



# SECTOR

## L'ECONOMIA DE LES DADES



# INFORME DE LES PRINCIPALS TENDÈNCIES EN EL SECTOR DE L'ECONOMIA DE LES DADES EN EL MARC DE L'ECONOMIA METROPOLITANA

## CONTINGUTS DE L'INFORME

0. PRESENTACIÓ	3
1. EL SECTOR EN XIFRES	5
2. COMPOSICIÓ DEL SECTOR	8
3. PRINCIPALS TENDÈNCIES	11
4. PRINCIPALS REPTES	15
5. BIBLIOGRAFIA	19



# 0. Presentació

**INTRODUCCIÓ**

**LA CIÈNCIA DE LES DADES (DATA SCIENCE)**





## INTRODUCCIÓ

Des de fa anys, la majoria de nosaltres ha estat facilitant dades personals a les grans corporacions sense masses complexos, i inclús sovint sense adonar-se'n. Els gegants tecnològics mundials, com Amazon, Facebook o Google, s'han beneficiat enormement d'aquestes dades. En pocs anys, han esdevingut les empreses més grans i poderoses a escala mundial. Això ha estat gràcies principalment a l'ús i a l'explotació intensiva de tota mena de dades sobre nosaltres i el nostre comportament quotidià.

**Les dades esdevenen l'ànima de l'economia global i representen un nou tipus d'actiu econòmic clau**

El valor de les dades continuarà creixent en un món on cada vegada més es registren digitalment tots els moviments. Mentre que, paral·lelament, el nombre de dades i de dispositius digitals creix exponencialment a tot el món any rere any. Les organitzacions que aprenguin a treure profit de les dades i a utilitzar-les de la millor manera, obtindran un avantatge competitiu decisiu en aquest món interconnectat. En aquest context, la gestió i l'anàlisi de dades serà indispensable per a qualsevol organització, tant pública com privada, en el procés de presa de decisions.

Alguns dels principals avantatges que proporciona una bona gestió de les dades a les organitzacions són els següents [1]:

-  Augmentar-ne el rendiment i l'eficiència
-  Oferir productes i serveis més centrats en l'usuari
-  Fomentar la innovació
-  Diferenciar-se davant d'altres competidors

L'ús massiu de les aplicacions mòbils o la naturalitat i la rapidesa en la manera que s'estan introduint en el llenguatge comú paraules com "el núvol" -per referir-nos a l'espai on ciutadans i organitzacions emmagatzemem les dades-, dona una idea de fins a quin punt aquest sector està penetrant amb força a la societat actual.

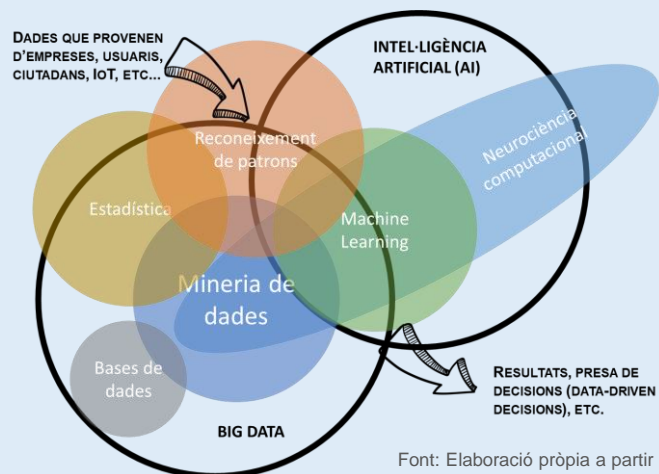
En aquest informe s'analitzen els efectes del sector de l'economia de les dades en l'actualitat i, com ja s'està avançant, el potencial impacte que tindrà en el futur. A l'apartat final de l'informe s'apunten alguns dels principals reptes, entre els quals els que ha d'afrontar el sector per resoldre algunes de les seves principals implicacions ètiques, socials i econòmiques a escala global.

## LA CIÈNCIA DE LES DADES (DATA SCIENCE)

Una de les principals potencialitats de l'economia de les dades, i que ja estan començant a explotar la majoria de les grans organitzacions, públiques i privades, és la possibilitat d'oferir un modelatge predictiu basat en la ciència de les dades.

