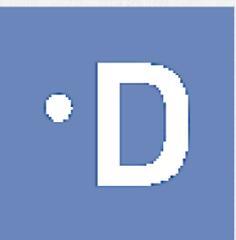


Raffaele Lillo
Chief Data Officer
@ Digital Transformation Team

Data & Analytics Framework

Come la pubblica amministrazione
può valorizzare il proprio patrimonio
di dati



Agenda

CAD & Piattaforma Digitale Nazionale Dati

Strategia e visione

Demo

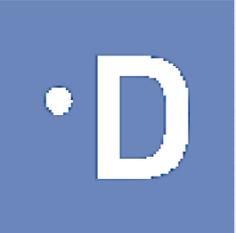
Che cosa é il DAF

Che cosa non è il DAF

A cosa serve

Piattaforma Digitale Nazionale Dati

(Art. 50-ter correttivo CAD)



1. La Presidenza del Consiglio dei ministri promuove la progettazione, lo sviluppo e la sperimentazione di una **Piattaforma Digitale Nazionale Dati** finalizzata a favorire la conoscenza e l'utilizzo del **patrimonio informativo** detenuto, per finalità istituzionali, dai soggetti di cui all'articolo 2, comma 2, lettera a), ad esclusione delle autorità amministrative indipendenti di garanzia, vigilanza e regolazione, nonché alla **condivisione** dei dati tra i soggetti che hanno diritto ad accedervi ai fini della **semplificazione degli adempimenti amministrativi** dei cittadini e delle imprese, in conformità alla disciplina vigente.

2. **In sede di prima applicazione**, la sperimentazione della Piattaforma Digitale Nazionale Dati è affidata al **Commissario straordinario** per l'attuazione dell'Agenda digitale non oltre il 31 dicembre 2018.

Piattaforma Digitale Nazionale Dati

(Art. 50-ter correttivo CAD)

3. Ai fini dello svolgimento delle funzioni di cui al comma 2, il Commissario straordinario per l'attuazione dell'Agenda digitale provvede, nel rispetto dei limiti, delle condizioni e delle modalità stabilite dal **Garante per la protezione dei dati personali** e dal decreto di cui al comma 4, ad acquisire i dati detenuti dai soggetti di cui all'articolo 2, comma 2, lettera a), ad esclusione delle autorità amministrative indipendenti di garanzia, vigilanza e regolazione, organizzarli e conservarli, nel rispetto delle norme tecniche e delle metodologie idonee a garantire la condivisione dei dati tra le pubbliche amministrazioni stabilite da AgID nelle Linee guida. I soggetti che detengono i dati identificati nel decreto di cui al comma 4, **hanno l'obbligo di riscontrare la richiesta del Commissario**, rendendo disponibili i dati richiesti senza nuovi o maggiori oneri per la finanza pubblica.

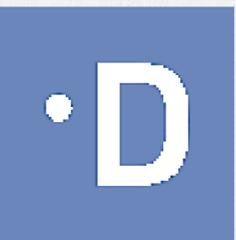


Piattaforma Digitale Nazionale Dati

(Art. 50-ter correttivo CAD)

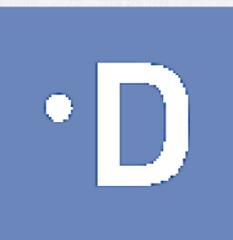
4. Con **decreto del Presidente del Consiglio dei ministri**, di concerto con il Ministro dell'economia e delle finanze, sentito il Garante per la protezione dei dati personali e le Amministrazioni titolari dei dati, sono stabilite le modalità di attuazione del presente articolo al fine di favorire la **condivisione dei dati** fra le pubbliche amministrazioni, di **semplificare l'accesso** ai dati stessi da parte dei soggetti che hanno diritto ad accedervi e di **semplificare gli adempimenti** e gli oneri amministrativi per i cittadini e le imprese, ed è identificato l'**elenco dei dati** che i soggetti di cui all'articolo 2, comma 2, lettera a), ad esclusione delle autorità amministrative indipendenti di garanzia, vigilanza e regolazione, sono tenuti a rendere disponibili per le finalità di cui al comma 3; l'elenco è aggiornato periodicamente con decreto del Presidente del Consiglio dei ministri, sentito il Garante per la protezione dei dati personali e le Amministrazioni titolari dei dati. Con il medesimo decreto sono stabiliti i limiti e le modalità di acquisizione, organizzazione e conservazione dei dati.

