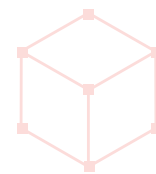
I Raccomandazioni

- **Adottare architetture logiche dei dati** per separare l'accesso dai vincoli fisici e abilitare l'integrazione distribuita nel rispetto delle norme EDA.
- **Costruire un modello semantico attivo** per armonizzare significato e utilizzo dei dati provenienti da attori differenti.
- **Centralizzare la governance dei dati** a livello logico per gestire accessi, mascheramenti e conformità in modo scalabile e indipendente dall'infrastruttura tecnica.
- **Abilitare la democratizzazione dei dati** attraverso strumenti di Data Catalog e Data Marketplace, consentendo un uso responsabile e diffuso del patrimonio informativo.
- **Supportare AI e RAG con** dati contestualizzati e strutturati logicamente, per migliorare l'efficacia degli assistenti virtuali e delle decisioni automatizzate.
- **Promuovere hub locali o settoriali** per federare dati a livello territoriale e semplificare l'adozione del Data Act da parte delle aziende più piccole e con scarse competenze tecnologiche.

I Conclusioni principali

Attraverso l'adozione di soluzioni come la Denodo Platform, le organizzazioni possono non solo garantire la conformità al Data Act, ma trasformare il dato in un asset strategico al servizio dell'innovazione, della competitività e del bene comune.





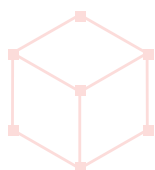
Introduzione

Nel pieno della trasformazione digitale, l'Unione Europea ha avviato una serie di iniziative normative volte a costruire un mercato unico dei dati, in cui questi ultimi possano circolare in modo sicuro, trasparente e controllato, generando valore per cittadini, imprese e istituzioni.

Tra queste iniziative, lo European Data Act (EDA), operativo in Italia dal 12 settembre 2025, rappresenta un tassello centrale: un regolamento pensato per regolare l'accesso, l'utilizzo e la condivisione dei dati in un'economia sempre più connessa.

Il presente documento si propone di analizzare il Data Act nei suoi principi ispiratori e nelle implicazioni operative, con particolare attenzione al ruolo che le architetture logiche di integrazione dei dati possono giocare nell'aiutare le organizzazioni a essere compliant e, al tempo stesso, a cogliere le opportunità di innovazione e competitività offerte da un uso più libero, governato e consapevole del patrimonio informativo.

Alla luce dell'impostazione del documento e della sua stretta connessione con gli aspetti tecnologici che definiscono le scelte in merito di integrazione, governance e fruizione dei dati, si opererà una riduzione rispetto agli ambiti trattati dal regolamento, concentrando la trattazione su quelle sezioni (CAPI, secondo la terminologia del regolamento) che, più di altre, vedono tali temi centrali per la loro effettiva operatività.



L'European Data Act, in pillole

Lo European Data Act, adottato dall'Unione Europea come parte della strategia per un mercato unico digitale, rappresenta un cambiamento epocale nel modo in cui i dati vengono trattati, scambiati e valorizzati all'interno dell'UE. Il regolamento si inserisce in un quadro normativo più ampio, che include il Data Governance Act, l'AI Act e il Digital Services Act, con l'obiettivo comune di rafforzare la sovranità digitale europea e stimolare un'economia dei dati equa e sostenibile.

Lo spirito del Data Act risiede nella volontà di riequilibrare i rapporti di forza nel mercato dei dati, garantendo che anche PMI, enti pubblici e cittadini possano accedere ai benefici derivanti dalla valorizzazione del dato. Attualmente, una parte significativa dei dati generati all'interno dell'Unione rimane inutilizzata o concentrata nelle mani di pochi attori dominanti, limitando la concorrenza e frenando l'innovazione. Il Data Act si propone quindi di rendere il dato un bene condiviso, promuovendo la creazione di un ecosistema di data sharing trasparente, interoperabile e sicuro.

Uno dei concetti chiave introdotti dal regolamento è quello di accesso equo e trasparente ai dati generati dagli oggetti connessi (Internet of Things). Il Data Act stabilisce che l'utente di un dispositivo connesso, sia esso un individuo o un'organizzazione, ha il diritto di accedere ai dati che tale dispositivo genera e, se lo desidera, di condividerli con terze parti. Questo principio estende il concetto di portabilità dei dati introdotto dal GDPR, adattandolo all'era dell'IoT e dell'intelligenza artificiale.

Un altro elemento centrale è la promozione dell'uso dei dati da parte del settore pubblico in situazioni di interesse generale (Business-to-Government Data Sharing, o B2G). Il Data Act prevede che le aziende private, in determinate circostanze e con adeguate garanzie, mettano a disposizione i propri dati alle pubbliche amministrazioni per affrontare emergenze, eventi straordinari o esigenze di pianificazione pubblica.

Sono questi i temi che saranno approfonditi nei successivi paragrafi, con un'ovvia focalizzazione sui dati, su come deve cambiare il modo di gestirli, sul perché serve farlo e sul contributo che le architetture logiche posso dare al riguardo, perché, in fondo, i dati sono come Giano, una delle divinità più antiche della religione romana e latina, in grado di guardare al passato e al futuro, esattamente come fanno i dati, che sintetizzano ciò che è accaduto e, attraverso la loro analisi, consentono di prevedere cosa potrebbe succedere nel futuro.

Gli obiettivi principali del Data Act possono essere sintetizzati in:

- favorire la **circolazione** dei dati all'interno dell'UE, abbattendo barriere tecniche e legali;
- garantire un **accesso equo** ai dati, con regole chiare per utenti, produttori e terze parti;
- stimolare la nascita di **nuovi modelli di business** basati sulla condivisione dei dati;
- rafforzare la **sovranità digitale** dei cittadini e delle organizzazioni europee;
- assicurare **trasparenza, sicurezza e fiducia** attraverso regole di governance condivise.



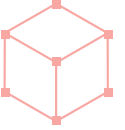
L'ambizione dell'EDA è quella di costruire un mercato unico dei dati, un'infrastruttura digitale dove i dati possano fluire agevolmente tra settori, territori e attori diversi, generando valore economico e sociale. In questo senso, il Data Act rappresenta una tappa fondamentale verso una democrazia del dato, dove ogni soggetto – pubblico o privato – può contribuire e beneficiare della circolazione dei dati, nel rispetto dei principi di proprietà, responsabilità e sostenibilità.

Per concretizzare questa visione, le organizzazioni devono dotarsi di modelli di gestione e condivisione dei dati che siano flessibili, scalabili e rispettosi delle regole stabilite dal regolamento e, non ultime, fare le opportune scelte per definire la base tecnologica ideale per l'attuazione dei principi ispiratori del Data Act.

Con riferimento alla struttura dell'EDA e agli obiettivi del presente documento, quindi, la trattazione prenderà spunto da quei CAPI che, più di altri, sottendono implicazioni tecnologiche volte a rispettare e facilitare quanto previsto dal regolamento. Nello specifico:

- **Capo II.** I prodotti connessi sono progettati e fabbricati e i servizi correlati sono progettati e forniti in modo tale che i dati, compresi i pertinenti metadati necessari a interpretare e utilizzare tali dati, siano, per impostazione predefinita, accessibili all'utente in modo facile, sicuro, gratuito, in un formato completo, strutturato, di uso comune e leggibile da dispositivo automatico e, ove pertinente e tecnicamente possibile, in modo diretto.
- **Capo III.** Il titolare dei dati che, nel quadro di relazioni tra imprese, è tenuto a mettere i dati a disposizione di un destinatario, concorda con il destinatario dei dati le modalità della messa a disposizione dei dati e lo fa a condizioni eque, ragionevoli e non discriminatori e in modo trasparente.
- **Capo V.** Qualora un ente pubblico, la Commissione, la Banca centrale europea o un organismo dell'Unione dimostri una necessità eccezionale di utilizzare taluni dati, ivi compresi i pertinenti metadati necessari per interpretare e utilizzare tali dati, per svolgere le proprie funzioni statutarie nell'interesse pubblico, i titolari dei dati li mettono a disposizione su richiesta motivata.
- **Capo VI.** I fornitori di servizi di trattamento dei dati adottano opportune misure per consentire ai clienti di passare a un servizio di trattamento dei dati, che copre lo stesso tipo di servizio ed è fornito da un diverso fornitore di servizi di trattamento dei dati, o a un'infrastruttura TIC locale, o, se del caso, di utilizzare più fornitori di servizi di trattamento dei dati contemporaneamente.
- **Capo VIII.** Per facilitare l'interoperabilità dei dati nonché degli spazi comuni europei di dati, i partecipanti agli spazi di dati che offrono dati o servizi di dati ad altri partecipanti devono rispettare i requisiti essenziali che consentano un'effettiva condivisione e circolazione dei dati.





