

# I vantaggi delle architetture logiche dei dati per la Pubblica Amministrazione

---

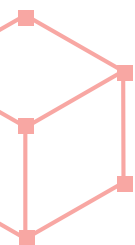
Un abilitatore per una interoperabilità sicura e per il “Once Only”





## Indice

|                                                             |    |
|-------------------------------------------------------------|----|
| Introduzione                                                | 5  |
| I dati sono una fonte energetica preziosa e inesauribile    | 6  |
| Le peculiarità del settore pubblico                         | 8  |
| Perché serve rappresentare un cittadino in modo nuovo       | 8  |
| Verso un Augmented Citizen View                             | 9  |
| Comprendere le architetture logiche dei dati                | 10 |
| Le architetture logiche in uno scenario <i>Augmented</i>    | 12 |
| Il ruolo delle architetture logiche nella PDND              | 13 |
| Il ruolo dell'intelligenza artificiale nel settore pubblico | 17 |
| Due storie di successo                                      | 18 |
| Conclusioni                                                 | 21 |



# Introduzione

Carl Shapiro e Hal R. Varian, nel loro libro del 1998 *“Information Rules: A Strategic Guide to the Network Economy”*, ci ricordano che *“Technology changes, but the laws of economics don’t!”* e, andando molto indietro nel tempo, Virgilio, in uno dei versi delle Georgiche, recita *“Sed fugit interea fugit irreparabile tempus”*<sup>1</sup> o, più semplicemente, nella forma che poi si è tramandata nel tempo, *“Tempus fugit”*.

Se i fondamentali dell’economia non cambiano, ciò che cambia – e lo fa sempre più velocemente – è il mondo con il quale le amministrazioni si confrontano, un mondo in continua accelerazione, spesso soggetto a eventi che perturbano la linearità del cambiamento, dove i fenomeni si propagano con la velocità di un’onda di marea, che per essere cavalcata necessita di capacità sempre nuove, che consentano di rimanere sulla sua cresta, senza esserne travolti.

Nulla di nuovo, potremmo dire, ricordando che già nel lontano 1987 l’U.S. Army War College aveva coniato l’acronimo VUCA – *Volatility, Uncertainty, Complexity, Ambiguity* – per sottolineare come, dopo il periodo di forte tensione dovuto alla guerra fredda, il mondo era ora in uno scenario fortemente entropico, che rendeva difficile la definizione di una strategia complessiva e che guidasse le decisioni politiche. Una situazione molto simile al mondo digitale con il quale le amministrazioni si confrontano quotidianamente, cercando di anticipare ciò che vi succede - perché agire è sempre meglio che reagire - cercando di definire una strategia il quanto più possibile a lungo termine, che guidi le decisioni che devono essere prese, sempre più rapidamente, perché ogni amministrazione, nel suo modo di essere e nei risultati che ottiene, si manifesta e si realizza, appunto, nelle decisioni che prende.

Il tempo è quindi la nuova unità di misura e pervade tutto ciò che un’amministrazione fa, o si trova costretta a fare, e il mondo con il quale l’amministrazione si confronta: l’auspicata velocità del successo di una iniziativa; quella di saper riconoscere una scelta sbagliata; la rapidità con la quale cambia il mercato; la velocità con la quale la notizia di una scelta sbagliata si propaga sulla rete; il desiderio di poter contrarre il tempo tra un investimento e il suo ritorno o tra un’idea e la sua realizzazione; la rapidità con la quale la concorrenza copia, rendendolo non più esclusivo, ciò che un’amministrazione ha ideato e messo sul mercato.

Ci si deve quindi confrontare quotidianamente con esso, cercando di governarlo, perché il tempo è prezioso e una sua gestione oculata ed efficiente è la chiave per governare i fondamentali economici ai quali si riferiscono Shapiro e Varian, tenendo anche presente che l’attesa non è consentita, perché come ancora una volta ci ricorda Seneca, *“Dum differtur, vita transcurrit”*<sup>2</sup>.

Più saranno in grado di sfruttare al meglio il (poco) tempo che le amministrazioni hanno a disposizione, tanto più sani saranno i macro-indicatori economici, perché la sfida è anche quella di andare contro alla saggezza popolare, quando ci dice che *“presto e bene, raro avviene”*, mostrando che, con le giuste scelte, si può invece fare bene e farlo rapidamente.

Una rapidità che, parlando di dati, deve pervaderne l’intero ciclo di vita, dalla loro rilevazione alla loro sintesi, affinché producano informazioni, consentano l’innesto di tali informazioni nel contesto dell’amministrazione, così che la conoscenza possa svilupparsi e arricchire quella saggezza alla base delle decisioni che vengono prese<sup>3</sup>, perché ogni amministrazione dovrebbe ambire a essere **fluida**, per adattarsi alle mutevoli forme che assume il contesto con il quale si confronta e alle altrettanto mutevoli esigenze e abitudini dei soggetti che lo vivono, e **non viscosa**, in modo che tale adattamento sia rapido, perché se così non fosse si rischierebbe di adattarsi a un contest che, nel frattempo, potrebbe già aver cambiato forma.

<sup>1</sup> *“Ma fugge intanto, fugge irreparabilmente il tempo.”*

<sup>2</sup> *“Mentre si rinvia, la vita passa.”*

<sup>3</sup> Qui il riferimento è alla ben nota piramide DIKW (Data, Information, Knowledge, Wisdom), la cui origine è considerata incerta, ma che rappresenta quella che normalmente viene indicata come la *gerarchia della conoscenza*.



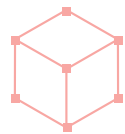
Fluidità e non viscosità, quindi, il cui raggiungimento dipende in modo significativo dai dati, che in un mondo digitale sono l'equivalente degli occhi in un mondo fisico e che, quindi, sono il viatico per cogliere ogni cambiamento, comprenderlo e reagire a esso.<sup>4</sup>

Qualsiasi tecnologia, pertanto, non deve mai di perdere di vista due elementi fondamentali, ontologici potremmo dire: essere un valido supporto affinché un'amministrazione sia in grado di essere economicamente sana e affiancare chi nell'amministrazione lavora, dandogli il giusto aiuto nel poter garantire quella fluidità e non viscosità nel reagire ai cambiamenti che inevitabilmente si verificheranno, nel prendere le decisioni che dovranno essere prese e nelle valutazioni che di queste se ne dovrà dare conto.

Ortogonalmente ai due elementi appena descritti, inoltre e non meno importante, il rispetto dell'ambiente perché, se da un lato la tecnologia è sempre più una commodity a basso costo, dall'altro questo non deve autorizzarne l'abuso, ricordando che ogni tecnologia consuma energia e, di conseguenza, non si può ignorare la sua *Carbon Footprint* e l'esigenza di ridurla quanto più possibile, secondo lo spirito di trovare l'equilibrio ideale tra esigenze e risorse necessarie per indirizzarle.

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.

## I dati sono una fonte energetica preziosa e inesauribile



Nel 2006, il matematico inglese Clive Humby coniò l'ormai celebre affermazione *"Data is the new oil"* e una decina d'anni dopo, nel 2017, la rivista The Economist affermò, sulla copertina del numero di maggio 2017, che i dati erano ormai *"The world's most valuable resource"*.

Due affermazioni forti, ma già allora perfettamente allineate con il ruolo che i dati stavano assumendo, un ruolo così centrale da essere appunto paragonabili a una nuova e potente fonte d'energia, così tanto rinnovabile e auto-rigenerativa<sup>5</sup> da lasciar presagire che, in futuro, questo valore non poteva che aumentare vertiginosamente, cosa che non solo s'è verificata, ma che probabilmente è andata ben oltre le più rosee previsioni, come evidenziato nel report **"Data Age – The Global Datasphere 2025"** di IDC and Seagate:

1. la *Global Datasphere* crescerà dai 33 zettabytes del 2018 ai 175 previsti nel 2025. Si stima che i soli sensori IoT produrranno, nel 2025, circa 90 zettabytes di dati;
2. nel 2025, il 49% di tutti i dati a livello mondiale risiederanno in ambienti Public Cloud, che di fatto diventerà il *New Core*;
3. nel 2025, circa il 30% dei dati a livello mondiale, coerentemente con la crescita prevista per l'*Edge Computing*, necessiteranno di una gestione in *realtime*.