Aquest modelatge recull una gran quantitat de dades històriques, identifica patrons de dades per dissenyar un model, i utilitza el model per fer prediccions sobre el que passarà en el futur. Això permet prendre decisions basades en les dades.

En aquest informe no s'aprofundeix en les disciplines i les eines que s'utilitzen per al desenvolupament d'aquest model. No obstant això, si que es presenten les dues principals disciplines del Data Science: el Big Data, basat en la mineria de dades, el tractament de bases de dades, l'ús de l'estadística i el reconeixement de patrons. I la Intel·ligència Artificial, que està en procés de plena expansió, basat en el Machine Learning i la neurociència computacional.



Font: Elaboració pròpia a partir de [2] i [3]

# 1. EL SECTOR EN XIFRES

EVOLUCIÓ I USOS DE LES DADES

APORTACIÓ DE VALOR DE L'ECONOMIA DE LES DADES

INVERSIÓ EN INFRAESTRUCTURA DE GESTIÓ DE DADES

## EVOLUCIÓ I USOS DE LES DADES

La quantitat d'informació generada diàriament a escala global, no només és de gran abast, sinó que està creixent a un ritme exponencial. La Figura 1 mostra com ha evolucionat l'emmagatzematge de dades al món entre 2005 i 2020.

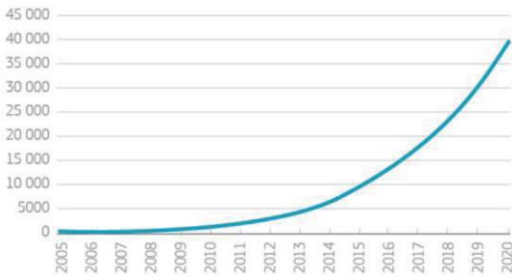


Figura 1. Evolució de l'emmagatzematge de dades mundial en exabytes [4]

**De tota la informació generada, s'estima que el 2020, s'obindrà informació de valor del 33% de l'univers digital un cop analitzada, comparat amb el 25% actual**

Aquesta evolució de la quantitat d'informació generada es deu principalment a l'increment del nombre d'aparells interconnectats a través de l'Internet of Things. Que es preveu que es multipliqui per 5 en només 6 anys, passant del 4,6 mil milions el 2015 als 16 mil milions el 2021 [5].

Els professionals de diferents sectors consideren que intensificaran l'ús de la gestió i anàlisi de dades en els propers anys, en diversos processos dins de les organitzacions. On ja en fan un ús més intensiu actualment és en l'àmbit del màrqueting. Ara bé, consideren que en els propers anys, on s'intensificarà més aquest ús, és en àmbits com l'anàlisi de riscos, l'optimització de processos o la predicció econòmica.

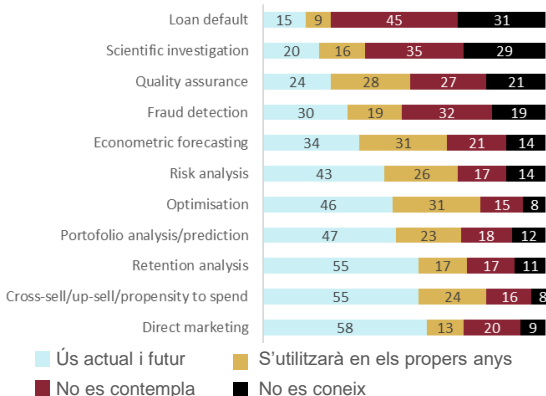


Figura 3. Ús de les dades per a la gestió i anàlisi en diferents processos a les organitzacions (%), a partir d'una enquesta global realitzada a càrrecs executius a l'any 2016 [6]

Pel que fa a les principals fonts d'on provenen la major part de les dades que es gestionen i s'analitzen, les més utilitzades actualment són les estructures de dades en forma de taules i registres. I continuarà sent així en els pròxims anys. Mentre que de moment, el tipus de fonts menys utilitzades per aquesta finalitat són les generades de forma automàtica per dispositius. Les altres fonts de dades més utilitzades són les demogràfiques, geoespaciales, sèries de dades, formularis/registres a través de la web, etc. La Figura 4 mostra les diferents fonts de dades utilitzades actualment i les que es preveuen que se n'incrementi l'ús en els pròxims anys