L'impatto sulle organizzazioni del Data Act

Il Data Act, introdotto dall'Unione Europea, rappresenta un pilastro fondamentale della strategia europea per i dati. Mira a favorire un'equa condivisione dei dati generati nell'ambito dell'economia digitale, promuovendo un accesso più aperto, sicuro ed efficiente tra imprese, cittadini e istituzioni. Per cogliere appieno lo spirito del regolamento, le aziende devono affrontare in modo integrato e armonico le dimensioni organizzativa, culturale e tecnologica.

DIMENSIONE ORGANIZZATIVA

Il Data Act richiede una ridefinizione dei modelli operativi e di governance, in particolare:

- l'istituzione di processi chiari per la gestione, condivisione e accesso ai dati,
- l'identificazione di ruoli e responsabilità specifici, come data steward, responsabili legali e IT, per garantire la conformità normativa;
- l'integrazione del data sharing nei processi di business, bilanciando trasparenza e protezione degli interessi aziendali.

Un'implementazione efficace non può essere improvvisata: servono strutture organizzative solide, orientate alla valorizzazione dei dati come asset condiviso e strategico.

DIMENSIONE CULTURALE

Il successo del Data Act dipende fortemente da un cambiamento culturale profondo:

- occorre superare la logica proprietaria e promuovere una cultura della collaborazione e della fiducia nell'uso e nella condivisione dei dati;
- è fondamentale sensibilizzare i dipendenti sui principi di equità, trasparenza e responsabilità che il regolamento incarna;
- serve un impegno nella formazione continua, per garantire che tutti gli attori coinvolti comprendano i diritti, i doveri e le opportunità legate all'economia dei dati. La cultura aziendale deve accompagnare ogni scelta organizzativa e tecnologica, facilitando l'adozione di nuovi comportamenti e mentalità.

DIMENSIONE TECNOLOGICA

La tecnologia rappresenta un potente abilitatore per l'attuazione del Data Act, ma non ne costituisce il punto di partenza.

Prima di investire in soluzioni IT, occorre costruire una visione strategica e condivisa del valore dei dati, supportata da processi e policy coerenti. In parallelo, è indispensabile:

- selezionare tecnologie aperte, interoperabili e scalabili, che facilitino l'accesso e la condivisione sicura dei dati;
- garantire piattaforme intuitive e user-friendly, affinché tutti gli utenti – tecnici e non – possano operare facilmente;
- monitorare costantemente l'efficacia delle soluzioni adottate rispetto agli obiettivi del regolamento.

Tecnologie complesse, chiuse o poco accessibili rischiano di compromettere l'efficacia e l'accettazione del Data Act. Solo se integrate armoniosamente con le dimensioni organizzativa e culturale, possono diventare leve concrete di facilitazione e successo.

CONCLUSIONI

L'attuazione del Data Act richiede una gestione bilanciata delle sue tre dimensioni chiave. Nessun ambito, da solo, può garantire il pieno rispetto dello spirito del regolamento. Solo un approccio coordinato e trasversale consente di trasformare la conformità normativa in un'opportunità di efficienza, innovazione e creazione di valore condiviso.

Le architetture logiche al servizio dello European Data Act

L'EDA è una iniziativa estremamente pervasiva e tocca, con le sue direttive, l'ambito organizzativo, legale e culturale delle diverse organizzazioni, sancendo di fatto la necessità di un cambio di passo in quello che può considerarsi il cammino verso un'economia dei dati che sia veramente tale, senza sproporzioni nei rapporti di forza e dove tutti abbiano le stesse possibilità di accesso e uso dei dati.

Quando una trasformazione è così profonda, il ruolo della tecnologia deve essere ben compreso e correttamente collocato, per evitare che a questo sia dato un ruolo eccessivamente centrale, cosa che potrebbe far perdere di vista l'essenza stessa della trasformazione, ma nemmeno relegarlo a elemento di sfondo, con il rischio di dimenticare il ruolo di abilitatore che la tecnologia ha e che, se non ben indirizzato, potrebbe da solo sancire il fallimento dell'iniziativa.

Comprendendo questo ruolo di abilitatore e in scenari in cui la condivisione dei dati è elemento centrale, come appunto lo è nell'EDA, e coinvolge diverse entità – pubbliche e private – e in cui l'accesso deve essere disponibile sostanzialmente a chiunque abbia l'esigenza di utilizzare i dati, l'adozione di architetture logiche per l'integrazione dei dati – e nello specifico la Denodo Platform – è essenziale, in quanto queste forniscono quella semplicità, flessibilità, sicurezza ed efficienza necessarie in ambienti complessi.

Un'economia dei dati fondata su un sistema democratico non può prescindere da un impianto tecnologico che consenta facilmente a tutti di esercitare i propri diritti in relazione ai dati – se un diritto non è facilmente esercitabile, allora diventa un privilegio – una facilità che deve pervadere l'intero ciclo di vita dei dati, dalla necessità di una loro messa a disposizione di chi ne ha bisogno, fino all'effettivo uso, che ne sprigiona il valore. Questa armonia, necessariamente omogenea su tutte le fasi che scandiscono la vita dei dati, è il risultato di un'alchimia – non banale – sulla quale si basano le architetture logiche e, nello specifico, la Denodo Platform, che si fonda su un approccio che non richiede la duplicazione dei dati, ma sfrutta la virtualizzazione e una sapiente gestione dei metadati, consentendo di unificare i dati provenienti da diverse fonti senza compromettere l'autonomia e la governance dei fornitori di tali dati.

L'essenza stessa delle architetture logiche, che si basa su una separazione logico-fisica dei dati e sul conseguente concetto di “connessione” alle diverse sorgenti, un approccio che si differenzia dal tradizionale approccio ETL, basato invece sulla replica dei dati in quello che poi sarà il loro punto di accesso, rende la messa a disposizione di nuovi dati estremamente più semplice e veloce – e al contempo sicura – così da poter rispondere efficacemente a richieste di soggetti esterni, così come normato dall'EAD.

Lo strato logico, che funge da livello di disaccoppiamento tra i Data Consumer e l'infrastruttura tecnica, spesso complessa e variegata e necessaria a gestire tutti i dati disponibili, è sede del modello semantico che realizza la concettualizzazione del contesto di riferimento e la mette a disposizione di tutti, in modo chiaro, coerente e non ambiguo, arricchendo i dati di tutti quegli elementi – o meglio, meta-elementi – che sono funzionali a una loro corretta individuazione, comprensione e corretto utilizzo.

Il modello semantico, inoltre, è luogo della gestione e della garanzia del rispetto delle leggi sulla privacy, dei protocolli di sicurezza e delle linee guida etiche, definendo esplicitamente i diritti di accesso e i vincoli di utilizzo, cosa che favorisce la fiducia tra tutti gli stakeholder, incoraggiando un più ampio utilizzo dei dati e preservandone al contempo la sovranità e l'integrità.