Se questa immensa fonte di energia rappresenta indubbiamente un potenziale enorme per le amministrazioni, che attraverso i dati saranno sempre più in grado di osservare e comprendere cosa accade, dall'altro è anche un elemento di attenzione, per la difficoltà di gestirne le quantità e abilitarne un uso efficace ed efficiente, perché a un aumento della quantità e della tipologia dei dati, corrisponde parimenti il rischio di perdersi nella vastità di ciò che è disponibile e di rendere complessa la comprensione del significato di ciò che i dati rappresentano, come peraltro si evince un'analisi condotta da **MarketLogic**, che evidenzia come:

<sup>4</sup> Naturalmente, se osservazione e comprensione sono profondamente dipendenti dai dati, la capacità di reazione e adattamento sono una questione organizzativa e in parte culturale, la cui analisi esula però dagli obiettivi del documento.

<sup>5</sup> A differenza del petrolio, i dati non solo sono una fonte energetica praticamente infinita, ma anche in grado di auto-rigenerarsi, visto che le decisioni che sono prese in base ai dati producono un cambiamento nel mondo sul quale sono dirette, cambiamento che sarà rilevato da nuovi dati, creando quindi un circolo virtuoso e un conseguente e costante aumento della capacità della fonte energetica stessa.

- **I lavoratori non riescono a individuare i dati aziendali dei quali hanno bisogno** - Uno dei motivi principali per cui così tanti dati aziendali non vengono sfruttati è semplicemente perché le persone non riescono a trovarli. Una inefficienza che crea sofferenza sui budget aziendali.
- **I lavoratori non riescono a dare un senso ai dati aziendali abbastanza velocemente** - Anche se i lavoratori riescono a trovare le informazioni, potrebbero ignorarle se non riescono a trasformarle in informazioni velocemente utilizzabili. In altre parole, c'è un divario tra i flussi crescenti di dati e informazioni e la capacità delle amministrazioni di dare a esse un senso e, quindi, di trasformarle in valore aziendale.
- **Il mercato si muove troppo velocemente per una gestione disorganizzata delle informazioni** - In uno studio dello scorso anno, due terzi delle aziende britanniche hanno ammesso di aver commissionato ricerche di mercato che non hanno mai utilizzato perché rese irrilevanti dal rapido cambiamento delle condizioni economiche e sociali.

Altrettanto interessante, o forse preoccupante, come emerge da uno **studio del 2020 di Seagate e IDC**, è la capacità delle amministrazioni di utilizzare fattivamente questa immensa mole di dati, considerando la velocità con la quale questi sono generati e, in molti casi, l'altrettanta velocità con la quale diventano obsoleti, aspetto che si ipotizza possa peggiorare, visto che le previsioni future sono per una crescita considerevole dei dati in tempo reale<sup>6</sup>.

La capacità di *catturare* tutti i dati potenzialmente catturabili e la conseguente capacità di utilizzarli fattivamente per determinare cosa sta accadendo e quali decisioni e azioni vadano di conseguenza poste in essere è, purtroppo, ancora ampiamente migliorabili, come evidenzia lo studio del 2020 sopra menzionato, dal quale emerge che:

solo il 56% dei dati generati viene effettivamente catturato e, quindi, diventa parte del patrimonio informativo di un'amministrazione;

di questo 56%, solo il 57% dei dati viene realmente usato per condurre analisi, prendere decisioni e definire azioni.

Si ha quindi la netta sensazione - la certezza, potremmo dire - che per gestire al meglio questo enorme potenziale, serva un modo nuovo di gestire i dati, sia per quanto riguarda le risorse necessarie, ma soprattutto per rendere quanto più possibile semplice comprenderli e usarli, perché, se una cosa è difficile da farsi, allora molto probabilmente non verrà fatta.

## Le peculiarità del settore pubblico

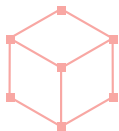
Nel settore pubblico, la capacità di integrare e analizzare in modo agile, veloce e semplice i dati provenienti da varie fonti è fondamentale, soprattutto in considerazione che ciò che a una prima e superficiale lettura è visto come un'unica entità è, al contrario, un insieme di soggetti distinti che operano per un fine comune.

Quest'operare in modo armonico, per consentire ai partecipanti alla vita pubblica, cittadini o imprese che siano, di poter esercitare i propri diritti e di farlo nel rispetto dei doveri, ha nei dati l'elemento fondante, soprattutto considerando che ciò che il partecipante è e fa richiede una rappresentazione necessariamente distribuita tra le diverse amministrazioni, ognuna delle quali responsabile di una parte parziale di tale rappresentazione.

Le architetture logiche dei dati forniscono un quadro strategico che consente di consolidare i dati provenienti dalle diverse sorgenti interne di un'amministrazione, da quelli detenuti dalle altre amministrazioni, da eventuali Data Provider, pubblici o privati, e dagli stessi cittadini e imprese: è proprio questa integrazione che consente di creare una visione completa e integrata di un cittadino o di un'impresa.

Nei paragrafi successivi, quindi, si detaglieranno meglio questi aspetti e – ma solo per semplicità narrativa – si

<sup>6</sup> Secondo uno **studio di IDC**, entro il 2025 quasi il 30% dei dati prodotti sarà costituito da dati in tempo reale.



farà riferimento al cittadino come elemento centrale, anche se quanto seguirà si applicherà, *mutatis mutandis*, agli altri soggetti con i quali un'amministrazione del Settore pubblico interagisce, come ad esempio il lavoratore, il pensionato, l'impresa, l'associazione no-profit o il professionista che opera da intermediario.

## Perché serve rappresentare un cittadino in modo nuovo

Ogni cittadino, in quanto essere umano, possiede alcune caratteristiche intrinseche che definiscono ciò che lui è, caratteristiche che si modificano nel corso della sua vita, dalla nascita fino al suo decesso. Lo stesso cittadino, tuttavia, sarà necessariamente collocato in un contesto specifico, connotato da elementi demografici, economici, sociali e culturali, che inevitabilmente plasmano il modo di essere e, soprattutto, che ne caratterizzano il modo in cui interagisce con la Pubblica amministrazione, nei termini dei servizi dei quali può beneficiare, dei diritti che può esercitare e dei doveri ai quali deve ottemperare.

Per una corretta gestione del cittadino, lungo tutto il suo ciclo di vita e nei diversi modi in cui questo interagisce con l'amministrazione, è quindi necessario considerare tutti questi elementi, che hanno un valore pari a quelli che caratterizzano il cittadino in quanto tale, esattamente lo stesso cittadino, in quanto essere umano, è plasmato dal contesto in cui cresce e vive.

Poiché nel mondo digitale ogni rappresentazione si nutre di dati, è necessario quindi definire i diversi livelli che consentono di estendere gradualmente tale rappresentazione, in modo che questa sia sempre più ricca e completa e consenta di generare quel valore che, da un lato il cittadino si aspetta in quanto parte di una comunità, dall'altro le amministrazioni ne hanno bisogno per poter fornire servizi in linea con le aspettative della comunità stessa.

Nei paragrafi che seguono il termine "Organization" va letto in senso generico, potendo indicare il generico soggetto che appartiene al Settore pubblico, come ad esempio un ente, un ministero o un'agenzia.



### ORGANIZATION LEVEL

Si tratta di dati sotto il controllo diretto dell'amministrazione, dati di cui l'amministrazione ha la proprietà e la responsabilità, indipendentemente dal fatto che si trovino on-premise o nel cloud.

### EXTENDED ORGANIZATION LEVEL

Si tratta di dati sotto il diretto controllo dell'amministrazione a cui si aggiungono dati forniti da Data Provider esterni, dati che dipendono dai servizi forniti da questi ultimi e che permettono di arricchire ulteriormente ciò che i dati interni consentono di rappresentare.

