

Sommario

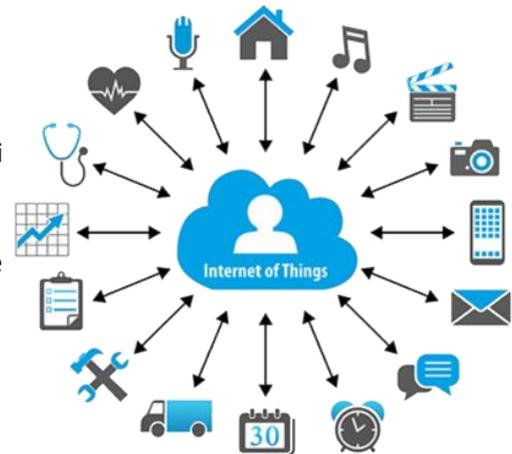
Internet Of Things	2
Che cos'è?	2
Funzionamento	2
Criticità	2
Il problema della privacy	2
Il problema della sicurezza	2
Smart city	3
Che cos'è?	3
Vantaggi	3
Svantaggi	3
Big Data	4
Introduzione	4
Le 5 V	4
La scienza dei dati	5
Tecnologie per i Big data	6

Internet Of Things

Che cos'è?

L'Internet Of Things (abbreviato in IoT), consiste nella connessione ad internet di oggetti della nostra vita quotidiana, in modo che questi oggetti "intelligenti" si interconnettano tra di loro permettendo lo scambio di dati.

Nello specifico, IoT indica qualsiasi sistema di dispositivi fisici che ricevono e trasferiscono i dati su reti wireless, che non richiede interventi manuali. Ciò si ottiene integrando semplici dispositivi di elaborazione con sensori in ogni tipo di oggetto.



Funzionamento

Un tipico sistema IoT, funziona grazie all'invio, alla ricezione e all'analisi dei dati in un ciclo continuo di feedback. Prendiamo l'esempio di una casa smart. Per prevedere il momento ottimale in cui controllare il termostato prima di arrivare a casa, il sistema IoT può connettersi a Google Maps per esaminare i modelli del traffico locale in tempo reale e utilizzare i dati dell'automobile acquisiti nel lungo periodo per conoscere le abitudini di viaggio dell'utente. I dati IoT acquisiti da ogni utente dei termostati smart possono quindi essere utilizzati dalle aziende dei servizi di pubblica utilità, per iniziative di ottimizzazione su più larga scala.

Criticità

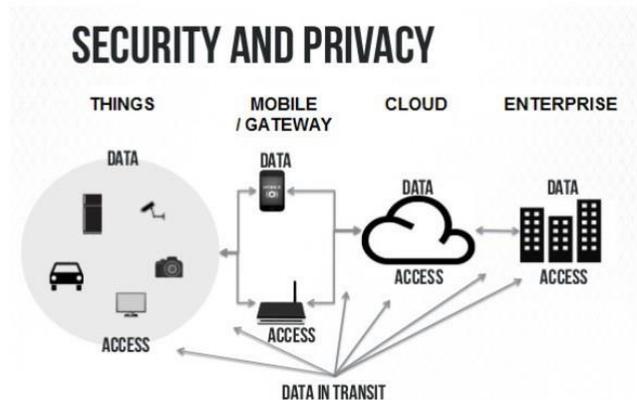
I problemi dell'IoT sono principalmente due, la sicurezza e la privacy, dovuti al fatto che i sistemi di tipo centralizzato sempre più in uso si sono rivelati in proporzione sempre più vulnerabili e manipolabili ed è per questo che la maggior parte delle aziende sta lavorando a soluzioni di sistema non più centralizzato ma distribuito, per render più sicuri i dispositivi da errori o da attacchi informatici.