5. Il trasferimento dei dati nella Piattaforma Digitale Nazionale Dati **non modifica la titolarità del dato.**



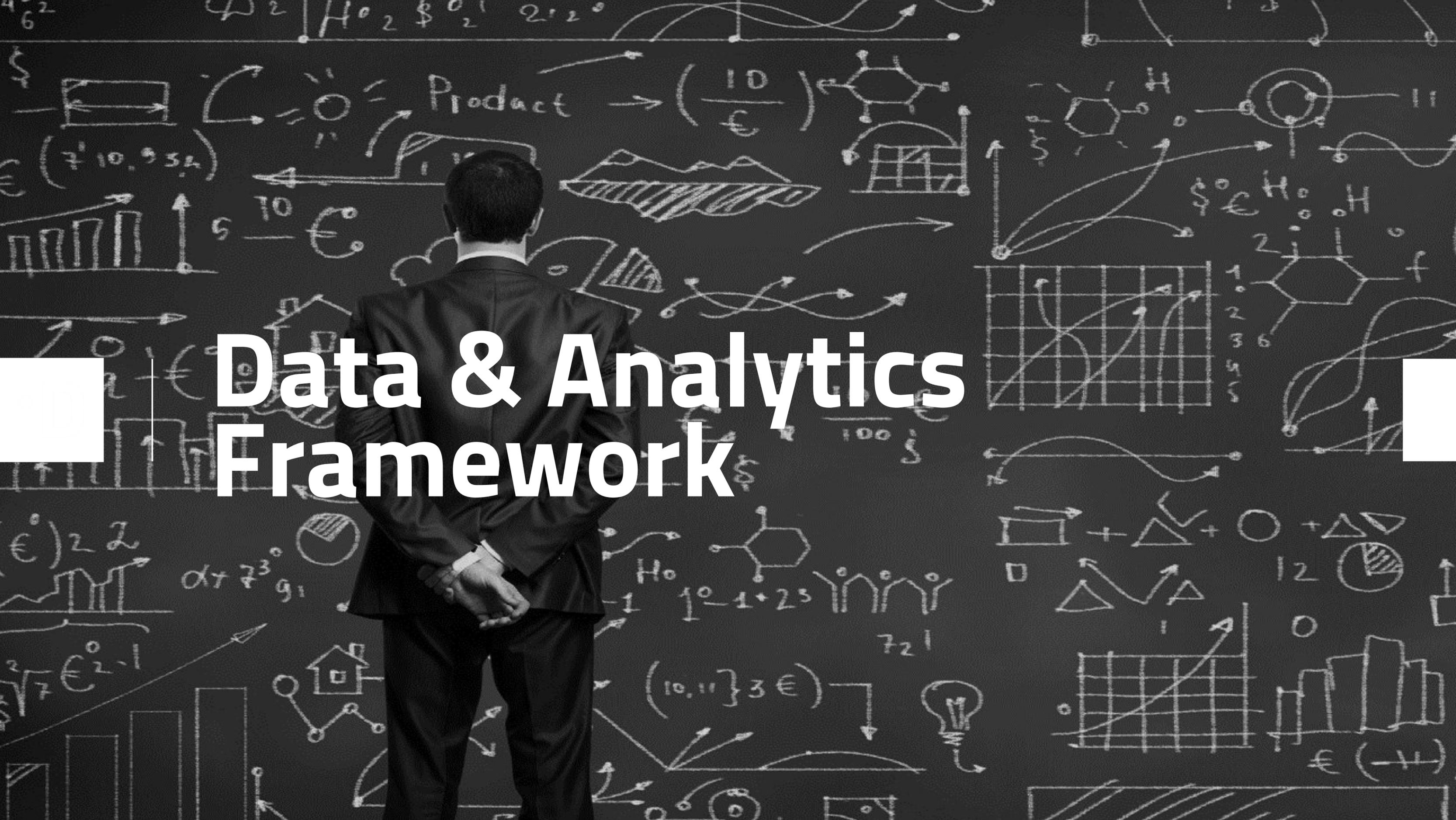
Visione

L'informazione é un asset fondamentale per interpretare i fenomeni sociali ed economici, prendere decisioni, migliorare i servizi ai cittadini, competere sullo scenario internazionale. Le nuove tecnologie permettono di estrarre sapere dall'enorme mole di dati di cui lo Stato dispone