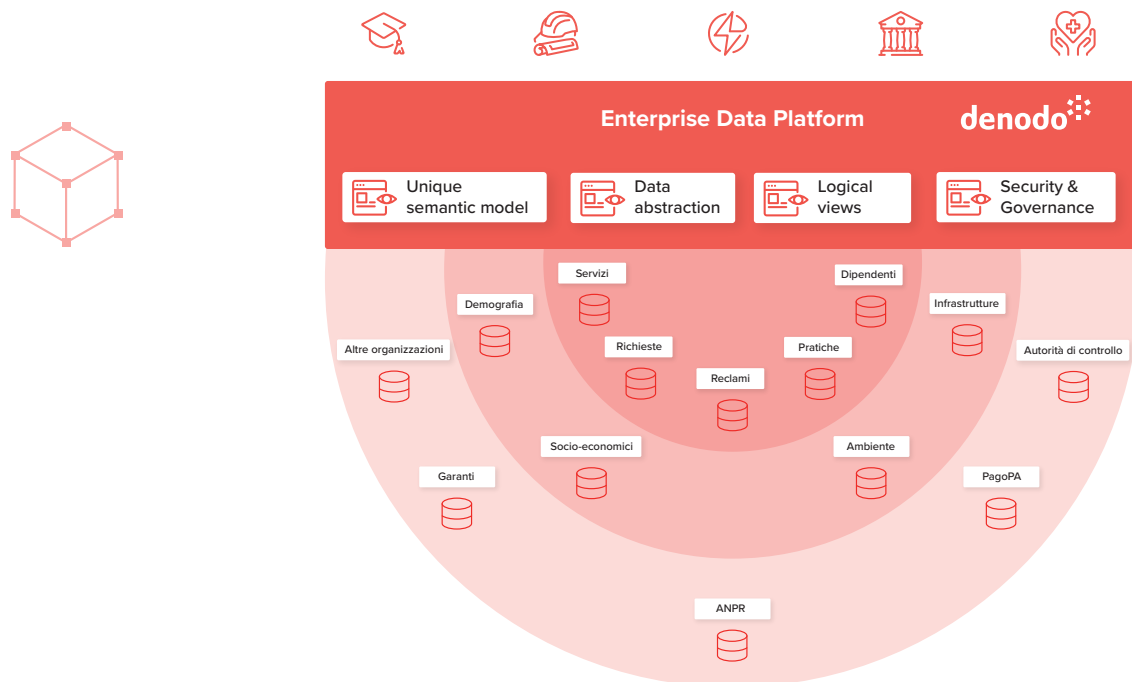
### AUGMENTED ORGANIZATION LEVEL

È composto da dati interni, dati forniti da provider di dati esterni, nonché da dati che l'amministrazione decide di condividere esternamente con altre amministrazioni o soggetti privati, così da creare un ecosistema i cui confini non sono più definiti strettamente dall'ambito legale che definisce l'amministrazione, ma da quello che determina il suo ruolo nella comunità, nei termini di cosa deve dare e cosa ricevere.



# Verso un Augmented Citizen View

In continuità con i livelli sopra descritti, che si riferiscono ai diversi “confini” di un’amministrazione rispetto ai dati che intende utilizzare, possiamo utilizzare lo stesso approccio, definendo livelli analoghi per la rappresentazione del cittadino, ottenendo quanto schematizzato dalla figura seguente, dove la rappresentazione del cittadino lungo tutto il suo ciclo di vita, e grazie a Denodo, può estendersi progressivamente, come mostrano i cerchi concentrici che rappresentano i tre diversi livelli e che consentono, in modo progressivo e laddove ritenuto necessario, di arricchire la rappresentazione del cittadino – ma come già detto, di qualsiasi concetto che sia ragionevolmente complesso – con tutti i dati che si ritiene utili o necessari per fornire un servizio, condurre analisi o prendere decisioni.



Augmented Citizen View

Naturalmente, questi livelli progressivamente più ricchi non sono un obbligo semantico – e non devono esserlo – ma un’opportunità da cogliere tutte le volte che lo si ritenga utile o necessario, secondo quel principio che ci dice che la rappresentazione di un concetto deve essere guidata dall’uso che di esso si vuole fare – strategico o tattico che sia – e non dalle loro specificità tecnologiche, che vanno ovviamente gestite, ma che non possono rappresentare un vincolo a ciò che si desidera fare.

Questa necessità di rappresentazione, inoltre, deve poter essere indirizzata in agilità, velocità e semplicità, non per un vizzo particolare, ma semplicemente perché viviamo in un mondo fluido, in continua accelerazione, che la velocità di reazione ai cambiamenti, necessariamente da doversi comprendere e, conseguentemente, gestiti attraverso le opportune decisioni, deve essere altrettanto rapida ed è qui che entrano in gioco le architetture logica per l’integrazione dei dati.



# Comprendere le architetture logiche dei dati

Le architetture logiche dei dati si riferiscono ai modelli concettuali, che definiscono il significato dei dati e il modo in cui i dati sono logicamente organizzati all'interno di un'amministrazione.

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.

Questa astrazione consente una maggiore flessibilità e scalabilità, garantendo l'integrazione e l'accesso ai dati da diverse fonti in modo agile e, 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 livello logico, consentendo di implementare un'architettura estremamente resiliente.

## COMPONENTI DELLE ARCHITETTURE LOGICHE DEI DATI

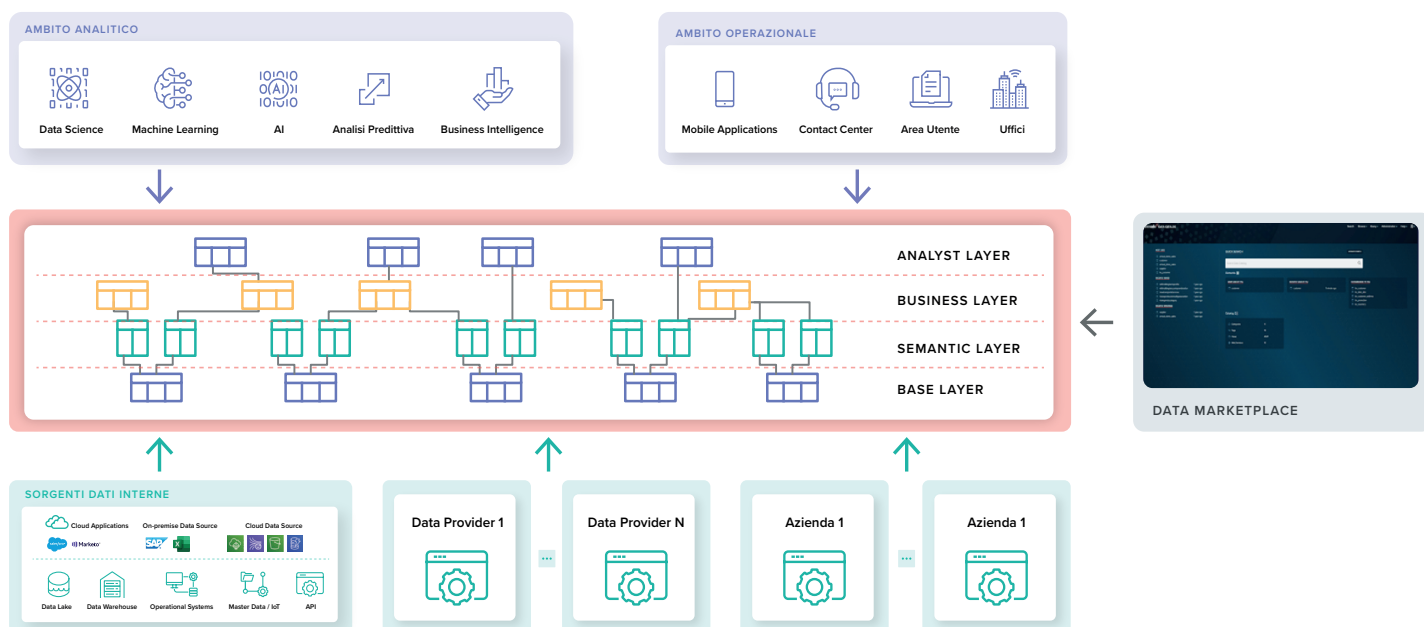
- **Integrazione:** il processo di combinazione di dati provenienti da fonti diverse per fornire una visione unificata.
- **Astrazione:** creare un livello che nasconda le caratteristiche tecniche dei dati, concentrandosi invece sul modo in cui i dati vengono rappresentati, compresi e utilizzati.
- **Governance:** 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.