Il problema della privacy

I dati creati e scambiati da un dispositivo IoT possono essere un rischio per la nostra privacy. Un attaccante intercettando i dati può scoprire, in questo modo, le loro abitudini e i loro comportamenti. Ad esempio tramite una lampadina smart è possibile sapere quando gli utenti sono in casa e quando vanno a dormire mentre dagli activity tracker è possibile ricavare dati sulle abitudini di movimento e di allenamento di chi le indossa.

Il problema della sicurezza

A causa dello sviluppo molto rapido che sta avendo l'IoT, non si presta la dovuta attenzione alla sicurezza. E ciò si traduce nella possibilità che i dispositivi possano essere soggetti ad attacchi informatici che li trasformino in oggetti per ottenere informazioni spiando i soggetti interessati.



Smart city

Che cos'è?

Una smart city (città intelligente) identifica una città che gestisce le risorse in modo intelligente, mettendo in relazione le infrastrutture materiali della città (capitale umano, sociale ed intellettuale)

tramite l'impiego diffuso delle nuove tecnologie.

In termini più pratici significa connessioni wi-fi nei luoghi più disparati, strade percorse da auto a guida autonoma, incroci regolati da semafori intelligenti, un alto livello di tecnologia high-tech.



Tutto questo è sviluppato con l'obiettivo di diventare economicamente sostenibile ed energeticamente autosufficiente.

Vantaggi

Ciò che rende vantaggioso questo tipo di città sono i bassi tempi di rientro degli investimenti, il miglioramento della qualità ambientale, che ha un impatto diretto sulla vita e la salute dei cittadini e cosa più importante un immediato e considerevole risparmio di tempo nello svolgere le azioni quotidiane: basti pensare al risparmio di tempo nelle interazioni con la pubblica amministrazione e con il sistema sanitario.

Oltre a ciò la smart city garantisce una maggiore sicurezza che si riflette in molti ambiti come ad esempio nella riduzione della criminalità e degli incidenti.

Svantaggi

Le principali criticità generate dallo sviluppo delle città intelligenti sono:

- la sopravvalutazione del loro interesse strategico potrebbe portare ad ignorare vie alternative di sviluppo urbano;
- Il fatto che potrebbero sottovalutare i possibili effetti negativi derivanti dallo sviluppo delle nuove infrastrutture tecnologiche e di rete di una smart city.

Big Data

Introduzione

Già nel 1975 si cominciò a parlare di Very Large Data Bases, recentemente è stato introdotto nell'uso comune il termine big data.

Oggi la produzione di dati ha raggiunto livelli straordinari e questo ha comportato la formazione di nuovi strumenti informatici sempre più evoluti in grado di memorizzare, interrogare ed analizzare questi dati.

Possiamo quindi dire che...

- Il termine **Big Data** è usato per indicare un insieme di nuove tecnologie atte ad acquisire, elaborare, mantenere e visualizzare dati che altrimenti sarebbero intrattabili con le odierne tecnologie.

Ciò è reso possibile grazie all'adozione storage distribuiti, **database NOSQL** e in generale grazie alle tecniche di calcolo distribuito.



Le 5 V

Il concetto di big data ha acquisito uno slancio solo *all'inizio degli anni 2000*, quando venne articolata l'attuale definizione di big data come "le 5 V":

VOLUME

È la principale caratteristica dei big data.

Con volume si fa riferimento quindi a quest' ingente massa di informazioni, che non è possibile raccogliere con tecnologie tradizionali.

Lo stesso quantitativo di dati prodotto dall'inizio dell'umanità ad oggi sarà a breve generato in un minuto.

In un minuto, ad esempio, si generano 2,3 milioni di interrogazioni a Google, 3 milioni di like e 3 milioni di share su Facebook; 2,7 milioni di video vengono scaricati da Youtube (e 139 mila nuovi video vengono caricati); 44 milioni di messaggi vengono processati, di cui 486 mila contenenti foto e 70 mila contenenti video

VELOCITA'

Ossia la velocità con cui i dati vengono generati e scambiati.