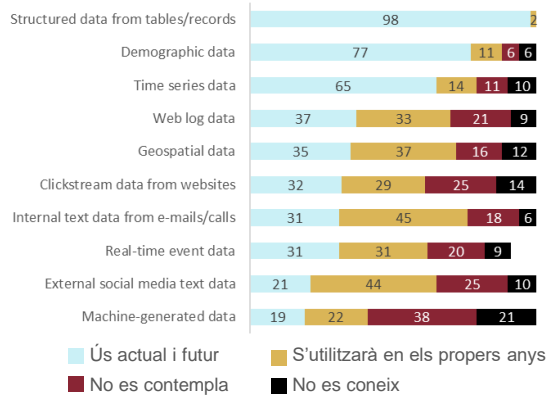


Figura 4. Ús de dades per a l'anàlisi predictiu (%), a partir d'una enquesta global realitzada a càrrecs executius a l'any 2016 [6]

L'Ajuntament de Barcelona ofereix un servei de dades obertes amb informació accessible per a la ciutadania, a través de bases de dades (432), sèries històriques (1.337) i altres recursos (2.152), classificats segons 5 temes (Administració, Ciutat i serveis, Població, Territori i Economia i empresa). Tal com mostra la Figura 5, el nombre de publicacions disponibles en forma de bases de dades ha anat incrementant en els darrers anys.



Figura 5. Evolució de les publicacions de bases de dades [7]

Les publicacions més visitades pels usuaris de la plataforma, són les que donen informació sobre: (1) les estacions de Bicing, (2) dutxes públiques a la ciutat, (3) incidències, queixes, suggeriments, consultes i agraïments gestionats per l'Ajuntament de Barcelona i (4) punts de recàrrega de vehicles elèctrics.

## APORTACIÓ DE VALOR DE L'ECONOMIA DE LES DADES

El mercat global de l'economia de les dades està experimentant un increment important des de fa anys, amb una previsió d'assolir els gairebé 50 mil milions d'euros aquest 2019, multiplicant per 10 la xifra de l'any 2010 [4].

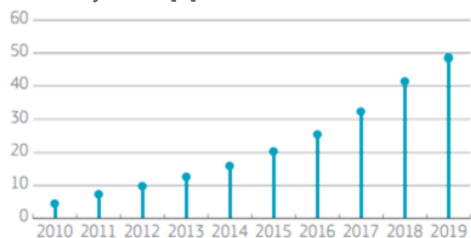


Figura 6. Evolució del mercat global de l'economia de les dades (€) [4]

Nombrosos estudis destaquen els beneficis econòmics que genera la gestió i l'ús intensiu de les dades. Es calcula que les organitzacions que adopten la presa de decisions basades en dades incrementen la seva producció i productivitat entre un 5-6% [8].

## INVERSIÓ EN INFRAESTRUCTURA DE GESTIÓ DE DADES

El país que l'any 2015 havia invertit més en infraestructura d'emmagatzematge de dades (data centers) va ser els Estats Units, amb més del 50% dels data centers dels països de l'OECD, seguit de lluny per el Canadà, Austràlia, França, Regne Unit, Alemanya i Índia [10].

En aquest sentit, des de la Unió Europea s'estan duent a terme diverses actuacions per a intensificar-ne el paper en infraestructura d'emmagatzematge de dades a escala global. Aquestes actuacions es consideren estratègiques per no perdre l'avantatge competitiu davant d'altres regions. La infraestructura per capturar, processar, emmagatzemar i analitzar grans volums de dades i generar-ne valor serà un recurs estratègic i es considera que una de les principals fonts de riquesa i de prosperitat futura.