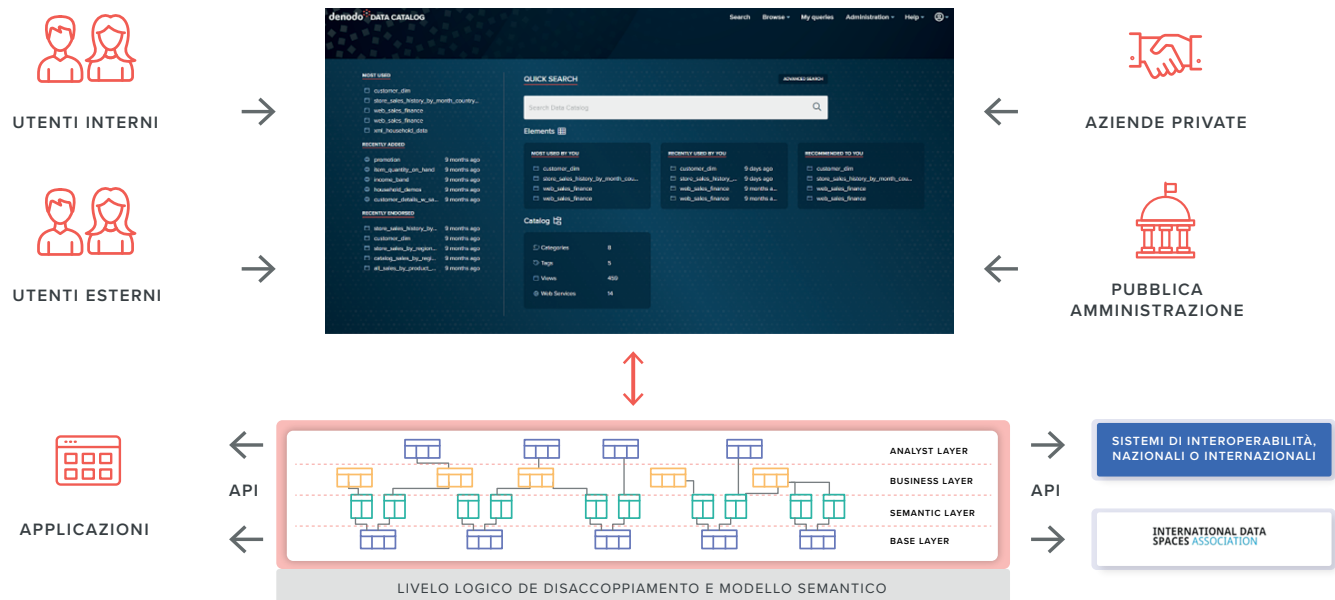
Il Data Catalog, un punto di accesso ai dati pensato specificatamente per i Data Consumer, non solo semplifica tutte le attività di esplorazione dei dati, ma rende possibile, in ogni momento, la connessione tra significato dei dati e loro occorrenze, cosa che rafforza sia la comprensione di ciò che i dati rappresentano e di quanto bene lo fanno, ma soprattutto la fiducia che in essi si può avere.

L'intelligenza artificiale, infine, opera anch'essa per semplificare la gestione, la comprensione e l'uso dei dati, assistendo e affiancando sia il Data Steward che il Data Consumer nelle loro attività quotidiane, perché – e vale sempre la pena ripeterlo – ogni trasformazione, che non sia affiancata da strumenti che ne rendano semplice l'adozione, è destinata al fallimento.

Pertanto, in ambienti complessi e multi-entità, l'integrazione logica dei dati non è solo un'opzione, ma è la soluzione definitiva per garantire che i dati siano condivisi in modo efficace, facilmente comprensibili, governati in modo sicuro e ampiamente utilizzati, così da essere conformi allo spirito dell'EDA ma, soprattutto, fare un deciso passo in avanti verso un'economia dei dati che sia etica, giusta e a beneficio di tutti.

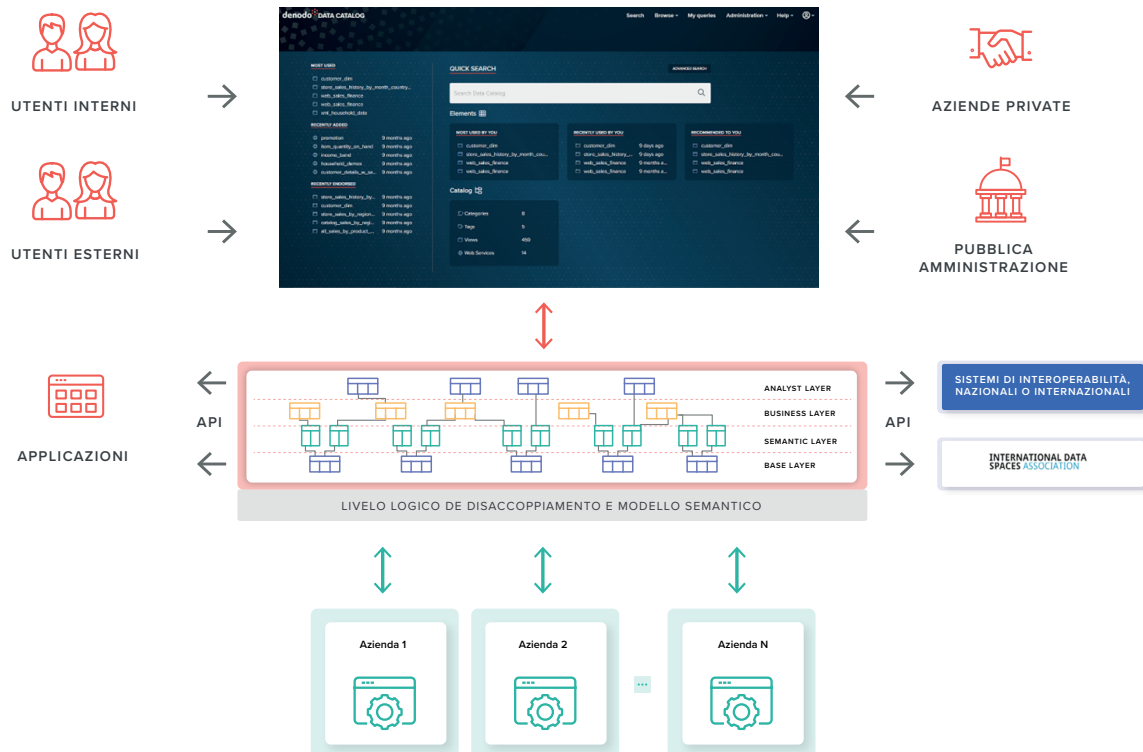
Va anche ricordato, uscendo per un momento dai confini dello EAD, che tutte le iniziative mirate a condividere dati di soggetti diversi, come ad esempio i Common European Data Spaces, peraltro menzionati nell'EDA, siano basati su un modello logico, dove viene centralizzato solamente il catalogo dei dati disponibili, mentre questi rimangono laddove sono originariamente presenti, per essere prelevati e consegnati solo quando richiesto.

Da questo punto di vista, pertanto, il ruolo di una architettura logica e, nello specifico, il suo modello semantico, diventa quello di livello di disaccoppiamento tra i potenziali fruitori dei dati, siano essi persone o applicazioni, e tutta l'infrastruttura tecnica sottostante, necessaria a gestire i dati che un soggetto, pubblico o privato, decide di mettere a disposizione dei cosiddetti portatori di interesse, con un evidente rimando al Data Governance Act, Nel prosieguo del documento saranno analizzati, con maggior dettaglio, gli aspetti e le componenti di un'architettura logica – e nello specifico della Denodo Platform – che più di ogni altro sono abilitatori essenziale per cogliere lo spirito del EDA, sia per quanto riguarda gli aspetti regolamentatori che quelli votati a creare una vera e propria economia dei dati, consapevole, governata e portatrice di valore per la comunità.



Il ruolo delle architetture logiche come snodo verso uno scenario transnazionale

In linea con questo spirito vocato alla condivisione si potrebbe pensare, come ulteriore passo in avanti, alla creazione di hub locali, il cui compito sarebbe quello di federare, sempre secondo lo spirito dell'EDA, i dati di tutte quelle aziende territoriali che, per vari motivi, non hanno le competenze o le risorse per operare da sole. Questo hub fungerebbe quindi da punto di smistamento verso un sistema più ampio, come rappresentato dalla figura seguente, che di fatto espande quanto già mostrato da quella precedente.



Il ruolo delle architetture logiche come nodo di instradamento locale di un sistema più ampio

CONCLUSIONI

Lo European Data Act (EDA) rappresenta un cambio di paradigma nella gestione, condivisione e valorizzazione dei dati all'interno dell'Unione Europea. Per le organizzazioni pubbliche e private, conformarsi al regolamento non significa solo rispettare obblighi normativi, ma anche cogliere una concreta opportunità di trasformazione digitale, collaborazione e innovazione.

Il documento ha evidenziato come l'adozione di una Logical Data Architecture, e in particolare della Denodo Platform, sia la scelta più efficace per affrontare le sfide poste dal regolamento. Grazie al suo approccio basato su data virtualization, modello semantico attivo e governance centralizzata, Denodo consente di semplificare l'accesso ai dati, garantendone al contempo qualità, sicurezza e tracciabilità.

Attraverso strumenti come il Data Catalog, la piattaforma rende i dati facilmente esplorabili e comprensibili da parte dei Data Consumer, favorendo una vera democratizzazione dei dati e una collaborazione intelligente tra attori diversi. L'integrazione fluida con tecnologie di intelligenza artificiale e il supporto a modelli come il Retrieval Augmented Generation (RAG) rafforzano ulteriormente la capacità delle organizzazioni di creare valore in modo etico, trasparente e sostenibile.