La sfida, con cui le aziende sono chiamate a confrontarsi, è la necessità non solo di raccogliere questi dati ma anche analizzarli in tempo reale, per poter prendere decisioni di business con la maggiore tempestività possibile, evitando così di memorizzare i dati nel database dove vengono memorizzati solo i risultati dell'analisi. A tal proposito ci vengono incontro tecniche di **data streaming**.

VARIETA'

Con varietà si fa riferimento proprio alle differenti tipologie di dati oggi disponibili, provenienti da un numero crescente di fonti non più solo interni all'organizzazione ma anche acquisiti esternamente.

In passato avevamo dati strutturati, tipicamente memorizzati in tabelle, oggi la maggior parte dei dati sono non strutturati (testi, immagini, voci, video). Si parla per tanto di **data integration**.

VERIDICITA'

E' possibile estrarre informazioni vere dai dati, nonostante essi contengano gravi errori, imprecisioni e incompletezze. Si parla di **data quality** come la tecnica per gestire la (mancanza di) veridicità che caratterizza molte raccolte di dati.

VARIABILITA'

Nella gestione di molti dati, in diversi formati e provenienti da diversi contesti, la mutevolezza del loro significato è un aspetto da tenere in considerazione nel momento in cui i dati vengono interpretati.



La scienza dei dati

La **Data Science** è tipicamente associata ai Big Data ed è una materia interdisciplinare, tra la statistica e l'informatica.

datacleaning: ha per obiettivo la costruzione di una raccolta dati che abbia un buon livello di qualità.

dataintegration: ha per obiettivo la costruzione di una raccolta dati integrata a partire da differenti sorgenti dati.

datamining: ovvero l'estrazione di informazione utile dai dati. Nel data mining sono molto utilizzate le regole di associazione, che consentono di dire quante volte, nel contesto di una specifica

operazione (ad esempio, una transazione di acquisto) sono coinvolte le stesse istanze di dati (ad esempio, gli stessi oggetti).

metodi predittivi: Ovvero metodi che consentono di prevedere, a partire da osservazioni nel passato, i dati caratterizzanti uno scenario futuro oppure alternativo. Tecniche statistiche consentono di selezionare i dati più utili (*feature selection*). Le predizioni vengono associate ad una probabilità.

Machinelearning o metodi di apprendimento automatico: sono particolari metodi predittivi per classificare i dati a partire da esempi di classificazione. Alcuni dati (*training set*) sono classificati a priori (ad esempio, da un esperto) e associati ad una etichetta (*label*) che ne indica la classe

Tra i metodi di machine learning assume sempre maggior rilevanza il cosiddetto *deep learning*, che consiste in un'imitazione del comportamento del cervello umano, in quanto dotato di moduli software che simulano i neuroni e le connessioni fra di essi.

Tecnologie per i Big data

Cloud Computing

pensati soprattutto per l'elaborazione batch di enormi mole di dati in parallelo, e sono scarsamente utilizzabili per rispondere ad interrogazioni online.

Sistemi NoSQL

rinunciano a gestire l'intera complessità del linguaggio SQL o le caratteristiche di piena transazionalità pur di gestire efficientemente alcune tipologie di dati con una semplice struttura. Consentono il calcolo parallelo (datacenter).

Perché non è possibile l'utilizzo di un database relazionale?

Perché i database relazionali hanno un approccio centralizzato e quindi possono scalare verticalmente ma non orizzontalmente. – Significa che, se aumentano utenti o dati, necessitano server più potenti: più CPU, più memoria, più memoria su disco. I server molto potenti sono molto complessi e costano molto.

Infatti i giganti del web, che si trovano a dover gestire database di dimensioni veramente imponenti, hanno sviluppato vari NRDBMS (**database non relazionali**) con approcci leggermente diversi, ma tutti con gli stessi principi di base, si raggruppano infatti sotto la dicitura di movimento NoSQL (Not Only SQL).