



Cofinanțat de
Uniunea Europeană



Proiectul Wallachia e-Hub este finanțat prin Digital Europe Programme (EC/101083410) - WeH și prin Programul Creștere Inteligentă, Digitalizare și Instrumente Financiare (POCIDIF/1147/2/1/161799)



MIHAI ANDRONIE,
TRAINER COORDONARE A TRANSFORMĂRII DIGITALE

APLICABILITATEA BIG DATA ÎN ADMINISTRAȚIA PUBLICĂ

WEH.SPIRUHARET.RO

APARIȚIA DATELOR MARI

Big data este unul dintre subiectele care a câștigat popularitate în ultimii ani, fiind rezultatul progreselor fără precedent în tehnologia informației și comunicațiilor. Cu toate acestea, chiar dacă subiectul este adesea discutat, se pare că majoritatea oamenilor au încă doar o idee generală despre ceea ce reprezintă big data, care sunt beneficiile sale, limitările și cum anume poate îmbunătăți activitățile de zi cu zi.

Big data, în general, reprezintă volume uriașe de date care nu sunt ușor de gestionat de instrumentele și tehnicile obișnuite de procesare a datelor. Astfel de seturi de date au nevoie de o infrastructură hardware/ software specială pentru a fi achiziționate, gestionate și procesate într-un interval de timp tolerabil.

De cele mai multe ori big data sunt compuse din diferite tipuri de date complexe și eterogene, care pot fi structurate, semi-structurate și chiar nestructurate. Printre astfel de tipuri de date pot fi:

- baze de date tabelare/ relaționale (care sunt de obicei bine structurate);
- depozite de date (integrate din multiple surse);
- date text;
- pagini web;
- imagini în diferite formate, audio, video etc.

Uneori, big data pot conține date duplicate, date care nu sunt întotdeauna exacte sau chiar date contradictorii. Big Data poate avea surse multiple, unele dintre cele mai cunoscute fiind rețelele sociale, imaginile, rețelele de senzori și web.

CE DIFERENȚIAZĂ BIG DATA DE DATELE OBIȘNUITE

Caracteristicile care diferențiază big data de datele obișnuite sunt acoperite de modelul 4V care evidențiază următoarele patru proprietăți ale big data:

- **Viteză** – reprezintă viteza care caracterizează big data, atât în ceea ce privește colectarea acestora, cât și perioada de valabilitate a acestora (de obicei scurtă);
- **Volum** – cantitatea de date mari este mai mult decât poate fi gestionată prin instrumente tradiționale de prelucrare a datelor; Big data necesită infrastructuri hardware-software speciale pentru a fi colectate, stocate și procesate;
- **Veracitate** – reprezintă proprietatea datelor mari de a nu fi total de încredere și exacte; din acest motiv, dacă instrumentele obișnuite de analiză a datelor vor fi utilizate pe big data, rezultatele pot fi eronate;
- **Variatate** – se referă la multitudinea de tipuri de date care urmează să fie întâlnite.

Potrivit altor surse, există și un al cincilea „V” care caracterizează big data, adică Valoarea și se referă la faptul că big data conține modele și cunoștințe

valoroase care pot fi extrase cu tehnici adecvate de data mining, aducând oportunități deținătorilor și alte părți care le pot accesa și exploata.

În practică, big data se referă la seturi de date de volum mare, complexe, în continuă creștere, cu surse multiple și autonome.

În ceea ce privește dimensiunea la care au ajuns datele mari, acestea pot fi de la TB la PB, de obicei mai mult decât poate stoca și procesa un computer obișnuit din cauza limitărilor de memorie și procesor. Din acest motiv, modelele tradiționale de procesare a datelor (cum ar fi stocarea datelor de intrare în memoria computerului înainte de procesarea lor) sunt inadecvate pentru manipularea datelor mari.

APLICAȚII BIG DATA ÎN ADMINISTRAȚIA PUBLICĂ

Deși aplicarea big data este încă limitată în sectorul public, utilizarea lor se dezvoltă dinamic, aducând rezultate tangibile în ceea ce privește îmbunătățirea eficienței și eficacității activităților desfășurate.

Deținătorii de volume mari de date care se pot încadra în categoria big data au posibilitatea de a le exploata în diferite scopuri, dintre care amintim în mod deosebit:

- **Supravegherea publică** – identificarea neregulilor (de exemplu, nerespectarea legii) și luarea de măsuri de răspuns:
 - Pregătirea unor seturi largi de date (big data) pentru a oferi baza de analiză;
 - Pregătirea modelelor digitale pentru nereguli;
 - Aplicarea modelelor la seturile de date mari, ceea ce presupune filtrarea datelor folosind modelele de nereguli;
 - Ca urmare a acestui proces, sistemul informatic selectează și prezintă autorităților datele care se conformează unui model de nereguli pentru a se decide asupra situației.
- **Reglementarea publică** – reglementarea comportamentului social și modelarea relațiilor sociale prin permise, interdicții sau ordine:
 - Metodele big data permit organismelor publice să monitorizeze starea unei zone administrate în timp real sau cu o scurtă întârziere de minute sau câteva ore;
 - Aceasta oferă autorităților o bază largă de informații pentru luarea deciziilor, permițând atingerea obiectivelor de reglementare.
- **Furnizarea de servicii publice** – eficientizarea furnizării anumitor servicii sau produse de interes public (inclusiv pe partea de infrastructură):
 - Analiza comportamentului clienților pentru a înțelege mai bine nevoile și a furniza servicii publice mai potrivite acestor nevoi (aceasta include și crearea de noi servicii).
 - Verificarea eligibilității pentru servicii publice, indemnizații etc. și depistarea fraudei.
 - Performanță operațională mai bună a furnizorului de servicii.