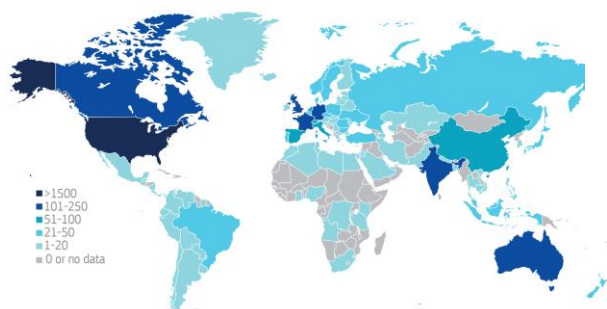


Figura 9: Mapa del nombre total de centres d'emmagatzematge de dades (Data Centers) per països, a tot el món (2015) [10]

Es preveu que aquest sector segueixi tenint un impacte important a l'economia en els pròxims anys. La seva repercussió té efectes en la majoria de sectors econòmics, especialment el de la indústria, però també en d'altres com els serveis professionals, les telecomunicacions, el retail o la banca. En el cas del Regne Unit per exemple, tal com mostra la Figura 7, l'impacte del sector de les dades a l'economia en els 5 anys del període comprès entre 2015-2020 es preveu que serà superior que el comprès entre 2012-2017.



Figura 7. Impacte econòmic de l'economia de les dades a diferents sectors entre els períodes 2012-2017 i 2015-2020 al Regne Unit [9]

Europa és la regió amb més fluxos de dades interregionals del món, com es pot observar en el mapa de la figura 10 [11]. I també la que intercanvia més dades amb altres regions, principalment amb Amèrica del Nord, seguit de lluny amb l'Orient Mitjà i Àfrica.

Amèrica del Nord és la regió amb més fluxos de dades amb altres regions, principalment Europa, Amèrica Llatina i Àsia, i en menor mesura amb Oceania.

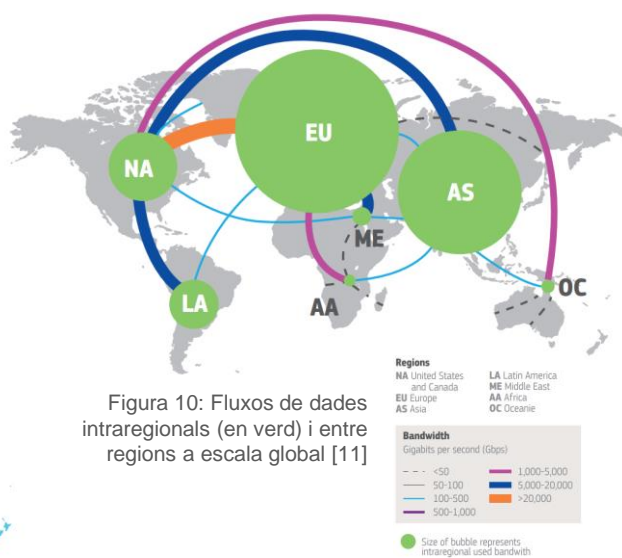
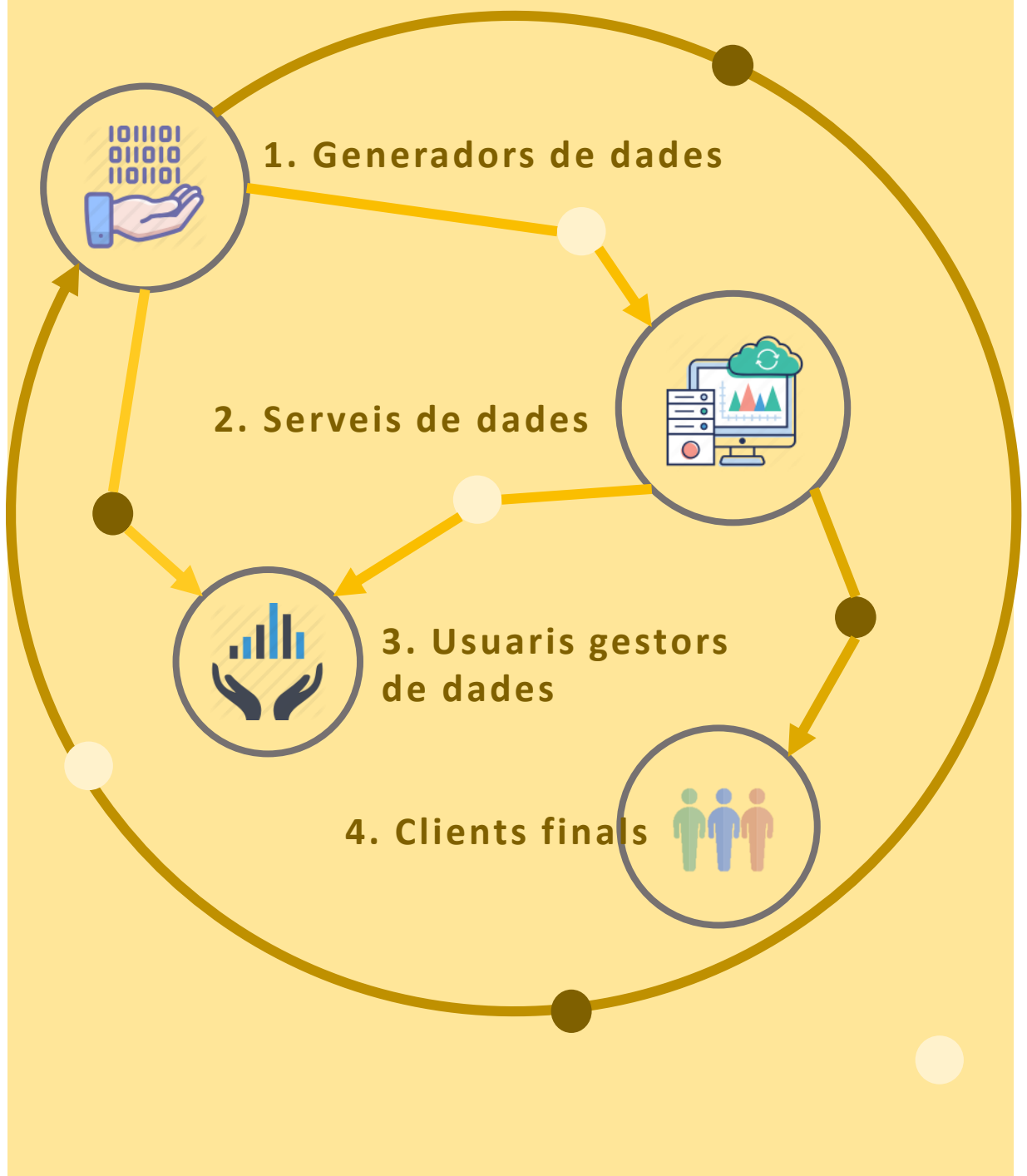