Strategia

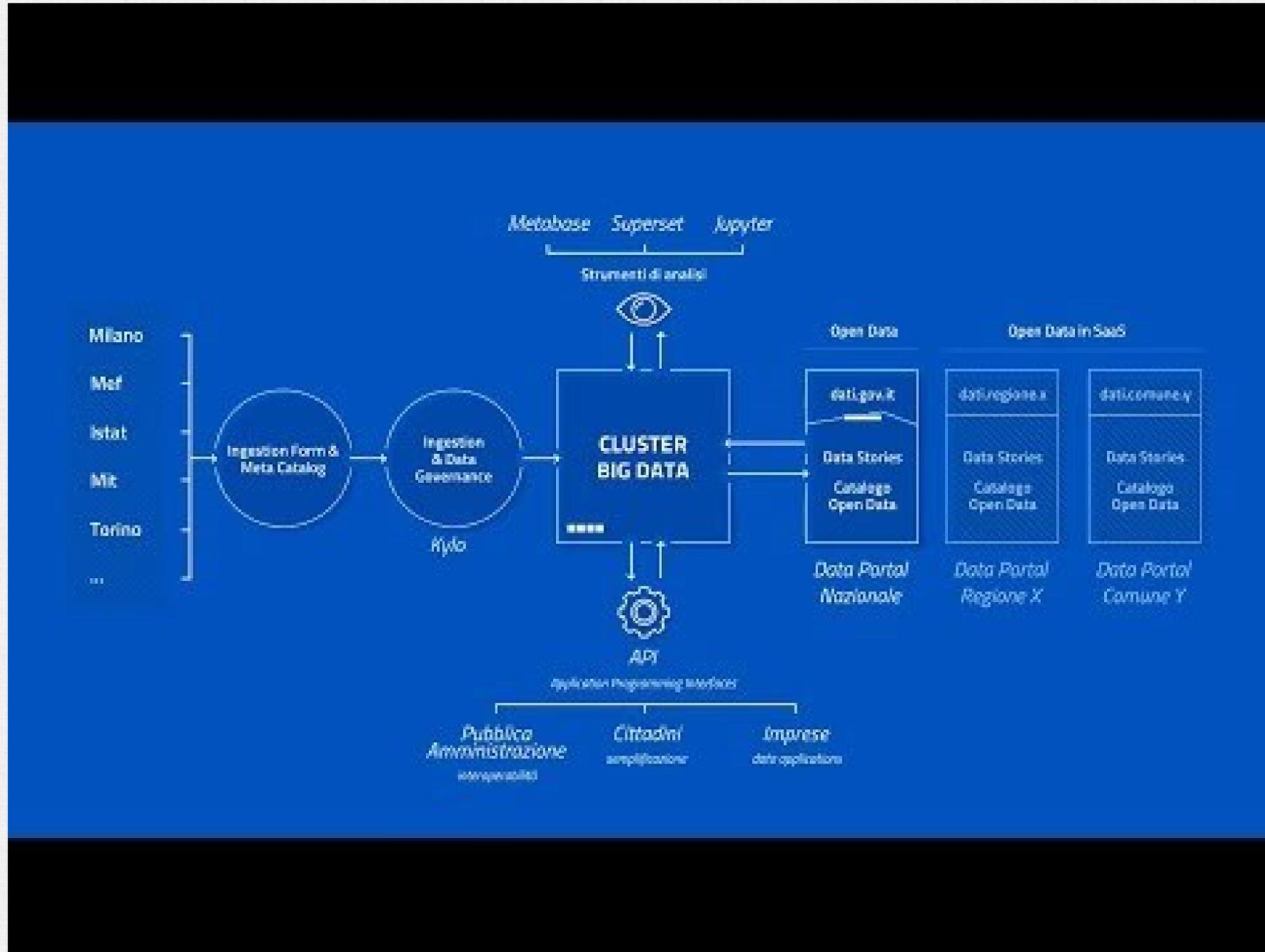
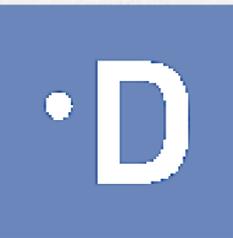
L'estrazione di valore dai dati necessita di una solida **piattaforma tecnologica**, un **team di esperti** e di una **governance** con **competenze** che coordini la **generazione**, **integrazione**, **standardizzazione** e **utilizzo** dei dati.