La Denodo Platform non solo consente la compliance al Data Act, ma diventa il motore di un'economia del dato più aperta, equa e interoperabile, dove il dato è finalmente un asset condiviso al servizio di tutti: aziende, cittadini, pubbliche amministrazioni e innovatori.

In conclusione, le architetture logiche come la Denodo Platform non solo rendono possibile la conformità al regolamento, ma offrono una base tecnologica solida per creare valore dai dati, in linea con una visione moderna, sostenibile e democratica dell'economia digitale.

Allegati

Le parti che seguono sono approfondimenti di quanto descritto nel documento con particolare riferimento alle già menzionate architetture logiche, che saranno approfondite in quelle loro componenti che, più di ogni altre, sono collegate al Data Act, nel senso di rappresentare funzionalità utili a far sì che quanto definito in termini organizzativi e culturali, trovi poi nella tecnologia un valido ed efficiente supporto.

COMPRENDERE LE ARCHITETTURE LOGICHE DEI DATI

A differenza delle architetture fisiche dei dati, principalmente basate sul ben noto paradigma ETL (Extract, Transform and Load) e che si basano su una duplicazione dei dati, trasformandoli e concentrandoli in un unico posto, come ad esempio un Data Warehouse o un Data Lake, le architetture logiche si occupano di come i dati sono collegati e di come possono essere utilizzati in modo efficiente indipendentemente dalla loro posizione fisica di archiviazione.

Le architetture logiche dei dati si riferiscono ai modelli concettuali, che definiscono il significato dei dati e il modo in cui questi sono logicamente organizzati, un'astrazione che garantisce flessibilità e scalabilità. Inoltre, il livello logico, all'interno del quale prende vita il modello semantico, funge anche da isolante rispetto alla tecnologia sottostante, i cui eventuali cambiamenti saranno assorbiti dal tale livello, consentendo di implementare un'architettura estremamente resiliente.

COMPONENTI DELLE ARCHITETTURE LOGICHE DEI DATI



Integrazione: il processo di combinazione di dati provenienti da fonti diverse per fornire una visione unificata.



Astrazione: creare un livello che nasconda le caratteristiche tecniche dei dati, concentrandosi invece sul modo in cui i dati vengono rappresentati, compresi e utilizzati.



Governance: stabilire politiche e procedure per gestire la qualità, la sicurezza e la conformità dei dati.



Modellazione: progettare modelli semantici che rappresentino le entità dei dati, le relazioni e i vincoli, in modo da creare la miglior concettualizzazione possibile del contesto di riferimento dell'amministrazione e, di conseguenza, progettare, erogare e misurare i servizi forniti.

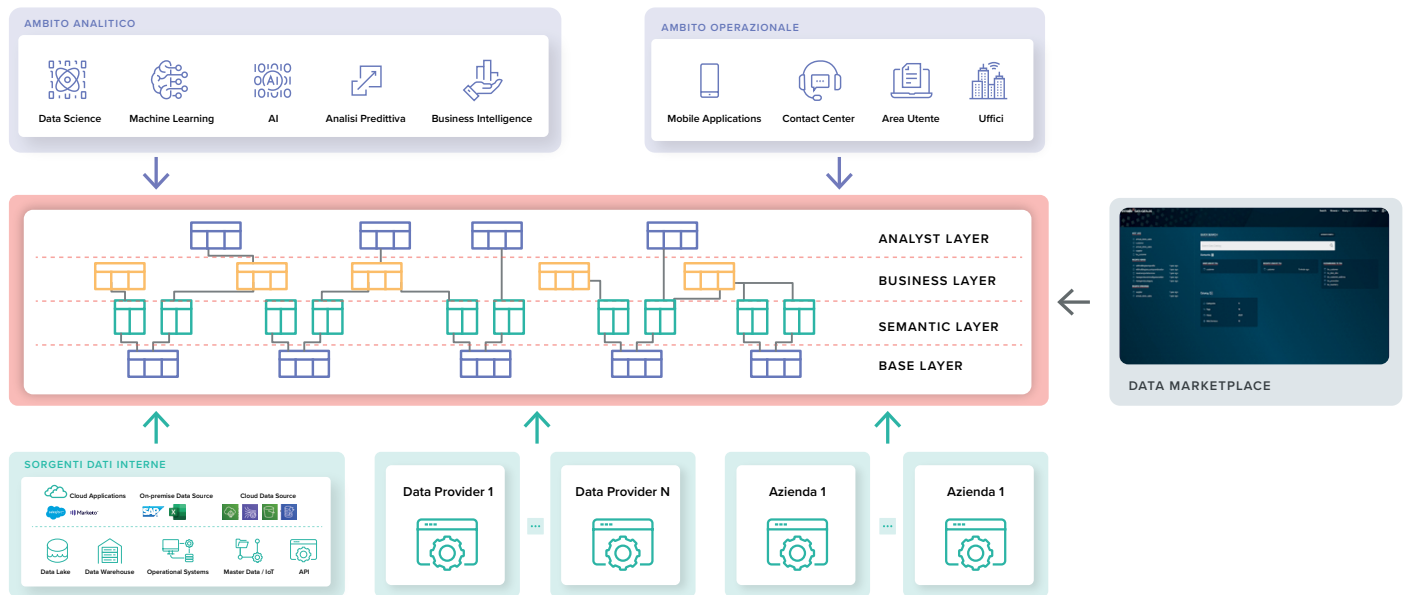


Accesso, consultazione e consumo: rendere effettiva la democratizzazione dei dati attraverso un Data Catalog o Data Marketplace, specificatamente rivolto ai Data Consumer, tipicamente competenti su ciò che i dati rappresentano e non sulle loro caratteristiche tecniche, in modo che questi possano facilmente esplorare il patrimonio informativo disponibile, comprenderlo e usarlo, senza doversi preoccupare degli aspetti tecnici dell'integrazione e della consegna di tali dati.

I modelli logici o semantici consentono quindi di dare concretezza alla concettualizzazione del contesto di riferimento, modellando i dati di interesse con un unico formalismo di rappresentazione, indipendentemente dalle caratteristiche tecniche dei dati, dalla loro distribuzione in rete e dal soggetto che li detiene e ne è responsabile, in modo da poter sfruttare dati di terze parti, come altre amministrazioni, aziende o data provider, senza dover duplicare, copiandoli, i dati da loro gestiti, con evidenti benefici sulla razionalizzazione complessiva e sulla loro governance.

Allo stesso tempo, il livello logico rappresenta un elemento di disaccoppiamento tra chi utilizza i dati – il Data Consumer – e l'infrastruttura tecnica necessaria per gestire tali dati. Questo disaccoppiamento diventa essenziale per gestire le modifiche a quest'ultima, poiché ogni cambiamento in questo senso sarà gestito dal livello logico, senza alcun riverbero sugli utenti, che continueranno a consumare i dati attraverso il modello semantico.

Graficamente, questo duplice ruolo del livello logico – rappresentazione del significato e disaccoppiamento della tecnologia – è rappresentato dalla figura seguente:



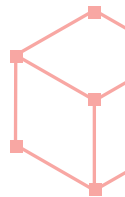
Panoramica dell'architettura di alto livello per l'integrazione logica dei dati

COME LE ARCHITETTURE LOGICHE DEI DATI SEMPLIFICANO LA GOVERNANCE DEI DATI

Le LDA migliorano significativamente la governance dei dati, consentendo la definizione di regole di sicurezza, visibilità e mascheramento dei dati esclusivamente a livello di astrazione logica. Questo approccio comporta notevoli vantaggi, in particolare la possibilità di centralizzare la gestione delle regole in un unico punto di controllo logico. In questo modo, le organizzazioni possono implementare sofisticate politiche di governance dei dati che si adattano dinamicamente ai ruoli e ai profili degli utenti, garantendo un accesso coerente e appropriato ai dati in tutta l'organizzazione.

Un vantaggio fondamentale della gestione logica della governance è la completa indipendenza dalle specifiche fisiche e tecniche delle fonti di dati sottostanti. Di conseguenza, se un'organizzazione intraprende cambiamenti tecnici, come la migrazione delle fonti di dati da ambienti on-premise a piattaforme cloud, le regole di governance stabilite rimangono inalterate. Poiché queste regole sono disaccoppiate dal livello fisico di archiviazione e sono definite in modo logico, si applicano immediatamente alla nuova fonte di dati senza richiedere modifiche o un'ampia convalida.