Non deve poi essere dimenticato il tema della sostenibilità, per il quale le architetture logiche, non richiedendo la duplicazione dei dati e, soprattutto, non richiedendo di conseguenza il loro periodico aggiornamento, operazione tipica dei modelli tradizionali basati sul paradigma ETL, sono intrinsecamente parche nell'uso di risorse, sia per lo storage che computazionali, prelevando e consegnando i dati solo quando richiesto e non in modo preventivo, come accade nel modello ETL, dove i dati sono raccolti in un unico punto (Data Warehouse o Data Lake) a prescindere da quali poi saranno effettivamente utilizzati.

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

## L'IMPORTANZA DI UN MODELLO SEMANTICO NEL CONTESTO PUBBLICO

Il livello logico, che tra le altre cose ospita i modelli semantici<sup>7</sup>, 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, come negli esempi fatti per illustrare il concetto di *Augmented View*, 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. Ad esempio, la rappresentazione del concetto di “cittadino” ragionevolmente differirà tra la lettura che ne dà l'Agenzia delle Entrate e quella dell'INPS, semplicemente perché queste due amministrazioni lo considerano secondo ciò che fanno e i servizi che offrono.

Considerando che la Pubblica Amministrazione si compone di soggetti differenti, con differenti ruoli e responsabilità, la possibilità – o meglio, la necessità – di avere questo punto di armonizzazione diventa fondamentale, essendo lì che si libera il potenziale insito nei dati, che può esprimersi solo all'interno di una rappresentazione che, da un lato si coerente, completa e non ambigua e, dall'altro, sia in rado comunque di rispettare quella soggettività rappresentativa necessaria alle diverse amministrazioni affinché queste possano adempiere al loro mandato.

Non è un caso, infatti, che sul tema della rappresentazione sono in corso iniziative di ampio respiro, come ad esempio il progetto **OntoPiA**, una rete di ontologie organizzate in livelli e che, coerentemente con lo spirito di ogni livello, definiscono e rappresentano i principali concetti, come schematizzato dalla figura seguente.

<sup>7</sup> Denodo non pone particolari limiti alla numerosità dei modelli semantici, potendo sceglierne se costruirne uno solo o più di uno, a seconda di come si preferisce concettualizzare il contesto di riferimento, che ad esempio potrebbe essere rappresentato come un tutt'uno o partizionato per domini semantico-funzionali, nello spirito del Data Mesh.



La rete **OntoPiA** come stack di ontologie

Altra iniziativa, sempre legata a fornire un sostrato semantico in grado di rendere efficace la comunicazione e condivisione tra amministrazioni e soggetti diversi, è il progetto “**Schema**”, un catalogo nazionale per lo scambio di dati e informazioni tra pubbliche amministrazioni, del quale **OntoPiA** è diventato parte e che si sviluppa in un insieme di ontologie e schemi dati che, complessivamente, altro non sono che una concettualizzazione del mondo pubblico.

## Le architetture logiche in uno scenario *Augmented*

L'idea alla base di un'Augmented Organization è quella dell'integrazione, condivisione e uso di dati a prescindere da chi tali dati possiede e gestisce e, ovviamente, dalle loro caratteristiche tecniche.

Se dal punto di vista del Data Consumer, l'ownership di un dato, le sue caratteristiche tecniche e le regole che ne consentono la corretta gestione, sono elementi utili ma secondari, in quanto il suo primario interesse è ciò che il dato rappresenta e quanto bene lo fa, differente è la questione per chi deve gestire tali dati, che benché debba ovviamente occuparsi del ruolo di rappresentazione che i dati hanno, deve anche garantirne la corretta governance, assicurandosi che ogni dato sia utilizzabile da chi ne abbia effettivamente diritto.

Finché tale gestione è relativa ai dati interni a un'amministrazione, la corretta gestione di tali aspetti è tutto sommato semplice e ben consolidata nelle prassi abituali di lavoro, la questione si complica quando l'integrazione si spinge al di fuori dei confini di un'amministrazione, attingendo a dati che sono di proprietà di soggetti terzi.

In tale scenario, infatti, l'idea di realizzare l'integrazione secondo un modello tradizionale, basato sul ben noto paradigma ETL e quindi basato sulla replica dei dati, copiandoli all'interno di quello che poi sarà il loro punto di accesso, benché tecnicamente fattibile – per raggiungere una sorgente dati è sufficiente un indirizzo di rete e le giuste credenziali – diventa quasi un incubo dal punto di vista della governance, in quanto si tratterebbe di assumersi la responsabilità di gestione di dati che appartengono a soggetti legalmente differenti e che, quindi, sono probabilmente soggetti a regole, norme e politiche di gestione diverse.

In uno scenario come questo, al contrario, una architettura logica consente una agilità e facilità di gestione decisamente maggiore e, di fatto, rappresenta l'unica alternativa possibile, visto che non richiedendo la replica preventiva dei dati, non presenta quelle complessità di gestione legate alla loro duplicazione, facilitando la

definizione di tutte le regole necessarie operando esclusivamente sul modello logico.

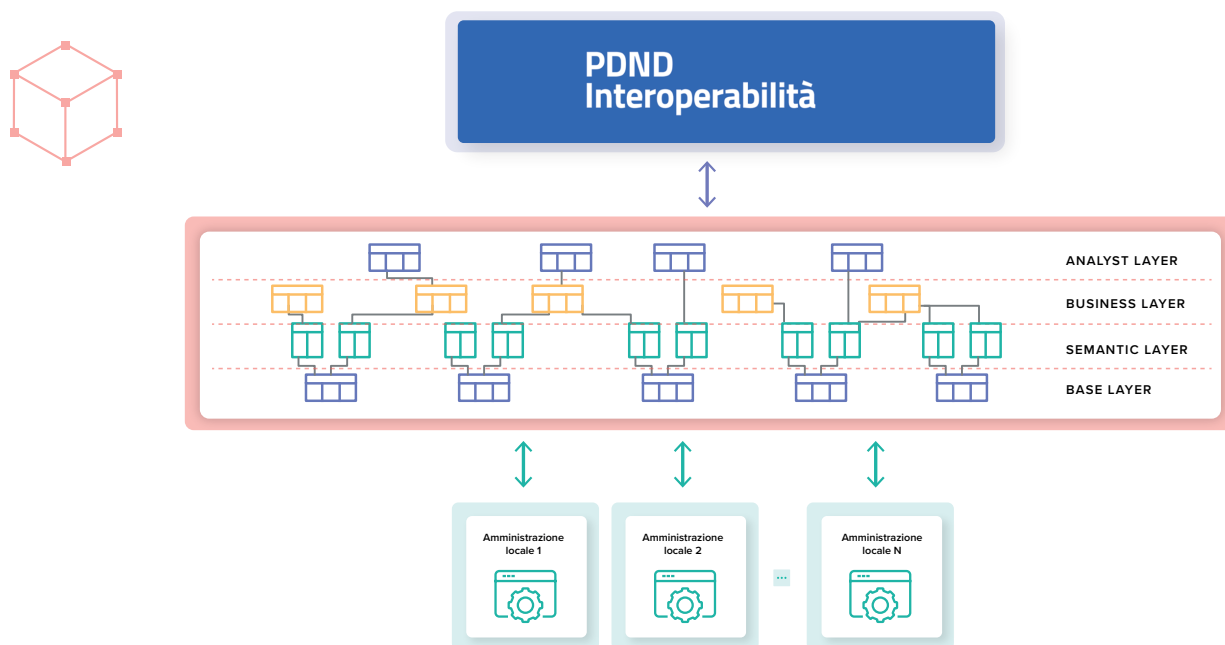
Non è un caso, infatti, che tutte le iniziative mirate a condividere dati di soggetti diversi, come ad esempio gli **International Data Space** o, in Italia, il **Sistema Pubblico di Interoperabilità** e la relativa **PDND**, sono basati su un modello logico, dove viene centralizzato solamente il catalogo dei dati disponibili, mentre questi rimangono laddove sono originariamente presenti, per essere prelevati e consegnati solo quando richiesto.

## Il ruolo delle architetture logiche nella PDND

Nel precedente paragrafo è stata menzionata la **PDND** come esempio di architettura logica, necessaria in conseguenza della distribuzione dei dati tipica del settore pubblico e della relativa responsabilità che su di essi hanno le amministrazioni.