The background features a chalkboard densely packed with hand-drawn diagrams and mathematical notations. These include bar charts, line graphs, flowcharts, and various symbols such as € , $\text{\$}$, and H_2O . A central diagram shows the word "Product" with an arrow pointing to a fraction $(\frac{10}{\text{€}})$. Other elements include a lightbulb icon, a house, and various geometric shapes and arrows.

Data & Analytics Framework

Demo



<https://dataportal.daf.teamdigitale.it>

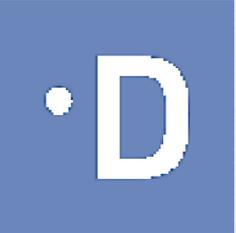
Che cos'è il DAF

- **Piattaforma Digitale Nazionale Dati (PDND):** per conservare e integrare (data lake), manipolare e standardizzare (data engine), implementare meccanismi di apprendimento automatico (machine learning), redistribuire (API & Data Applications) dati e informazioni e visualizzarli (interfaccia)
- **Data Office** trasversale a tutta la PA, con il compito di disegnare la strategia sui dati dello Stato e aiutare le PA nella sua implementazione. È composto da *data scientists* e *data engineers* che utilizzano ed evolvono una piattaforma big data, offrono consulenza alle PA, e creano *data applications*
- **Sistema normativo e regolatorio** che istituzionalizzi l'attività di valorizzazione del patrimonio informativo pubblico, definisca una governance efficace, tuteli i profili di privacy mitigando i rischi senza ostacolare l'operatività.

Che cosa non è il DAF

- non è un nuovo database
- non si sostituisce alle PA nello svolgimento delle proprie funzioni istituzionali, ma è al servizio delle PA per aiutarle ed assisterle in materia di data science.
- non acquista la titolarità dei dati, ma è responsabile della gestione della piattaforma tecnologica.
- non obbliga la PA all'utilizzo esclusivo dei suoi servizi, ma offre capacità aggiuntive alle possibilità delle PA in ambito di utilizzo e analisi dei dati.

Privacy e Security



Il Data & Analytics Framework dovrà essere compliant con i più stringenti requisiti di **sicurezza tecnologica** e di **privacy** previsti dal **GDPR**

A cosa serve il DAF

Interoperabilità e Standardizzazione (basta Silos!)

Integrazione banche dati by design: i dati vengono trattati in fase di caricamento sulla piattaforma per seguire regole e standard comuni in un unico data lake

Efficacia, Semplificazione, Risparmio (uniamo le forze!)

Condivisione di esperienze e riuso di risorse e tecnologie di data science difficilmente reperibili se cercate autonomamente dalle PA (economie di scala). Non si "re-inventa la ruota".