Le LDA offrono quindi soluzioni di governance robuste, flessibili e a prova di futuro, semplificando in modo significativo la gestione dei dati e migliorando l'agilità organizzativa e la conformità.



LA DEMOCRATIZZAZIONE DEI DATI COME ELEMENTO CHIAVE PER LA DATA DRIVEN TRANSFORMATION

Per diventare ed essere realmente data-driven, un'azienda deve necessariamente abbracciare la democratizzazione dei dati, il principio secondo cui i dati devono essere accessibili, comprensibili e utilizzabili da tutti i membri dell'organizzazione. Questo concetto va oltre la semplice concessione dell'accesso, affermando che l'utilizzo dei dati è un diritto, non un privilegio. Se l'uso dei dati è complesso o limitato, questi diritti rischiano di diventare dominio di pochi eletti, minando le fondamenta stesse di una cultura data-driven.

La democratizzazione implica che tutti – non solo gli analisti o il management – devono essere messi in grado di utilizzare i dati per l'uso che intendono farne. Questa accessibilità deve essere accompagnata dalla semplicità: strumenti intuitivi, governance chiara e supporto all'uso, che riducano le barriere all'ingresso e incoraggino un utilizzo dei dati diffuso.

Tuttavia, ai diritti devono corrispondere le responsabilità e deve essere altrettanto facile per gli individui adempiere ai propri doveri, senza che questi diventino un ostacolo all'uso dei dati, ma ne siano solamente strumento per un uso responsabile. Ma soprattutto, un'azienda veramente data-driven è una comunità, dove i dati non vengono solo consumati ma anche condivisi, arricchendo l'intelligenza collettiva dell'organizzazione.

La circolazione dei dati – la libertà che questi possano fluire all'interno dell'organizzazione – è essenziale per sbloccare il loro potenziale e quanto intuito o scoperto in un'area organizzativa può e deve alimentare o suggerire opportunità in altre aree e ciò tendenzialmente accade solo se i dati circolano in modo libero e responsabile. In questo contesto, la tecnologia svolge un ruolo abilitante fondamentale e un'architettura logica dei dati ben progettata consente di:

- Avere un accesso facile e sicuro ai dati rilevanti;
- Definire e applicare chiaramente e semplicemente diritti, doveri e responsabilità;
- Garantire un supporto scalabile per la collaborazione e la condivisione dei dati.

La democratizzazione dei dati non è solo un obiettivo tecnico o procedurale. È una pietra miliare culturale per le aziende che cercano di trasformare i dati in un patrimonio condiviso e in un motore di progresso collettivo.

L'IMPORTANZA DI UN MODELLO SEMANTICO

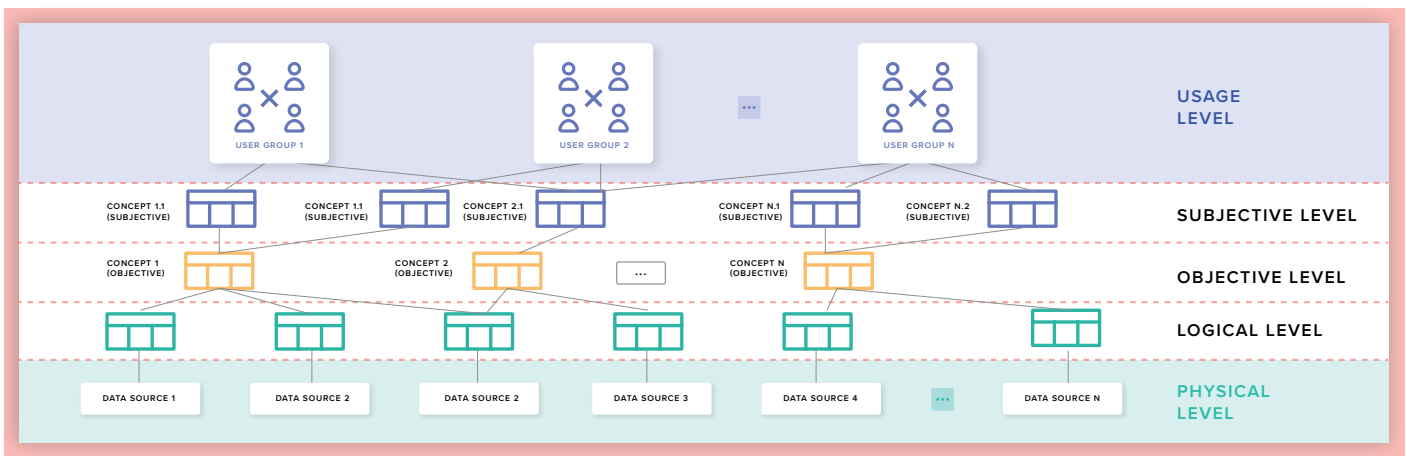
Il livello logico, che come già detto ospita i modelli semantici, è uno degli elementi più importanti, se non il più importante in assoluto, di una architettura logica, in quanto rappresenta il significato dei dati, nei termini di cosa rappresentano, di come lo fanno e di quali sono le relazioni reciproche.

La possibilità di realizzare tali modelli è fondamentale per rendere agevole a chi deve usare i dati – i Data Consumer – la loro ricerca e comprensione, ovviamente nel pieno rispetto delle regole che ne governano visibilità e uso.

Un modello semantico, inoltre, funge da collante semantico tra dati che appartengono a soggetti differenti, in quanto diventa il punto di armonizzazione di dati che, pur rappresentando lo stesso concetto, potrebbero farlo in modo differente, in virtù della diversa interpretazione che ne dà il soggetto che ne ha curato la genesi, interpretazione che dipende dal ruolo e dalle responsabilità del soggetto stesso.

Considerando che oggi praticamente ogni settore di mercato si trova a gestire processi caratterizzati da una forte condivisione di dati, posseduti e gestiti da soggetti diversi, con ruoli e responsabilità differenti, la possibilità – o meglio, la necessità – di avere questo punto di armonizzazione diventa fondamentale, poiché è lì che si sprigiona il potenziale insito nei dati, che può essere espresso solo all'interno di una rappresentazione che, da un lato, sia coerente, completa e univoca e, dall'altro, sia comunque in grado di rispettare quella soggettività rappresentativa necessaria a un'organizzazione per poter efficacemente gestire il proprio modello di business.

Volendo dare una rappresentazione di quanto detto, la figura seguente schematizza un modello semantico che rappresenta sia la componente oggettiva che quella soggettiva.



Una esemplificazione del modello semantico

La figura identifica diversi livelli, secondo un percorso di astrazione crescente:

- il **livello fisico** rappresenta le sorgenti dati oggetto di integrazione;
- il **livello logico** è dove si effettua la separazione logico-fisica delle sorgenti dati, elemento fondante delle architetture logiche;
- il **livello oggettivo** è dove vengono modellati quei concetti in modo tale che la loro rappresentazione sia considerata oggettiva all'interno dell'organizzazione, cosa che normalmente si traduce nel definire un concetto caratterizzandolo solo con quegli attributi che hanno un significato condiviso tra tutte le strutture dell'organizzazione;
- il **livello soggettivo** rappresenta la derivazione, a partire dal livello oggettivo, dei dati che catturano lo specifico punto di vista dei diversi Data Consumer, punto di vista che dipenderà dal loro particolare ruolo e dall'uso che ne devono fare in relazione all'obiettivo che stanno perseguendo;
- il **livello dell'uso**, infine, rappresenta i diversi Data Consumer che usano i dati, tipicamente usando quelli definiti secondo il loro punto di vista, che ragionevolmente meglio si adattano a ciò che devono fare.

Naturalmente, quelli che nella figura sono descritti genericamente come concetti, assumeranno una specifica caratterizzazione quando il modello sarà realizzato all'interno di una data organizzazione e con fini ben determinati, potendo ad esempio rappresentare il cliente, il cittadino, il punto vendita, il prodotto o l'impianto di produzione. Allo stesso modo, i gruppi di utenti potranno corrispondere, ad esempio, alle Vendite, al Customer Care, al Marketing, all'Amministrazione o alla Progettazione.