Poiché le architetture logiche, in virtù del loro approccio e dell'agilità che ne consegue, sono la scelta migliore qualora si vogliano integrare dati eterogenei e distribuiti, risulta piuttosto naturale pensare a un ecosistema più articolato, che veda sicuramente la **PDND** come uno degli elementi centrali, ma nel quale si possa prevedere un'articolazione ulteriore, che consenta di dare maggior flessibilità a un modello già di per sé valido e votato alla condivisione e valorizzazione dei dati.

Con tale spirito di ulteriore potenziamento, ad esempio, si potrebbe pensare alla creazione di hub locali, che abbiano il compito di federare, sempre secondo un paradigma logico, i dati di tutte quelle amministrazioni territoriali che, per diverse ragioni, non abbiano le competenze o le risorse per operare in proprio. Questo hub fungerebbe poi da punto di instradamento verso la **PDND** vera e propria, come rappresentato, in modo sintetico, dalla figura che segue.

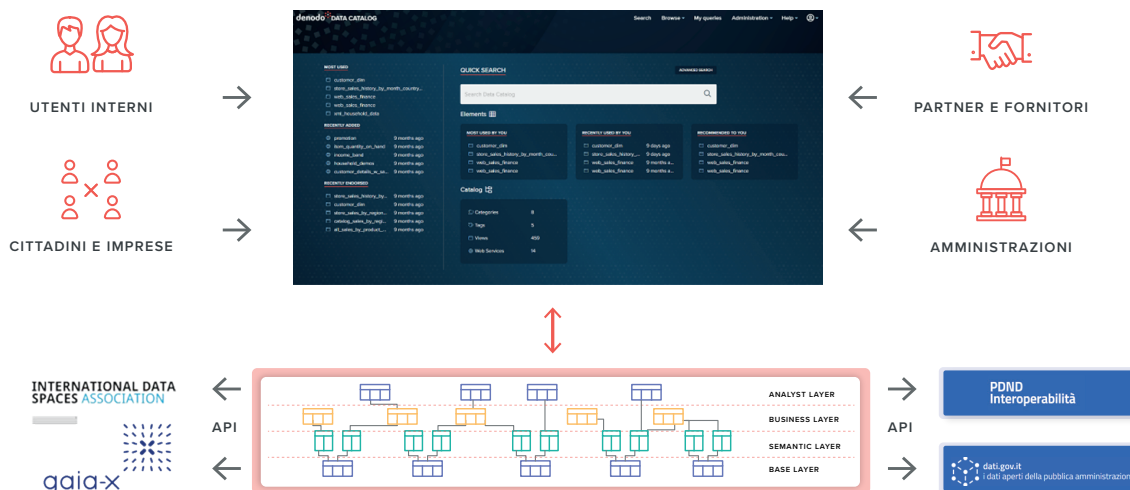


Il ruolo delle architetture logiche come hub locale di instradamento verso la PDND

Volendo invece fare un esempio, per così dire di verso opposto, si potrebbe pensare a un'architettura logica che veda la **PDND** come una delle possibili fonti dati, insieme ad altre di respiro internazionale, come i già citati Data Space, così da realizzare una vera e propria piattaforma transnazionale di condivisione e valorizzazione dei dati,

secondo quello spirito ben catturato dal **Data Governance Act**.

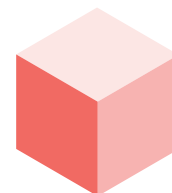
La disponibilità del Data Marketplace, poi, consentirebbe di istituire, su questo modello, un punto di accesso, esplorazione, consultazione e uso dei dati da parte di tutti i portatori di interesse, liberando tutte le potenzialità del concetto di *altruismo dei dati*, come sintetizzato dalla figura seguente.



Il ruolo delle architetture logiche come hub verso uno scenario transnazionale

Naturalmente, un approccio logico in grado di favorire l'interoperabilità dei dati, favorendone e semplificandone la governance, può portare analoghi benefici quando i dati vengono messi a disposizione della comunità, secondo quello spirito di restituzione del valore che anima gli Open Data, che non vincola l'uso dei dati a specifici servizi o all'ottemperanza di particolari obblighi di chi è parte di tale comunità, ma semplicemente li mette a disposizione affinché questi, usati nei modi opportuni e previsti dalla normativa, possano generare ulteriore valore attraverso l'uso che ne può fare la comunità stessa.

Analogamente a quanto descritto per la PDND, quindi, le architetture logiche possono rappresentare un elemento fondante per la creazione, ad esempio, di hub di aggregazione di specifici contesti, siano questi geografici o tematici, in modo da realizzare un modello federato multilivello, senza che sia necessario duplicare i dati ma semplicemente *connettendoli*, centralizzandone solamente il significato, così che siano facilmente esplorabili prima che siano effettivamente utilizzati, secondo quel modello già oggi catturato dal modello italiano, che vede il **portale dei dati aperti della pubblica amministrazione** come punto di ingresso, consultazione e consumo di ciò che le amministrazioni mettono a disposizione della comunità.



Cerca nei metadati 

[Ricerca avanzata](#)

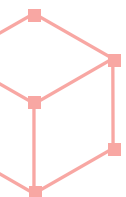
## naviga i dati per categoria tematica



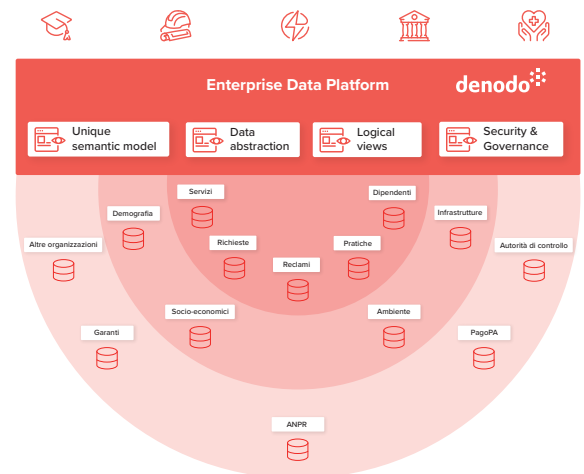
Il portale dei dati aperti della pubblica amministrazione

## Vantaggi delle architetture logiche nel Settore pubblico

Il settore pubblico può trarre notevoli benefici dall'implementazione di una architettura logica, come riassunto nella figura seguente, che suggerisce alcuni casi d'uso che possono positivamente derivare dalle Augmented View, sia questa quella di un cittadino, di un'impresa o di un generico asset complesso che una data amministrazione pubblica deve gestire.



|                                                         |
|---------------------------------------------------------|
| POLIZIA PREDITTIVA                                      |
| GESTIONE DELLA DOMANDA E OFFERTA DEL MONDO DEL LAVORO   |
| ANTITERRORISMO                                          |
| RILEVAMENTO DELLE FRODI                                 |
| MONITORAGGIO DELLE ATTIVITÀ SOSPETTE                    |
| RAFFORZAMENTO DELLA SICUREZZA NAZIONALE                 |
| SORVEGLIANZA DELLA SICUREZZA PUBBLICA                   |
| OTTIMIZZAZIONE DEL WELFARE SOCIALE                      |
| COORDINAMENTO DELLA RISPOSTA ALLE EMERGENZE             |
| PROTEZIONE AMBIENTALE                                   |
| OTTIMIZZAZIONE DEL TRASPORTO PUBBLICO                   |
| SMART BUILDING E SMART CITY                             |
| MIGLIORAMENTO DELL'EROGAZIONE DELL'ASSISTENZA SANITARIA |



#### Possibili casi d'uso per il Settore pubblico

Le architetture logiche dei dati possono poi apportare ulteriori vantaggi, di natura più generale, come brevemente illustrato nei paragrafi seguenti.

### 1. INTEGRAZIONE DEI DATI MIGLIORATA

Le architetture logiche facilitano l'integrazione di dati provenienti da diverse fonti, tra cui quelle interne all'azienda, i dati forniti da fornitori esterni e quelli condivisi da partner e fornitori. Questa integrazione consente alle pubbliche amministrazioni, come già ampiamente descritto, di creare una visione completa dei propri asset.