A cosa serve il DAF

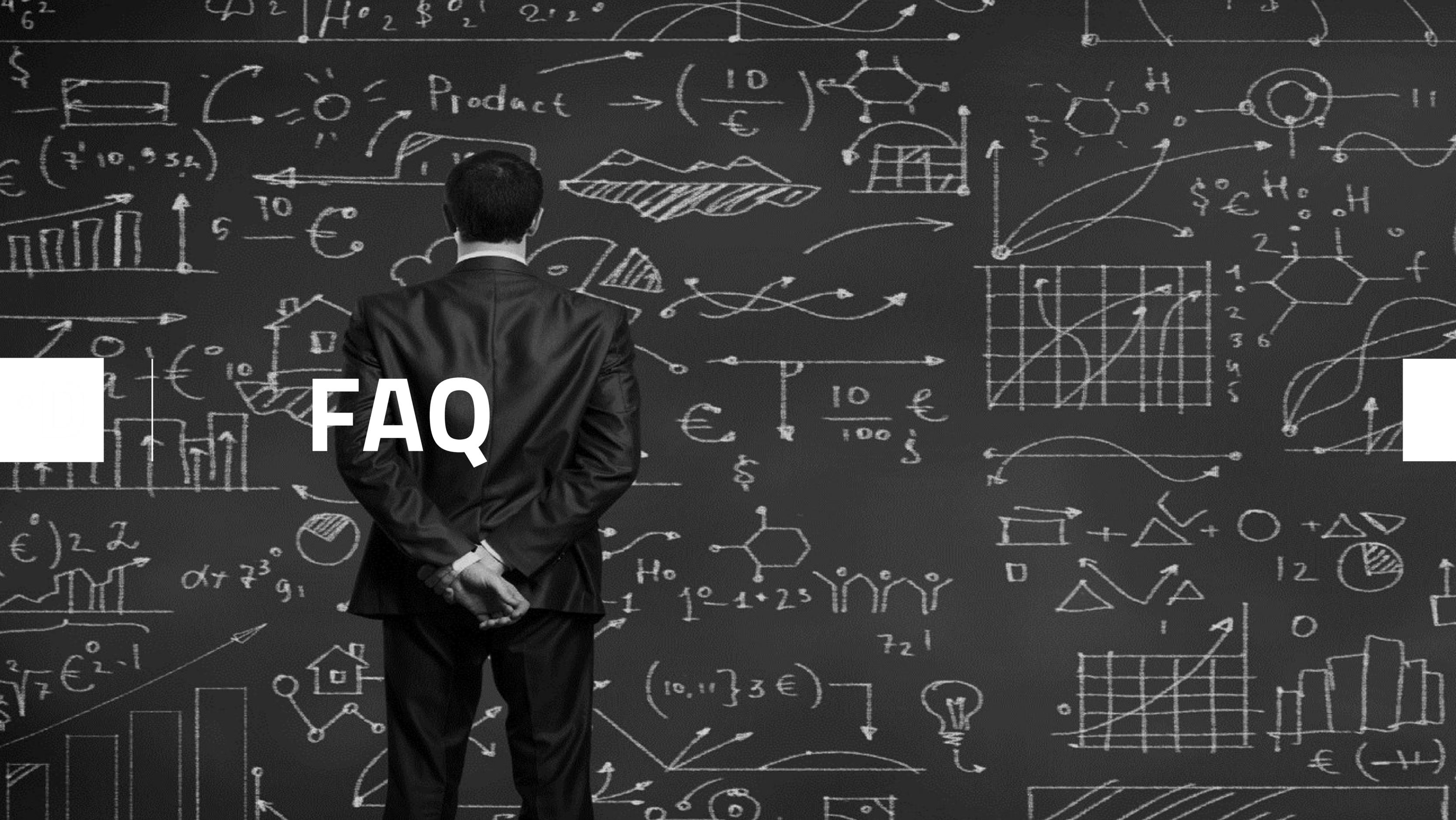
Analytics & Data Products (creiamo servizi dai dati!)

Dati sempre aggiornati all'interno di un'unica piattaforma come il DAF facilita analisi e costruzione di applicazioni che utilizzano dati per fornire funzionalità a PA e cittadini. Modelli di apprendimento automatico gestiti centralmente per tutte le PA

Open Data & Crowdsourcing (trasparenza e partecipazione!)

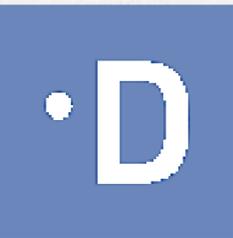
Open Data facilmente utilizzabili e automatizzabili per utilizzo anche via API sia dal pubblico che dal privato.





FAQ

Domande Frequenti



- **Update su roadmap DAF**
- **Update su Dataset Chiave**
- **Come faccio a aderire al DAF?**
- **Cosa è possibile fare dal dataportal?**

CONTATTI



RAFFAELE LILLO
Chief Data Officer

 raffaele@teamdigitale.governo.it





Seguici su:



teamdigitale.governo.it
pianotriennale-ict.italia.it



Twitter: @diegopia, @teamdigitaleIT

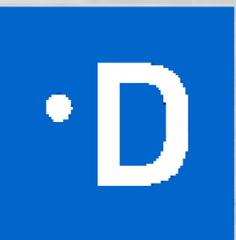


Medium: @diegopia, @team-per-la-trasformazione-digitale

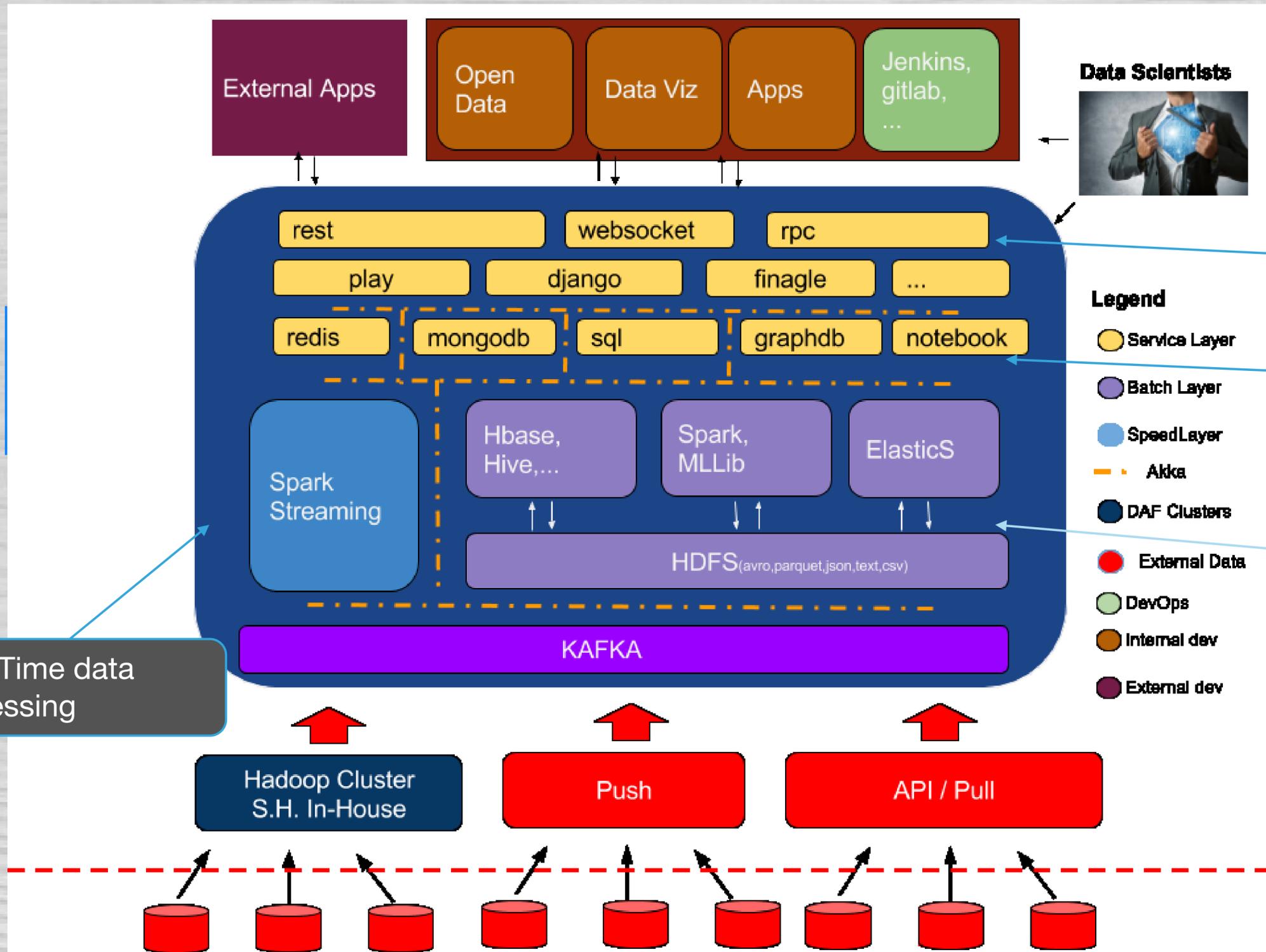


LinkedIn: @company/teamdigitale

Backup



High-level Architectural Design



Legend

- Service Layer
- Batch Layer
- SpeedLayer
- - - Akka
- DAF Clusters
- External Data
- DevOps
- Internal dev
- External dev