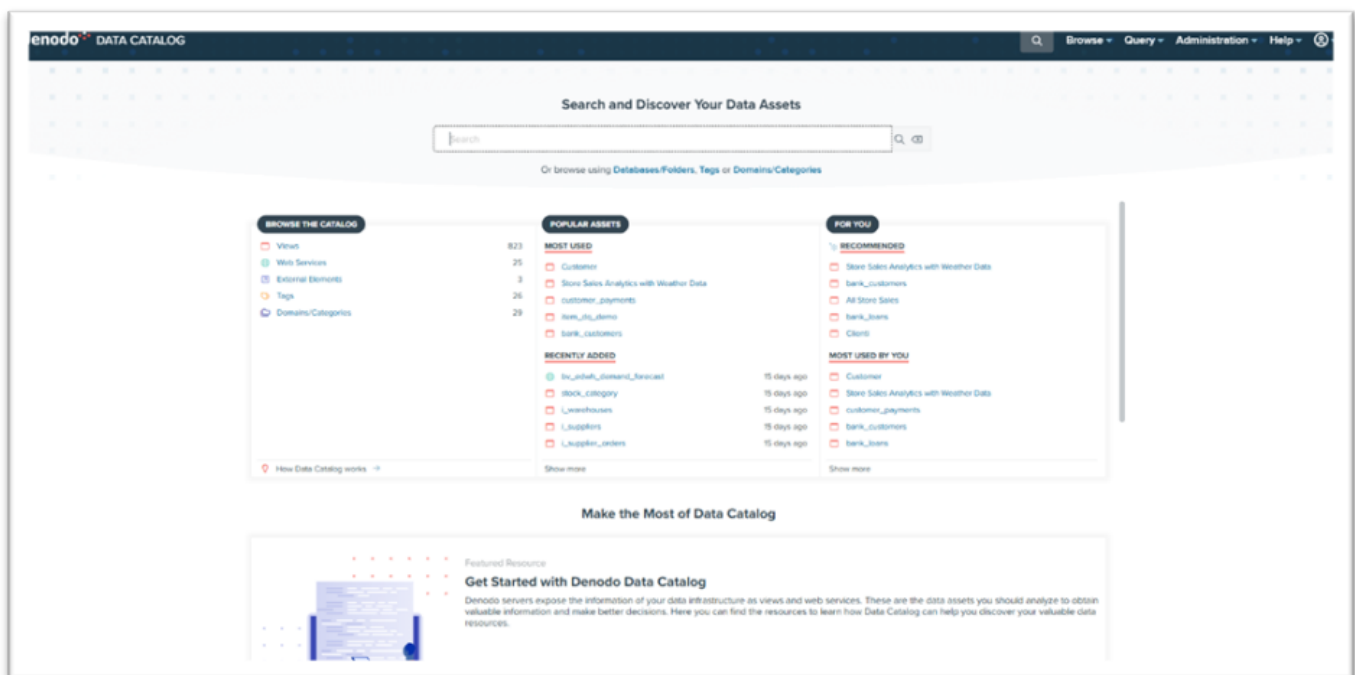
Va anche sottolineato che, sempre per non rendere troppo complessa la figura, si è deciso di evidenziare, da parte dei diversi utenti, solamente l'uso dei concetti del livello soggettivo, cosa che ovviamente non è un vincolo, dato che in situazioni reali ogni utente potrà usare uno qualsiasi dei concetti presenti nel modello semantico, indipendentemente dal livello nel quale è collocato.

L'IMPORTANZA DI UN CATALOGO DATI

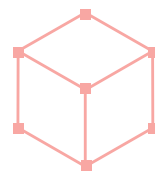
In un mondo sempre più orientato e fondato sui dati, è essenziale disporre di un catalogo di dati specificamente concepito per i Data Consumer. Tale catalogo semplifica il processo di esplorazione, comprensione e utilizzo efficace dei dati, astraendo dalla complessità tecnica sottostante. Offrendo un'interfaccia intuitiva e facile da usare che non richiede competenze specifiche, il catalogo dei dati elimina le barriere tradizionali che spesso ostacolano il processo di democratizzazione dei dati.

Questa facilità d'uso è fondamentale per promuovere una cultura orientata ai dati in cui tutti gli utenti, indipendentemente dalla loro formazione tecnica, possano accedere autonomamente ai dati rilevanti per il loro ruolo e sfruttarli al meglio. Quando i Data Consumer hanno la possibilità di interagire in modo indipendente con i dati, la dipendenza da un Team che gestisce centralmente i dati diminuisce, liberandolo e consentendogli di concentrarsi su iniziative strategiche di maggior valore. Inoltre, gli utenti che possono accedere autonomamente ai dati e gestirli responsabilmente sono maggiormente propensi ad essere partecipanti attivi, informati e responsabili nella più ampia comunità di dati dell'organizzazione.

In definitiva, un catalogo dei dati accessibile e orientato al consumatore è una pietra miliare per le organizzazioni che mirano a democratizzare i dati in modo efficace. Assicura che ogni consumatore di dati possa adempiere alle proprie responsabilità in modo efficiente e sicuro, contribuendo positivamente all'intelligenza collettiva e alla competitività dell'intera organizzazione.



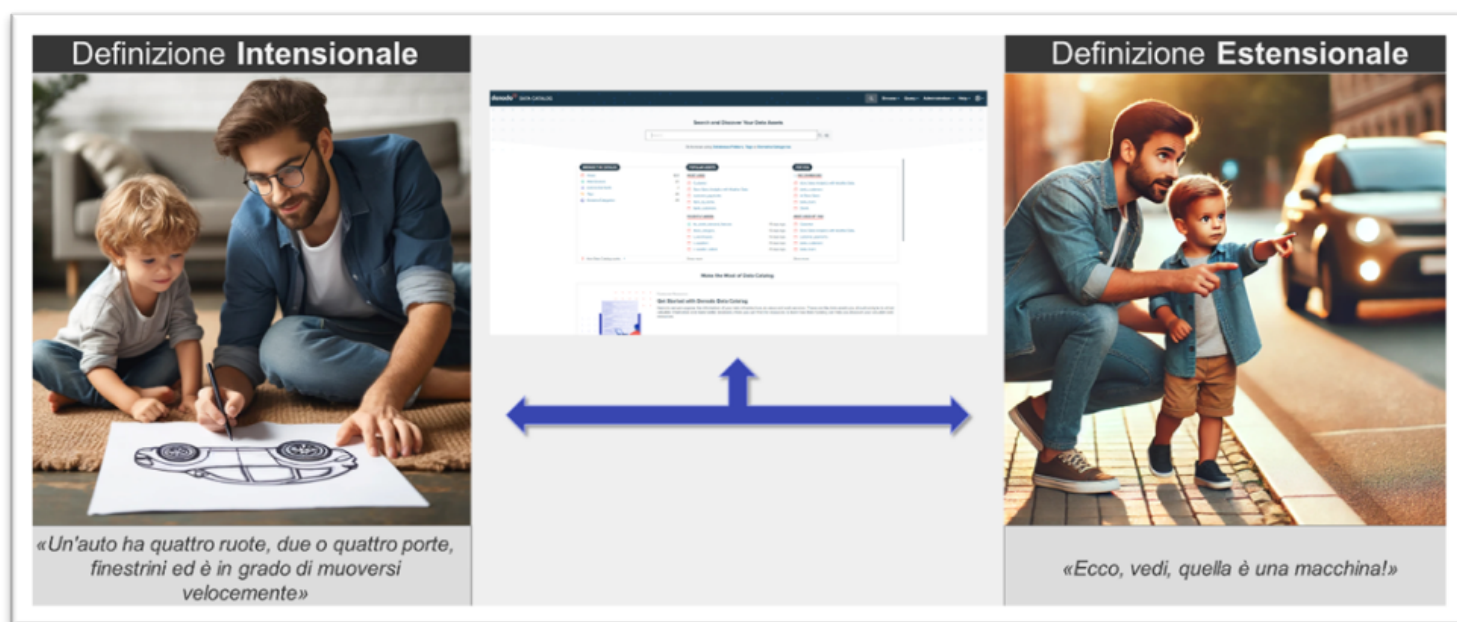
Il Catalogo dei dati, parte integrante della Piattaforma Denodo



RENDERE ATTIVE LE ONTOLOGIE CON LE ARCHITETTURE LOGICHE DEI DATI

Il modello semantico, cuore pulsante delle architetture logiche per l'integrazione dei dati, non solo mette al riparo da potenziali incoerenze nella definizione dei dati, riducendo – se non eliminando – il rischio di avere dati con gli stessi nomi che rappresentano concetti diversi, ma è anche lo strumento ideale, grazie al modo in cui il modello viene gestito in un contesto logico, per rendere viva un'ontologia, potendo, da un lato, importarla se è già stata definita altrove, in modo che il modello semantico sia automaticamente popolato dai concetti precedentemente definiti, e dall'altro collegarla alle fonti di dati che rappresentano la realizzazione di questi concetti. Questa connessione sarà responsabilità del Data Steward che, grazie alle sue competenze, sarà in grado di definire le cosiddette “viste di interfaccia”, arricchendo ciò che è rappresentato dall'ontologia con i dati che ne costituiscono la realizzazione.