Figura 10: Fluxos de dades intraregionals (en verd) i entre regions a escala global [11]

## 2. COMPOSICIÓ DEL SECTOR



En aquest apartat es descriuen els quatre actors principals en el sector de l'economia de les dades: Els generadors de dades, els serveis de dades, els usuaris gestors de dades i els clients finals [12]. No obstant això, és important assenyalar que aquesta classificació es fa amb finalitats analítiques. En realitat, és difícil observar actors que només encaixin en una de les categories. El més habitual és que les categories es superposin, en la mesura que els generadors de dades poden ser també clients finals, o empreses que fan ús de l'anàlisi de dades per generar les seves pròpies dades i poden executar els seus propis serveis d'anàlisi de dades.



## 1. GENERADORS DE DADES

Són la principal font de dades: els usuaris que naveguen per Internet; els consumidors que paguen amb la seva targeta de crèdit; dades d'ubicació de lectures GPS o la funció d'una nevera intel·ligent per reduir el consum energètic.

Qualsevol acció o no acció per part d'éssers humans o dispositius elèctrics transmet informació que pot ser recopilada i agregada per a una anàlisi de dades. Les dades es poden generar de dues maneres:



### DE FORMA PASSIVA

Quan els subjectes que generen dades no són necessàriament conscients que la seva acció transmet la informació capturada per una altra entitat. Per exemple un conductor que utilitza una aplicació de GPS per desplaçar-se amb cotxe.



### DE FORMA ACTIVA

Quan els subjectes transmeten les dades de manera conscient. Per exemple, un usuari que facilita la seva adreça de correu electrònic en un formulari online per subscriure's als serveis que ofereix una xarxa social.