### 2. MIGLIORAMENTO DEL PROCESSO DECISIONALE

Con una visione integrata, i responsabili delle diverse funzioni possono prendere decisioni più informate. La possibilità di accedere a dati in tempo reale da più fonti consente di migliorare la pianificazione degli interventi, la gestione dei rischi e l'allocazione delle risorse.

### 3. AUMENTO DELL'EFFICIENZA E DELLA PRODUTTIVITÀ

Le architetture logiche semplificano l'accesso ai dati e riducono il tempo dedicato al loro reperimento e analisi. Fornendo una visione unificata dei dati, le amministrazioni possono accedere rapidamente alle informazioni di cui hanno bisogno senza dover navigare in più sistemi. Questa efficienza si traduce in una maggiore produttività e in tempi più rapidi di completamento delle diverse iniziative.

### 4. COLLABORAZIONE RAFFORZATA E MIGLIORATA

Per il settore pubblico è vitale la collaborazione tra le diverse amministrazioni e tra queste e tutti i soggetti che con esse interagiscono. Le architetture logiche facilitano la condivisione dei dati e la collaborazione senza soluzione di continuità, fornendo un quadro comune di dati. Questa comunanza garantisce che tutti abbiano accesso agli stessi dati, riducendo le incomprensioni e migliorando la comunicazione.

# Il ruolo dell'intelligenza artificiale nel settore pubblico

Come discusso nei paragrafi precedenti, il settore pubblico genera e usa grandi quantità di dati, dalla pianificazione e definizione delle iniziative e degli interventi, fino alla loro esecuzione e alla misura degli effetti che queste producono. L'utilizzo efficace di questi dati è fondamentale per ottimizzare i risultati, soprattutto per quelle iniziative di ampio respiro e diluite nel tempo e le tecnologie AI forniscono strumenti sofisticati che migliorano le capacità decisionali.

Tuttavia, non dobbiamo dimenticare che ogni algoritmo di intelligenza artificiale, tradizionale o generativo, utilizza i dati esattamente come fa ogni altro Data Consumer, ad esempio un Data Scientist nello svolgere le sue attività, motivo per cui è fondamentale adottare un approccio all'integrazione e al consumo dei dati agile ed efficiente, che consenta rapidamente di aggiungere nuove sorgenti dati, che consenta, laddove necessario, di utilizzare dati in real-time e che ne consenta un'agevole governance, indipendentemente da come i dati siano fatti, da dove risiedano e della loro ownership.

Da questo punto di vista, sono proprio le architetture logiche la migliore risposta possibile, che grazie alla separazione logico-fisica dei dati, consentono di centralizzarne la rappresentazione, così che sia agevole costruire i dataset necessari agli algoritmi di intelligenza artificiale, lasciando i dati nelle rispettive sorgenti, prelevandoli solo quando sia necessario.

## IA TRADIZIONALE

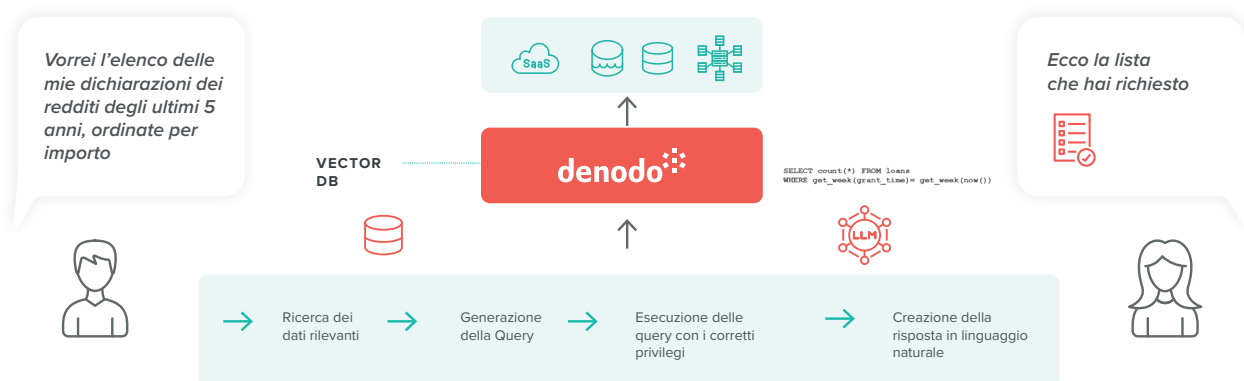
L'IA tradizionale, che comprende l'apprendimento automatico e l'analisi predittiva, svolge un ruolo fondamentale nel predisporre modelli predittivi che consentano di prevedere ciò che potrebbe accadere e, di conseguenza, anticipare interventi e misure al fine della mitigazione dei rischi e alla massimizzazione dei risultati.

## IA GENERATIVA

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 cittadini, imprese e dipendenti pubblici informazioni dettagliate, sia per un uso immediato, che per fornire più rapidamente le risposte.

## RETRIEVAL AUGMENTED GENERATION (RAG)

Il RAG integra i modelli di intelligenza artificiale con fonti di dati strutturati e non strutturati, migliorando sia il processo decisionale che la capacità degli assistenti virtuali di fornire risposte contestuali, nelle quali è necessario dare risposte che richiedono l'integrazione tra il linguaggio naturale e i dati pertinenti con la richiesta fatta, come ad esempio sintetizzato dalla figura seguente.



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

## L'INTELLIGENZA ARTIFICIALE PER IL DATA MANAGEMENT

In aggiunta agli specifici casi d'uso per il settore pubblico, non va dimenticato che l'intelligenza artificiale rappresenta una fondamentale leva di miglioramento per la gestione dei dati in quanto tale, aiutando sia il Data Consumer che il Data Steward o il Data Engineering, nei compiti nei quali sono abitualmente coinvolti.

- **Query in linguaggio naturale.** Ogni utente può semplicemente digitare 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.
- **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 dei 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.

## UNO SGUARDO AL FUTURO

Nonostante i vantaggi, l'implementazione dell'IA deve affrontare sfide quali i silos di dati, la complessità dell'integrazione e la necessità di professionisti qualificati. La ricerca futura dovrebbe concentrarsi sul miglioramento dell'interoperabilità tra gli strumenti di IA, sul potenziamento dei quadri di condivisione dei dati e sulla risoluzione delle problematiche etiche legate al processo decisionale guidato dall'IA, tutti temi che sono agevolmente risolti dalle architetture logiche.

In definitiva, l'intelligenza artificiale sta rivoluzionando il settore pubblico, come peraltro gli altri settori, consentendo l'uso ottimale dei dati e facilitando decisioni informate, ma presenta ancora ambiti di studio e miglioramento, come la riservatezza dei dati, la gestione di quelli personali, la rispondenza alle normative in materia, le questioni etiche e di copyright dei dati.

# Due storie di successo

L'approccio logico all'integrazione dei dati è stato scelto e implementato con successo da diverse amministrazioni Pubbliche, in tutto il mondo e anche in Italia, dimostrando come tale paradigma possa effettivamente agevolare l'integrazione, la valorizzazione e la condivisione dei dati in un settore che, come già ampiamente discusso, si trova a dover gestire un significativo patrimonio informativo, fortemente distribuito, sia per quanto riguarda aspetti tecnici che di governo e controllo.

Un patrimonio informativo che, oltre a essere funzionale alle attività delle diverse amministrazioni, rappresenta un valore per la comunità, dalla quale peraltro molti dei dati sono generati, alla quale tale valore deve essere restituito, nell'ovvio rispetto delle regole e della normativa, secondo quello spirito ben catturato dal **Data Governance Act**.

Nei successivi paragrafi, saranno quindi approfondite due di queste storie di successo.



*“Partendo dall’esperienza con Denodo, abbiamo intenzione di creare database virtuali incrociando tutti gli archivi oggi a disposizione. Dopo aver lavorato su quelli “civili”, passeremo alla virtualizzazione dei “sanitari”, sempre nell’ottica del rispetto della privacy e della corretta gestione delle informazioni, con opportunità di integrazione ora poco immaginabili.”* - **Diego Antonini**, Amministratore unico di Insiel.

