



Bloc 1 • Dades per tot arreu

Unitat 1 • Píndola 3

## **INTEL·LIGÈNCIA DE DADES: QUAN LES DADES ES TORNEN INTEL·LIGENTS**



**BIGDATA**

LA INTEL·LIGÈNCIA DE LES DADES

# INTEL·LIGÈNCIA DE DADES: QUAN LES DADES ES TORNEN INTEL·LIGENTS

Les dades i metadades que es troben a la xarxa no servrien gaire si no s'emmagatzemessin, es classifi-  
quessin i s'analitzessin. La disciplina que s'encarrega d'aquest procés rep el nom d'intel·ligència de  
dades o *Big Data*.

A partir d'aquestes macrodades, podem generar coneixement, productes i serveis. Per exemple, ens  
permeten predir el temps, analitzar paràmetres de salut, millorar l'eficiència energètica o vendre més i  
millor. Per això, la intel·ligència de dades està revolucionant la ciència, l'economia, la política i el nostre  
estil de vida.

S'utilitza la intel·ligència de dades quan:



Introduïm malament una paraula d'una cerca i Google ens  
la corregeix.



Amazon ens ensenya productes que podrien interessar-nos  
segons el que hem comprat abans.



Snapchat ens fa descobrir usuaris i notícies.



Facebook ens suggereix amistats amb gent que és probable  
que coneixem.



Spotify elabora una llista setmanal de cançons segons  
els nostres gustos.



# LES QUATRE VE BAIXES DE LA INTEL·LIGÈNCIA DE DADES

Aquestes primeres quatre ve baixes (més endavant en veurem alguna més) serveixen per recordar i descriure les característiques principals de la intel·ligència de dades: la gestió d'un gran volum de dades, a la velocitat més gran possible, emmagatzemats juntament amb una extensa varietat d'informació, que s'ha de verificar.

## VOLUM

el creixement de les dades a la xarxa és constant. Per això, es diu que «les dades mai no dormen». Tan és que sigui de dia o de nit: es generen dades les 24 hores a qualsevol part del món. Aquesta producció ininterrompuda es duplica cada 40 mesos, cosa que vol dir que es generen més dades en un dia de les que hi ha hagut en els darrers 20 anys.

## VELOCITAT

Imprescindible en la creació i anàlisi de dades. Tenint en compte que estem generant dades constantment, necessitem rapidesa per disposar d'aquesta informació en temps real. La informació es processa tan ràpidament que podem conèixer dades del present immediat, i consultar què està passant al món ara mateix. Això ens permet fer anàlisis prou detallats i complexos que sovint s'integren a altres processos de treball i sistemes.

## VARIETAT

Hi ha dades de molts formats i tipologies diferents, segons la seva procedència. Podem classificar les macrodades en:

### DADES PÚBLIQUES

Dades que tenen les administracions públiques (per exemple, dades sobre transport, ús d'energia, sanitat, etc.).

### DADES PRIVADES

Dades derivades de transaccions comercials, de la navegació web, de l'ús de la telefonia mòbil, etc.



## DADES COMUNITÀRIES

Dades produïdes principalment a les xarxes socials, continguts generats per l'usuari, etc.



Les dades obtingudes han de ser fiables, íntegres i autèntiques, per la qual cosa és necessari confirmar-ne la veracitat. I com sabem si les dades són vàlides? Això dependrà de les fonts i els recursos que hàgim utilitzat per obtenir-les.

## DADES AUTOQUANTIFICADES

Dades obtingudes i proporcionades per les mateixes persones que mesuren i quantifiquen els seus comportaments i accions. Per exemple, les dades monitorades sobre les pulsacions durant la realització de l'exercici físic que són recollides per dispositius mòbils.