Aquesta distinció és rellevant per tal de determinar si el generador de dades ha donat el seu consentiment per al processament de les dades i per conèixer quin és el propòsit de la recopilació d'aquestes dades. Això és així, atès que la informació creada pot ser valuosa pels "serveis de dades", que és un altre dels actors del sector. I, els generadors de dades, poden aprofitar-ho per beneficiar-se'n i "monetitzar" aquesta informació.

Un exemple clàssic és el del model de negoci publicitari, que permet als consumidors utilitzar els serveis de forma gratuïta, mentre utilitzen la informació proporcionada per a finalitats de publicitat orientada.

En aquest sentit, els generadors de dades poden estar exposats a riscos, com ara violacions de la privacitat i la seguretat, fins i tot quan les dades s'emmagatzemen amb seguretat.



## 2. SERVEIS DE DADES

Són els qui generen la creació real del valor que es produeix quan es processa i s'analitzen les dades proporcionades pels generadors de dades. Això inclou la recopilació, l'enregistrament, l'organització, l'estructuració, l'emmagatzematge, l'adaptació o l'alteració, la consulta, l'ús, la divulgació per mitjà de la transmissió, la difusió o la posada a disposició, l'alineació o la combinació, la restricció, l'esborrament o la destrucció de les dades.

Es poden distingir principalment 4 etapes en la cadena de creació de valor de les dades:









En general, com més dades s'introdueixen a l'anàlisi, millor i més eficient serà el servei de dades. Per exemple, com més usuaris facin servir un motor de cerca web, més retroalimentació i més forta serà la capacitat del motor de cerca de retornar resultats que responen al que realment busquen els usuaris.



### 3. USUARIS GESTORS DE DADES

Aquestes són les empreses i les administracions públiques que utilitzen el resultat de l'anàlisi de dades per millorar el seu rendiment. Això pot incloure la millora del seguiment intern, l'adaptació de les estratègies de màrqueting, la reducció dels costos en diferents processos interns, el desplegament de nous productes o, fins i tot, canvis en el seu model de negoci.

Per tal d'incorporar l'anàlisi de conjunts de dades a la seva organització, les empreses i les administracions públiques han de superar una sèrie de barreres. Algunes de les més rellevants són les següents:

-  La manca de consciència del funcionament i dels beneficis potencials.
-  La resistència i la por al possible canvi organitzatiu que comporta l'anàlisi de grans dades.
-  L'absència de mitjans tecnològics disponibles, especialment pel que fa a la capacitat d'integrar i gestionar grans conjunts de dades.
-  L'escassetat de personal qualificat i expert en la informació.
-  La manca de mitjans financers per fer les inversions tecnològiques i de personal.
-  La manca de claredat en el marc normatiu en matèria de privacitat, protecció de dades i seguretat.



### 4. CLIENTS FINALS

L'últim grup d'actors afectats són els consumidors i les empreses o administracions públiques que han implementat l'anàlisi de grans dades a la seva organització.

L'efecte més directe de l'anàlisi de dades per a aquesta categoria són les possibles reduccions de preus, d'adquisició de serveis o productes de millor qualitat, incloent-hi la possibilitat d'accedir a productes nous que no existirien sense l'anàlisi de dades.

A més, també es pot beneficiar d'ofertes més adequades, basades en estratègies de màrqueting més precises. Les aplicacions que mapen el trànsit en temps real i permeten als conductors seleccionar la millor ruta per evitar la congestió de la carretera són un exemple d'això. Aquestes aplicacions es basen en un gran nombre de controladors que comparteixen informació sobre el trànsit durant la conducció. Sense grans dades no existiria aquest servei.

Tot i això, aquests avantatges també poden comportar inconvenients: perquè els clients finals són sovint també generadors de dades i estan exposats a riscos de seguretat i privacitat.