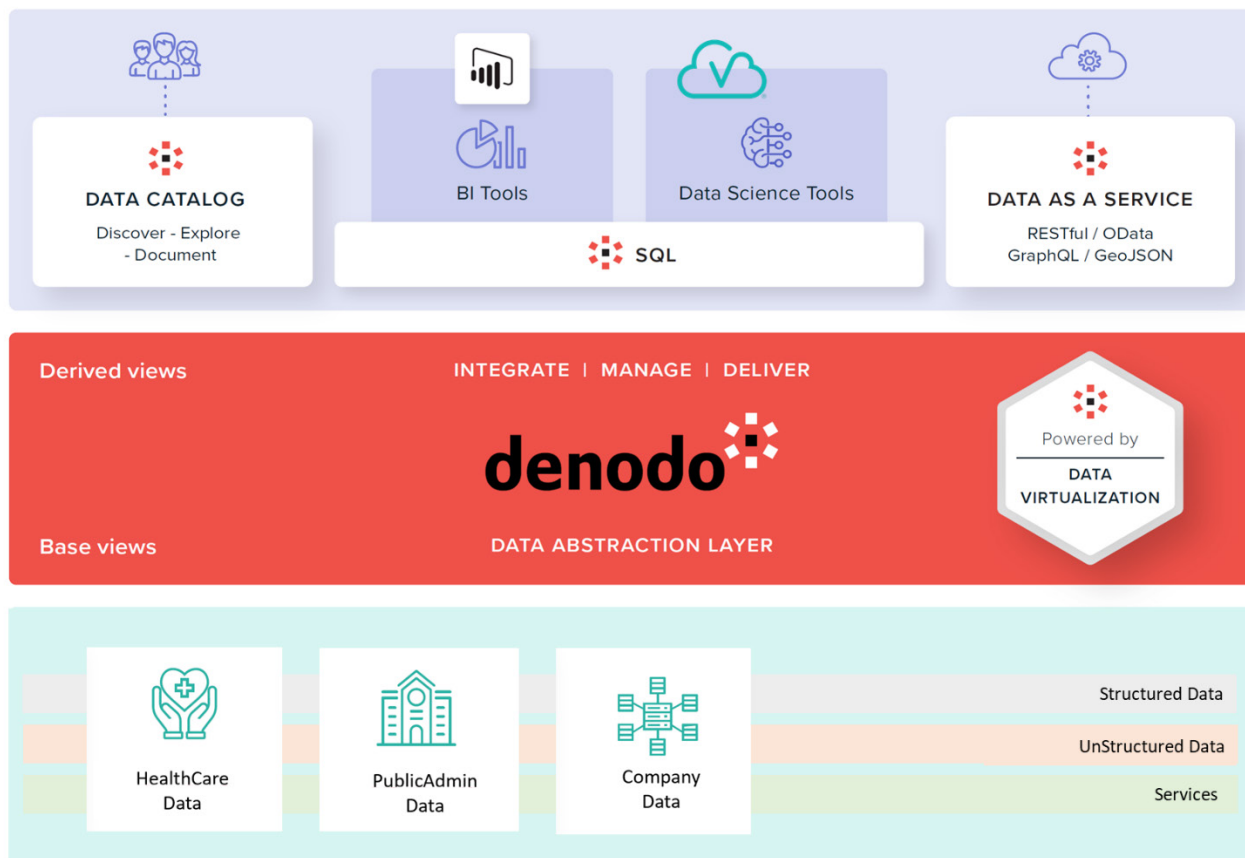
Fin dalla sua nascita, Insiel ha offerto servizi per la regione Friuli-Venezia Giulia, compresi quelli rivolti alla Pubblica Amministrazione, agli Enti Locali e alla Sanità Pubblica e, per restare fedeli alle linee guida di Insiel – utilizzare la tecnologia per fornire servizi utili a cittadini e imprese; implementare e gestire sistemi informativi per facilitare il lavoro e migliorare la qualità della vita; favorire l'innovazione in un mondo in costante cambiamento – Insiel ha voluto offrire ai propri utenti un sempre maggiore accesso ai dati, per permettere loro di prendere decisioni più informate, con la tranquillità di potersi fidare della qualità dei dati, così da rendere più efficaci ed efficienti i servizi pubblici.

*“Internamente questo concetto è stato denominato “valorizzazione” dei dati, vale a dire razionalizzarne e semplificarne l’accesso e l’uso. La valorizzazione si fonda anche sull’esplicitazione della semantica dei dati, per creare nuova conoscenza attraverso la loro analisi e l’uso di algoritmi appropriati. Per raggiungere questo obiettivo, Insiel ha compreso che avrebbe avuto bisogno anche di un catalogo dati che fungesse da interfaccia semplificata per comprendere e accedere alle diverse fonti dei dati.”* – **Andrea Soranzio**, responsabile dell'innovazione.

Con l'obiettivo quindi di sfruttare il grande volume di dati, Insiel ha scelto la Denodo Platform per creare un Data Marketplace e avviare, al tempo stesso, un processo di *standardizzazione semantica*.

Come testimoniato ancora una volta da **Andrea Soranzio**: *“avevamo bisogno di una piattaforma tecnologica con la quale gli utenti non dovessero preoccuparsi degli aspetti tecnici, potendo così dedicare più tempo all’analisi dei dati e trarne valore aggiunto, consapevoli dei diversi tipi di set di dati a loro disposizione”,* aggiungendo che *“spesso parliamo troppo di integrazione dei dati e ci concentriamo un po' troppo sugli aspetti tecnici, come ad esempio come connettersi a molti database dai formati diversi, e questo ci porta a perdere di vista i nostri veri obiettivi. Tuttavia, grazie alla Denodo Platform, possiamo integrare il nostro enorme volume di dati eterogenei, strutturati e non strutturati, e fornire i dati integrati attraverso un'unica piattaforma alle applicazioni di business intelligence e ai vari utenti”*.

**Andrea Soranzio** ha inoltre spiegato: *“quando abbiamo dovuto definire questa nuova architettura fondata sulla virtualizzazione, non l’abbiamo considerato come un progetto una tantum, ma come una scelta strategica. Inoltre, non siamo partiti da un requisito funzionale classico, come ad esempio la vista unica del cittadino, ma siamo partiti dalle fondamenta: abbiamo voluto creare una base così solida e flessibile che potesse adattarsi a grandi cambiamenti tecnologici futuri. La Denodo Platform ci ha aiutato proprio in questo modo.”*



L'architettura di Insiel per il Data Marketplace  
Maggiori informazioni sull'esperienza di Insiel sono **disponibili a questo indirizzo**.