Communication layer - API

Operational DB for API and Data applications

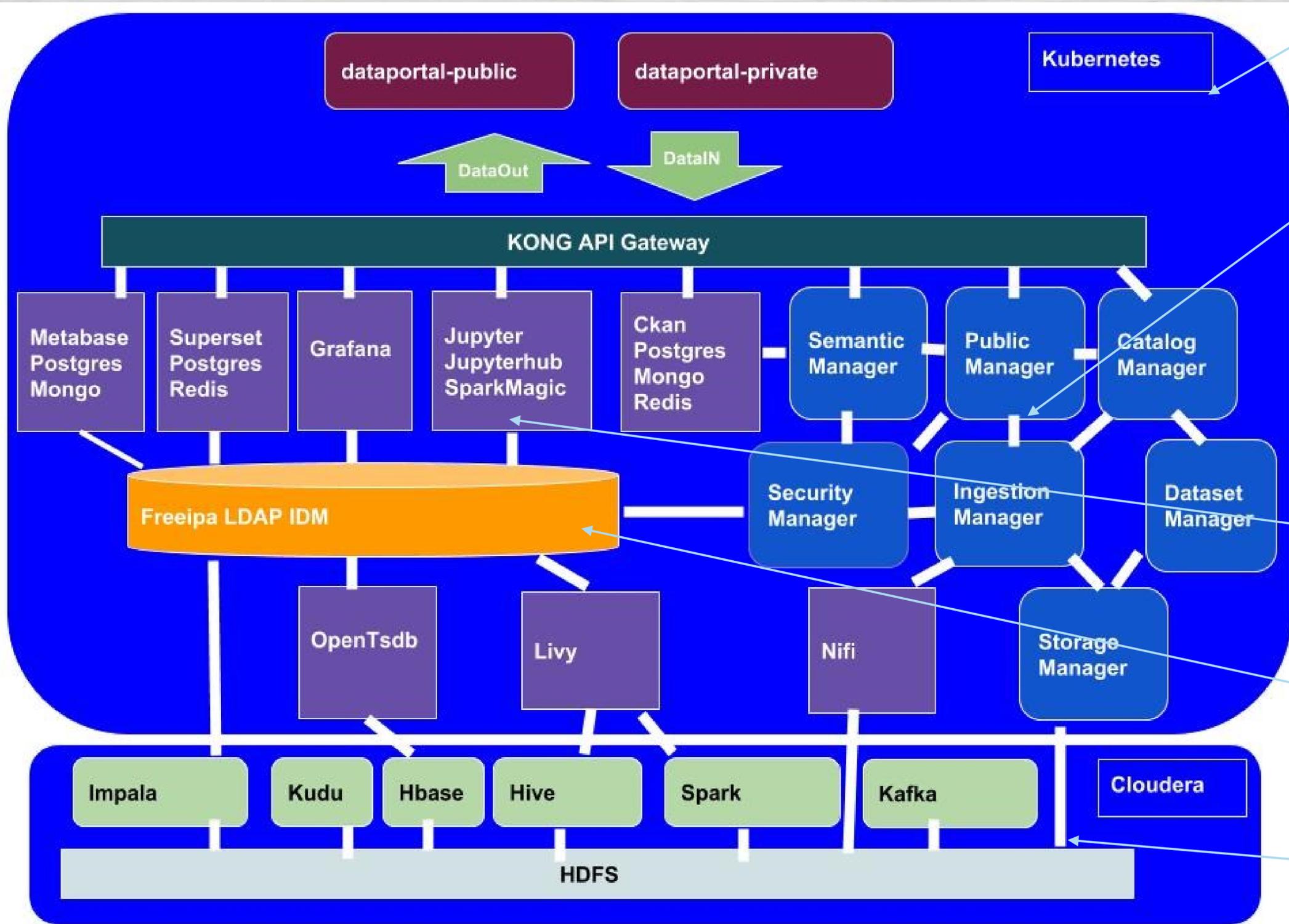
Persistence & Offline analysis (analysis, data processing, model training)

Real-Time data processing

Data Ingestion



(Less) High-Level Architectural Design



Kubernetes cluster manages dockerized microservices and external applications.

Core Managers: microservices managing core functionalities of DAF

External applications natively integrated in DAF

Unique identity management system, integrated with HDFS

Hadoop cluster for distributed persistence and processing



Dataportal - Public Version

Secure | <https://dataportal.daf.teamdigitale.it>

AgID + Team Digitale Piano Triennale Developers **DAF** Designers Forum Docs Github

Seguici su

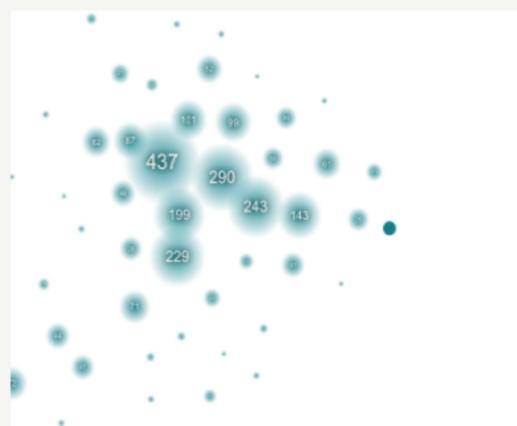
it /daf alpha
Il framework dei dati pubblici del Paese

cerca nei dataset

Il Progetto Esplora Community Partecipa [Accedi](#)

Incidenti a Roma

Quali sono le strade più a rischio incidenti?