In questo modo sarà possibile passare dalla componente logica (il significato dei dati) a quella fisica (le loro occorrenze) per fornire a chi deve utilizzare i dati la possibilità di valutarne, in fase di ricerca ed esplorazione, sia la correttezza della rappresentazione che la qualità dei dati raccolti, rafforzando così la fiducia che si può riporre in tali dati.



La connessione tra significato e occorrenze realizzata dal Data Catalog

In conclusione, le architetture logiche, con il loro concetto di modello semantico attivo, sono il ponte tra la componente intensionale dei dati e quella estensionale, una connessione continua e armoniosa che non richiede duplicazioni o spostamenti di dati.

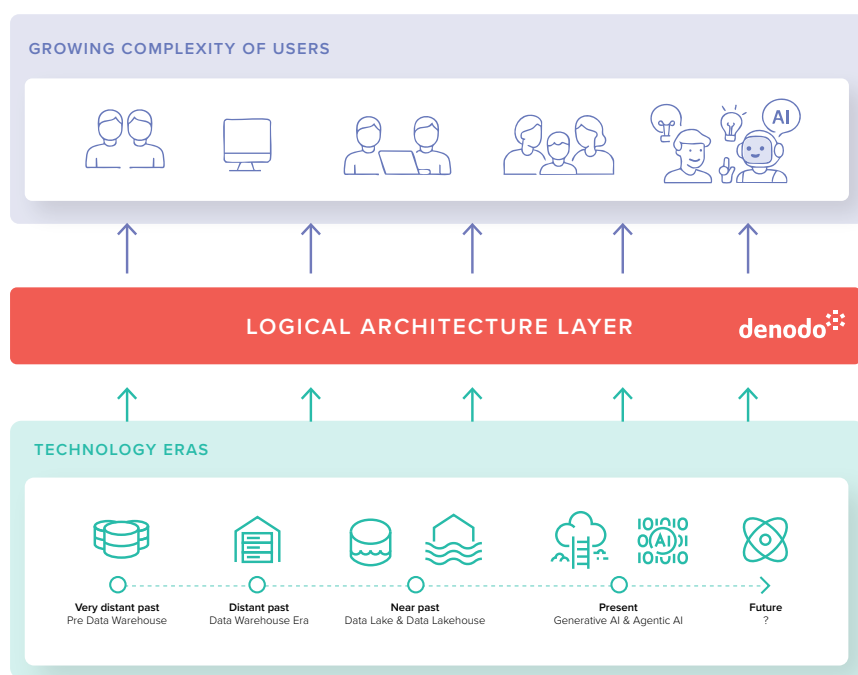
LA SEMANTICA NON BASTA, SERVE ANCHE LA RESILIENZA

Se il ruolo dei dati è tutto sommato invariante nel tempo, visto che il loro compito è quello di essere lo strumento per una concettualizzazione del mondo osservato, sia esso quello con cui un'azienda si relaziona o quello interno, che definisce il modo in cui essa si organizza e funziona, lo stesso non si può dire della tecnologia che ne permette la gestione, decisamente più mutevole, probabilmente in accelerazione e soggetta a salti quantici in corrispondenza di quelle innovazioni che possono essere considerate rivoluzionarie.

Una tecnologia soggetta a cambiamenti, a volte entropici, se da un lato permette di abilitare l'innovazione e il progresso, dall'altro crea inevitabili perturbazioni su chi deve utilizzare tale tecnologia, dove ogni cambiamento, ogni evoluzione, se non ben gestita, si riverbera inevitabilmente sugli utilizzatori di tale tecnologia, rallentando il loro lavoro e costringendoli a un continuo percorso di adattamento e formazione, con potenziali impatti negativi sull'efficienza complessiva e sulla capacità di governare quel tempo che, come detto all'inizio del documento, è ormai la nuova unità di misura.

Questo disturbo indotto dai cambiamenti tecnologici è efficacemente ridotto, se non addirittura eliminato, dalle architetture logiche per l'integrazione dei dati, che disaccoppiano l'accesso ai dati dai livelli fisici di storage e tecnologici sottostanti, creando un'astrazione flessibile che migliora significativamente la resilienza ai cambiamenti tecnologici. A differenza degli approcci tradizionali strettamente accoppiati, le architetture logiche garantiscono che gli aggiornamenti, le migrazioni o addirittura le sostituzioni delle tecnologie sottostanti avvengano in modo trasparente, senza interrompere l'esperienza del Data Consumer. Di conseguenza, gli utenti continuano ad accedere ai dati e a utilizzarli senza soluzione di continuità, al riparo da qualsiasi cambiamento, il che consente alle organizzazioni di evolvere tecnologicamente senza influire sulle operazioni quotidiane o sui flussi di lavoro degli utenti.

La figura seguente riassume graficamente la capacità delle architetture logiche di gestire un mondo che si muove a due velocità, dove all'invarianza delle motivazioni che guidano l'utilizzo dei dati ha corrisposto una significativa evoluzione tecnologica, che ha introdotto nuovi paradigmi per gestire uno scenario in cui i dati continuano a crescere in volume, eterogeneità e velocità.



Come un livello logico isola l'utilizzo dei dati dalla tecnologia di backend

IL RUOLO DELL'INTELLIGENZA ARTIFICIALE

Non va dimenticato che ogni algoritmo di intelligenza artificiale, tradizionale o generativo, utilizza i dati esattamente come fa ogni altro Data Consumer nello svolgimento delle sue attività, per cui anche in questo caso le architetture logiche sono lo strumento ideale, centralizzando la rappresentazione dei dati in modo che, da un lato, sia agevole costruire i dataset necessari agli algoritmi di intelligenza artificiale e, dall'altro, prelevando e consegnando i dati solo quando richiesto.

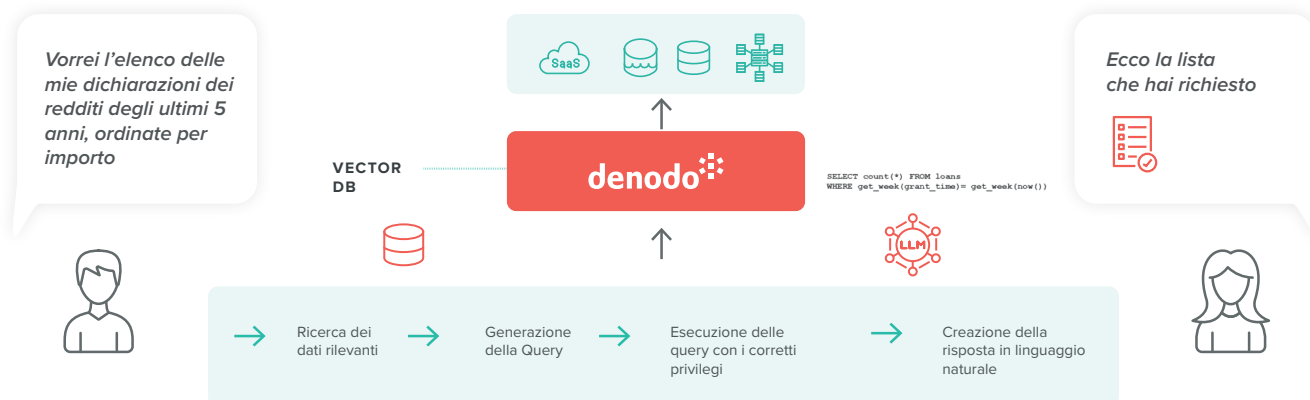
RETRIEVAL AUGMENTED GENERATION (RAG)

L'IA generativa riveste un ruolo sempre più centrale per la valorizzazione delle informazioni di natura testuale, rappresentando la base per assistenti virtuali di nuova generazione, in grado di sostenere conversazioni articolate, fornendo così a clienti, cittadini, collaboratori e dipendenti informazioni dettagliate, sia per un uso immediato, che per fornire più rapidamente le risposte.