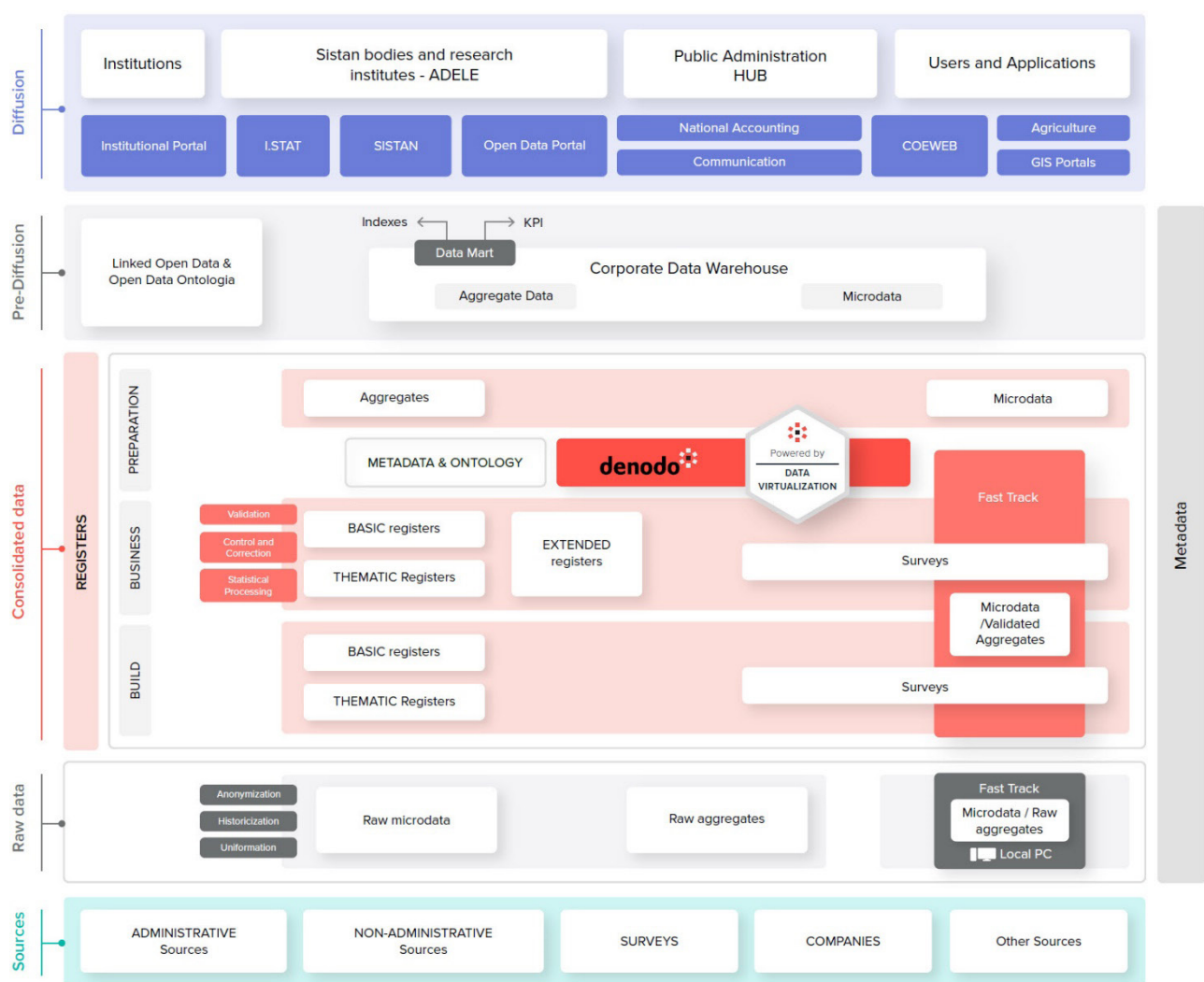
*“La piattaforma Denodo ha facilitato il processo di integrazione dei dati dell'Istat e ha consentito di collegare sistemi di dati più complessi e rigidi per l'analisi statistica dei dati demografici, economici e sociali.” – **Massimo Fedeli**, Direttore del “Dipartimento per lo sviluppo di metodi e tecnologie per la produzione e diffusione dell'informazione statistica”.*

Per ridurre al minimo le inefficienze e migliorare la qualità dei dati generati, l'Istituto nazionale di statistica italiano (Istat) ha introdotto un nuovo framework tecnologico, denominato *Register-Based Analytics Framework* (RAF), rinnovando il proprio approccio nei confronti della raccolta dei dati e dell'analisi statistica.

La scelta di Denodo, la cui Data Platform rappresenta uno degli elementi più importanti sui quali si basa questo nuovo framework tecnologico, ha consentito all'Istat di integrare agevolmente le sue innumerevoli fonti di dati eterogenee, mettendo fine ai silos di informazioni, migliorando l'efficienza della raccolta e dell'uso dei dati e permettendo l'implementazione di nuovi casi d'uso.

Il progetto di rinnovamento, partito nel 2016, ha comportato inizialmente lo sviluppo di un linguaggio comune per un unico dizionario di riferimento, seguito dalla creazione di sette registri tematici (imprese, persone, luoghi, ecc.). L'Istat ha poi ridefinito l'architettura di riferimento lungo l'intero ciclo di vita dei dati: raccolta, validazione, manipolazione e diffusione e riorganizzato l'intero framework dati, partendo dalla logica che integra le fonti dati amministrative e le indagini.

Il RAF, in sostanza, ha ridefinito l'intero processo di acquisizione, elaborazione e governance dei dati, adottando la virtualizzazione dei dati come approccio all'integrazione, senza quindi la necessità di doverli replicare.



La nuova architettura di Istat a supporto del progetto RAF

Maggiori informazioni sull'esperienza di Istat **sono disponibili a questo indirizzo**.

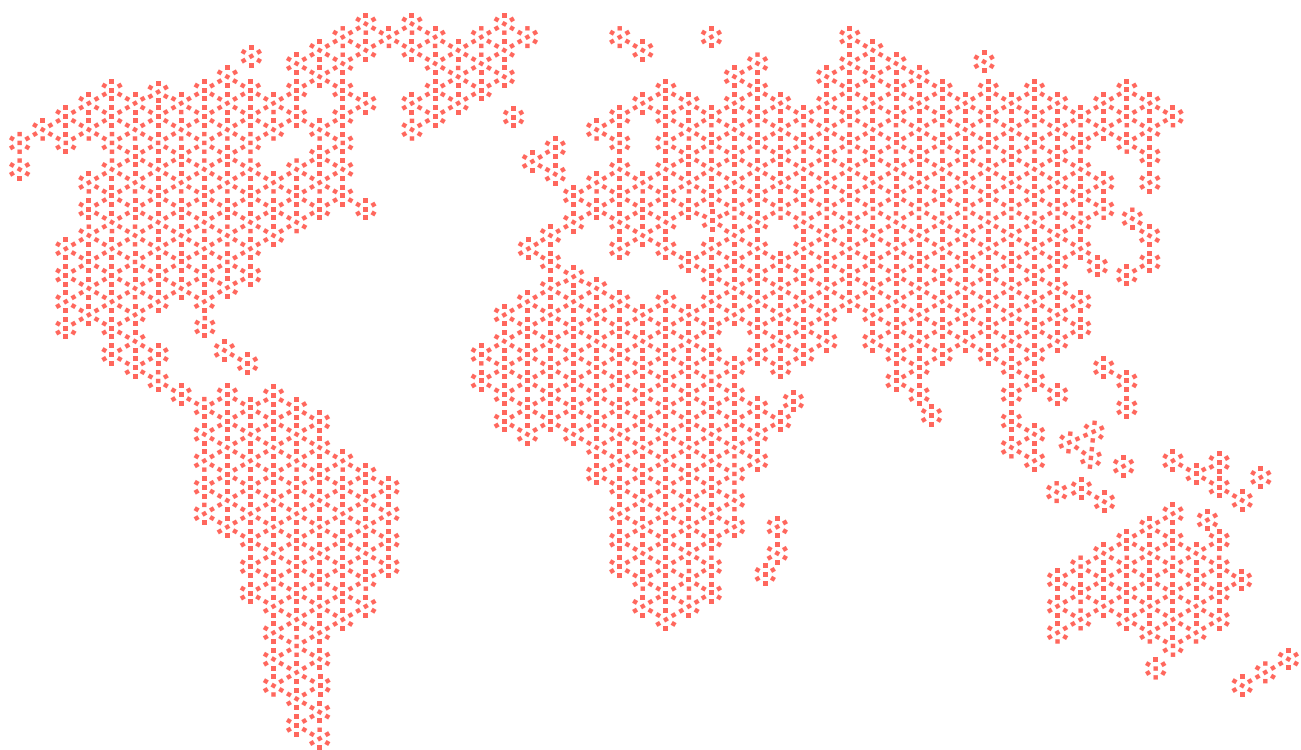
## Conclusioni

Le architetture logiche per la gestione dei dati offrono vantaggi significativi al settore pubblico, facilitando l'integrazione dei dati, migliorando il processo decisionale, aumentando l'efficienza, potenziando la collaborazione e consentendo una migliore gestione degli interventi.

Creando una visione integrata, come ad esempio quella del cittadino, utilizzando dati interni, esterni e forniti da terzi, queste architetture consentono di ottimizzare le diverse operazioni e di ottenere un maggiore successo nelle iniziative a beneficio della comunità.

Inoltre, l'integrazione di AI tradizionale, AI generativa e RAG offre numerosi vantaggi, dall'ottimizzazione dell'uso dei dati al miglioramento delle decisioni relative alla gestione delle risorse e, con la loro evoluzione, queste tecnologie svolgeranno un ruolo sempre più vitale nel plasmare il nostro futuro, guidando l'efficienza, la sicurezza e la sostenibilità.

L'adozione di una architettura logica per la gestione dei dati è una mossa strategica che può trasformare il settore pubblico e guidare l'innovazione per gli anni a venire.



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](http://www.denodo.com) | E-mail [info@denodo.com](mailto:info@denodo.com) | Scopri [community.denodo.com](https://community.denodo.com)