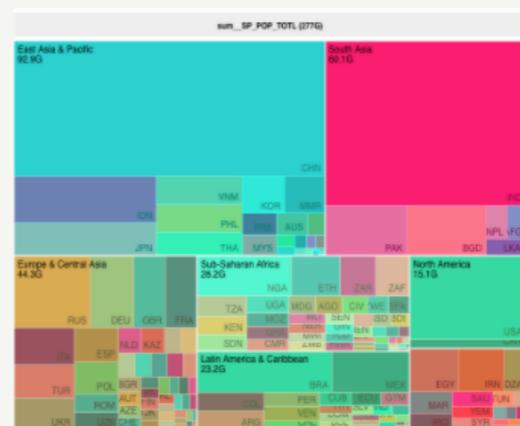


Publicato da: marcop
Data di ultima modifica:

[Per saperne di più >](#)

Esempio di storia

La popolazione nel mondo si divide

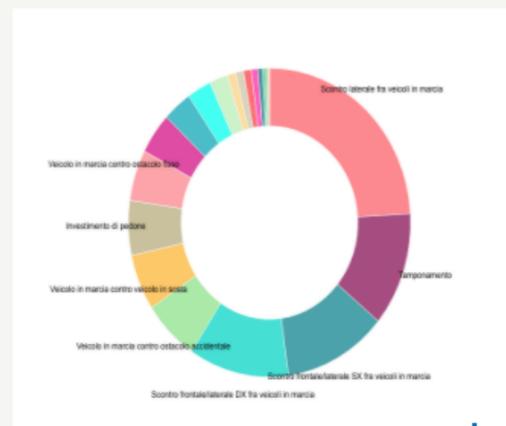


Publicato da: alessandro
Data di ultima modifica:

[Per saperne di più >](#)

Tipologia di incidenti

Tipologia di incidenti a roma



Publicato da: alessandro
Data di ultima modifica:

[Per saperne di più >](#)

<https://dataportal.daf.teamdigitale.it>

ESPLORA

Dataportal - Private Version

The screenshot displays the Dataportal Private Version interface. The top navigation bar includes the DAF logo, a search bar, and the user name 'Lillo'. The main content area is titled 'Ultime Dashboard' and shows a map of Rome with incident locations, labeled 'Incidenti a roma'. A Jupyter notebook is open in the foreground, showing code for loading Spark magic and managing a Spark session. The notebook output includes a table of Spark sessions and a code snippet for loading data from a CSV file. A network diagram is visible in the bottom left corner, showing various industrial and agricultural processes connected to a central 'Carbon Dioxide' node.

Home / Home

Ultime Dashboard

17-10-2017 14:47
alessandro

Incidenti a roma

Jupyter Untitled Last Checkpoint: 09/29/2017 (autosaved)

```
In [1]: %load_ext sparkmagic.magics
```

```
In [3]: %manage_spark
```

Manage Sessions Create Session Add Endpoint Manage Endpoint

Address: Auth type: Basic_Access

Added endpoint <http://livy.default.svc.cluster.local:8998>
Starting Spark application

ID	YARN Application ID	Kind	State	Spark UI	Driver log	Current session?
64	None	spark	idle			✓

SparkSession available as 'spark'.

```
In [9]: val df = (spark.read.format("csv")  
.option("header", "true")  
.option("inferSchema", "true")  
.load("/daf/opendata/_19luoghidellacultura_a_35e6edb1_4dec_455
```

df: org.apache.spark.sql.DataFrame = [LUOGO: string, INDIRIZZO: string, ...]

sum SP POP TOTL (277G)

<https://dataportal-private.daf.teamdigitale.it>

It... Could... Work!





Raffaele Lillo

Chief Data Officer

raffaele@teamdigitale.governo.it

Twitter, Medium: @lilloraffa



Grazie!

Cooperate with us, please :)

Website

<http://teamdigitale.governo.it>

Forum

<https://forum.italia.it/c/daf>

Twitter

#DatiPubblici #DAF

Google Group - Open Data & Analytics

data@teamdigitale.governo.it