Tuttavia, affinché questi assistenti virtuali siano in grado di rispondere in modo preciso e contestualizzato, è necessaria un'integrazione tra i Large Language Model (LLM) e le fonti di dati strutturati e non strutturati, necessari sia alla corretta comprensione del contesto nel quale le domande sono poste, che a reperire i dati che rappresentano l'oggetto della richiesta, dati che saranno utilizzati dai LLM per comporre la risposta.

Affinché un sistema di RAG possa associare correttamente gli elementi linguistici di una query in linguaggio naturale ai dati sottostanti, deve comprendere non solo i valori dei dati, ma anche il loro significato e il contesto al quale si riferiscono. Questa comprensione si basa su metadati ricchi e ben strutturati, che definiscono, con una terminologia vicina al business, le relazioni, le gerarchie e il significato dei dati.

Le Logical Data Architectures, in particolare quelle costruite sulla Denodo Platform, sono ideali per fornire questi metadati attraverso i loro modelli semantici, che offrono una visione astratta dei dati, orientata al business e indipendente dalle fonti fisiche, costituendo di fatto un ponte tra il linguaggio naturale e i dati strutturati. Di conseguenza, consentono ai sistemi RAG di interpretare con maggiore precisione l'intento dell'utente e di recuperare o generare risposte in linea con le aspettative dell'utente e con la vera semantica del panorama dei dati aziendali.



Un esempio di come il RAG consenta di rispondere a domande complesse

L'INTELLIGENZA ARTIFICIALE PER IL DATA MANAGEMENT

L'intelligenza artificiale, sia tradizionale che generativa, offre vantaggi sostanziali alla gestione dei dati. I metodi tradizionali di intelligenza artificiale, come gli algoritmi di apprendimento automatico, semplificano i processi automatizzando la classificazione, la convalida e la garanzia di qualità dei dati, alleggerendo notevolmente il carico di lavoro dei Data Steward. L'IA generativa espande queste capacità consentendo interazioni in linguaggio naturale e sintesi dei dati, facilitando notevolmente la scoperta, la comprensione e la fruibilità dei dati da parte dei Data Consumer.

GESTIONE SEMPLIFICATA DEI DATI

Le tecnologie di intelligenza artificiale automatizzano le attività di gestione, tra cui l'estrazione dei metadati, il monitoraggio della qualità dei dati e il rilevamento delle anomalie. Per i Data Steward, questa automazione riduce l'impegno manuale, assicura la coerenza delle pratiche di gestione dei dati e migliora l'accuratezza delle attività di governance. Il risultato è un approccio più efficiente, affidabile e agile alla gestione di ecosistemi di dati ampi e diversificati.

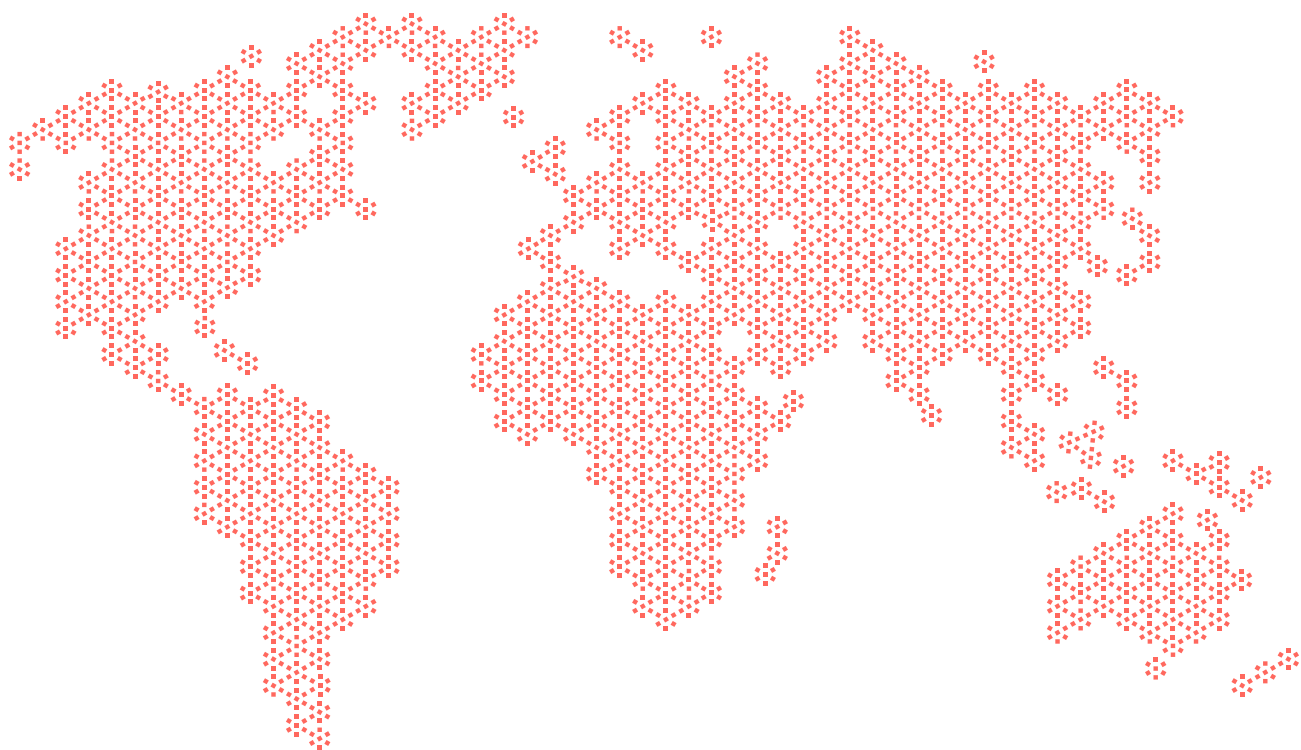
Con specifico riferimento alla piattaforma Denodo, il Denodo AI Assistant offre supporto, tra l'altro, per:

- Ottimizzazione delle query. La piattaforma Denodo utilizza da tempo l'intelligenza artificiale per l'ottimizzazione delle query, tra cui l'ordine di join, le associazioni e le aggregazioni. Ciò contribuisce a garantire prestazioni ottimali con costi ridotti, riducendo la richiesta di competenze specifiche, necessarie quando tale lavoro viene fatto manualmente.
- Preparazione e arricchimento dei dati. La piattaforma Denodo dispone di funzionalità di preparazione dei dati che sfruttano le potenzialità dell'intelligenza artificiale. Ad esempio, al Data Steward che sta lavorando su uno specifico insieme di dati, viene presentata una serie di trasformazioni consigliate in base ai suoi comportamenti precedenti e di quelli di utenti simili che hanno utilizzato lo stesso set di dati o altri dati simili, una capacità che migliora con l'uso, dato che la piattaforma Denodo apprende e aggiorna i metadati sulla base del loro effettivo uso.

MIGLIORAMENTO DELL'ACCESSIBILITÀ DEI DATI E DELL'ESPERIENZA UTENTE

L'intelligenza artificiale generativa consente di creare interfacce intuitive e di facile utilizzo, facilitando notevolmente la ricerca e l'interpretazione di dati complessi da parte dei consumatori di dati. Traduce i metadati tecnici in approfondimenti comprensibili, consentendo agli utenti privi di competenze tecniche approfondite di sfruttare i dati in modo efficace, accelerando così il processo decisionale informato in tutte le funzioni aziendali.

Con specifico riferimento alla piattaforma Denodo, il Denodo AI Assistant consente a ogni utente di formulare una domanda in linguaggio naturale, che viene poi convertita in codice SQL ed eseguita dalla piattaforma. Si tratta di un grande passo avanti nella democratizzazione dei dati, che consente ai Data Consumer di esplorare, comprendere e usare i dati con facilità.



Denodo è leader nella gestione dei dati. La pluripremiata Denodo Platform è la piattaforma leader per l'integrazione, la gestione e la distribuzione dei dati che utilizza un approccio logico per una business intelligence self-service, data science, integrazione dei dati ibrida/multi-cloud e data service aziendali. Con un ritorno sull'investimento superiore al 400% e profitti di milioni di dollari, i clienti di Denodo di grandi e medie imprese in oltre 30 settori hanno ammortizzato l'investimento in meno di sei mesi.

Visita www.denodo.com | E-mail info@denodo.com | Scopri community.denodo.com

