



Area 5 - LA PA NELLA TRASFORMAZIONE DIGITALE

Modulo 5 – Le tecnologie di frontiera e il loro uso per i servizi pubblici

Pillola 5.5a – Big Data per una PA data driven

Introduzione

Ciao, benvenuto in questo nuovo video in cui ti spiegherò quale enorme importanza hanno assunto ai nostri giorni i dati posseduti dalla PA.

Vedremo, in particolare cosa si intende per Big Data e perché rappresentano una grandissima opportunità per la PA.

Dati delle PA come patrimonio digitale

Cominciamo da questa affermazione: i dati sono uno dei principali patrimoni digitali della PA!

Per questo l'AgID (Agenzia per l'Italia Digitale), nel suo Piano Triennale per l'Informatica nella Pubblica Amministrazione, ha incluso la valorizzazione del patrimonio informativo pubblico nazionale tra gli obiettivi strategici da perseguire.

La Piattaforma Digitale Nazionale Dati

L'art. 50-ter del Codice dell'Amministrazione Digitale, per assolvere questo obiettivo, ha istituito la Piattaforma Digitale Nazionale Dati (PDND).

Questa piattaforma era precedentemente nota con il nome di Data & Analytics Framework (DAF).

La Piattaforma Digitale Nazionale Dati/2

Attraverso la Piattaforma Digitale Nazionale Dati si intende sviluppare e semplificare l'interoperabilità dei dati pubblici tra le diverse PA, sfruttando le opportunità offerte dalla gestione e dall'analisi dei Big Data.

Cosa sono i Big Data

Ma cosa si intende col termine Big Data?

Purtroppo non è possibile dare una definizione univoca di Big Data, in quanto il concetto varia a seconda degli ambiti.

In ogni caso, la definizione più comune si rifà all'enorme quantità di dati che l'attività odierna (di un cittadino o di un'organizzazione) produce.

Alcuni esempi di questi dati possono essere:

- i dati sul clima;
- le immagini satellitari;
- le immagini e i video digitali;
- le registrazioni di informazioni su moduli on line;
- o, infine, i segnali GPS.

Tieni presente, comunque, che oggi, quando si parla di Big Data, ci si riferisce anche all'insieme di tecnologie e metodi analitici utilizzati per estrarre valori di interesse da questa messe di dati.



Le 5 V

Per descrivere meglio i Big Data esistono 5 caratteristiche che li contraddistinguono, le cosiddette 5 V:

- 1) Volume;
- 2) Velocità;
- 3) Varietà;
- 4) Variabilità;
- 5) Veridicità.

Vediamole in dettaglio...

Le 5 V: Volume

Il Volume è sicuramente la caratteristica più nota associata ai Big Data e si riferisce all'enorme quantità di dati che vengono prodotti ogni giorno.

Tieni presente che spesso, per trattare e gestire questi dati, i requisiti di elaborazione superano le capacità di un singolo computer. Proprio dalla considerazione di questo aspetto, quindi, deriva una nuova sfida tesa a risolvere il problema di come procedere per mettere in comune, allocare e coordinare le diverse risorse provenienti da gruppi di computer.

Le 5 V: Velocità

La seconda V è relativa alla Velocità.

Con questo termine si intende indicare:

- sia la rapidità con cui le informazioni si muovono attraverso il sistema;
- sia la necessità di elaborazione in tempo reale di questi dati, che arrivano da più fonti.

Le 5 V: Varietà

La terza V si riferisce alla Varietà.

Questa caratteristica si riferisce alle numerose fonti da cui i dati provengono, quali:

- applicazioni e server;
- social media;
- dispositivi;
- ecc.

Nonostante questa varietà di fonti (che tuttavia va tenuta in considerazione), ai fini della loro elaborazione non è così importante la loro provenienza, in quanto le tecnologie Big Data cercano comunque di gestirle tutte con un unico sistema.

Le 5 V: Variabilità

La quarta V si riferisce alla Variabilità.

Con questo termine si rimanda alle difficoltà di gestione dei grandi flussi di dati. L'interpretazione dei dati, infatti, non è univoca, ma può cambiare in base ai contesti di raccolta ed analisi.

Le 5 V: Veridicità

Infine, la quinta V fa riferimento alla Veridicità.



Infatti, molteplici fonti ed elaborazioni complesse possono causare difficoltà nel valutare i dati come veritieri.

Questo aspetto, quindi, rimane uno degli aspetti più complicati e delicati da gestire, soprattutto poi in riferimento al trattamento dei dati sensibili.

L'evoluzione dei Big Data: Data Warehouse e Data Lake

L'avvento dei Big Data, se da un lato ha complicato (come abbiamo appena visto) alcuni aspetti gestionali, dall'altro lato ha semplificato la raccolta di nuove tipologie di informazioni, sia all'interno che all'esterno dell'organizzazione.

Per questo, il Data Warehouse tradizionale (cioè il grande archivio di dati puliti tradizionale) si è evoluto per forza di cose nei cosiddetti Data Lake.

Vediamo questi due oggetti in maggiore dettaglio...

Cos'è il Data Warehouse

Il Data Warehouse (come abbiamo già detto) è un grande archivio di dati puliti, ordinati e integrati con altre fonti, che possono essere utilizzati per l'analisi e/o la reportistica.

Il Data Warehouse viene perciò solitamente considerato come una fonte attendibile di dati certificati all'interno dell'organizzazione.

Cos'è il Data Lake

Il Data Lake, invece, indica un grande deposito di dati raccolti in uno stato relativamente grezzo. I dati che vi sono conservati, infatti, in genere non sono strutturati e spesso sono soggetti a variazioni.

Opportunità offerte dai Big Data

In ogni caso, indipendentemente dalle loro caratteristiche e forme di conservazione i Big Data costituiscono una vera e propria ricchezza, in quanto:

- se ne possono ricavare informazioni di valore,
- efficientano i processi interni,
- impostano nuove politiche di sviluppo,
- offrono informazioni utili a chi deve prendere decisioni,
- processano dati in tempo reale,
- permettono di estrarre informazioni dall'incrocio di molteplici basi di dati,
- consentono più prospettive di analisi su un dato fenomeno.

La sesta V: Valore

In questo senso è opportuno aggiungere una sesta V alla definizione di Big Data: il Valore.

Lo scopo ultimo dei Big Data, infatti, è quello di fornire un valore aggiunto

Riepilogo

Bene, siamo giunti alla fine di questo video.

Prima di chiudere, ti ricordo i punti salienti che di cui ti ho parlato:

- i Big Data sono una grande mole di dati, che se elaborati possono aiutare le PA nel definire strategie di sviluppo;



- i Big Data sono definiti da Volume, Velocità, Varietà, Variabilità, Veridicità;
- a queste 5 caratteristiche, si aggiunge il Valore come sesta dimensione;
- l'evoluzione dei Big Data ha portato ai Data Lake, ovvero a un deposito di dati raccolti a uno stato relativamente grezzo.

Spero di essere stato chiaro e d'aiuto. Buono studio....