A continuació es mostra una taula amb els principals beneficis i riscos dels diferents actors de la cadena de generació de valor de les dades [12]:

	 Generadors de dades	 Serveis de dades	 Usuaris gestors de dades	 Clients finals
<b>Beneficis</b> 	Possibilitat d'utilitzar dades pròpies com a producte	Generació de beneficis	Augment de la productivitat Innovació incremental i disruptiva	Preus més baixos i de major qualitat, ofertes a mida, nous serveis / productes
<b>Riscos</b> 	Infraccions en seguretat i en privacitat	Colls d'ampolla en els fluxos de dades Costos de compliment del marc normatiu	Costos d'inversió davant la incertesa en la regulació Resistència o incapacitat per implementar l'anàlisi de grans bases dades	Discriminació de preus Infraccions en seguretat i en privacitat

# 3. TENDÈNCIES

## Privacitat de les dades

GESTIÓ DE LES DADES PERSONALS

DADES SINTÈTIQUES AMB INTEL·LIGÈNCIA ARTIFICIAL

BLOCKCHAIN PER A UNA IDENTITAT DIGITAL

## Open Data

IDEAMB, DADES GEOGRÀFIQUES OBERTES

KAGGLE, COMPETICIONS PER ANALISTES DE DADES

OPEN DATA PER A LA TRANSICIÓ A SMART CITIES

## Les dades a les empreses

CONSTRUCCIÓ DE DATA CENTERS EN AMBIENTS FREDS

OPTIMITZACIÓ DE PROCESSOS AMB BUSINESS INTELLIGENCE

ANALISTA DE DADES, UN PERFIL AMB DEMANDA CREIXENT

## PRIVACITAT DE LES DADES

Avui dia, les dades sobre hàbits i comportaments de la població són molt preuades per a les grans empreses a l'hora de prendre decisions comercials. Al mateix temps, cedir dades personals pot ser una amenaça per la privacitat de les persones. És per això que cada vegada més usuaris de les plataformes digitals reclamen mesures per garantir una gestió transparent i adequada de la informació personal que no vulneri els drets de privacitat.



### 1. Gestió de les dades personals

La quantitat d'informació generada pels usuaris a través de xarxes socials, cerques per internet, correus electrònics, etc. està creixent a un ritme vertiginós en els darrers anys [13]:

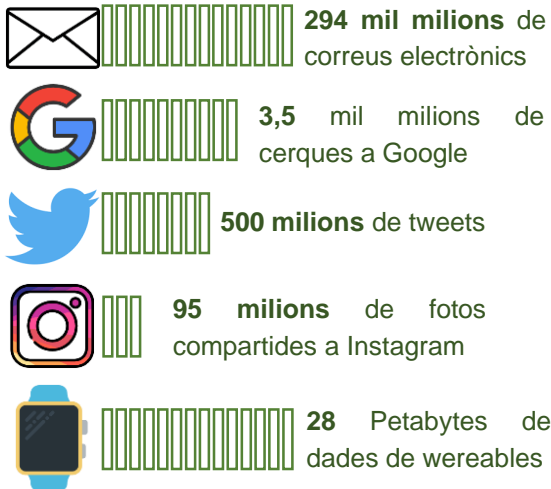


Figura 12: Quantitat de dades generades en un sol dia pels usuaris de la xarxa [13]

Existeixen diverses companyies dedicades a recopilar informació sobre els usuaris, a través de les seves cerques per internet i l'ús dels telèfons mòbils. Aquesta informació després és venuda a tercers. Davant d'això, estan apareixent companyies de "data brokerage" que ajuden a gestionar les dades personals dels usuaris amb la intenció de protegir-ne la privacitat. Algunes simplement permeten bloquejar l'accés a dades de diverses aplicacions. D'altres, permeten cedir les dades a canvi d'una recompensa econòmica. És el cas de l'aplicació Wibson [14], una companyia de Silicon Valley que ha creat un Marketplace entre agents interessats a obtenir dades i persones disposades a donar-les. D'aquesta manera, l'usuari és conscient de qui utilitza la seva informació.

### 2. Dades sintètiques amb intel·ligència artificial

En l'àmbit de la investigació és bàsic disposar de dades sobre les quals treballar. Especialment en l'ús de tècniques de Machine learning, on cal un gran nombre de dades per entrenar els algorismes. En l'àmbit mèdic, sovint és difícil obtenir dades sense posar en risc la privacitat dels pacients. És per això que diversos investigadors estan treballant amb mètodes d'intel·ligència artificial per crear algorismes capaços de replicar dades similars a les de pacients reals. D'aquesta forma, els metges poden investigar sobre dades representatives de la població, però que, en realitat, aquestes no pertanyen a ningú. Un dels grups de recerca en aquest àmbit pertany a la Universitat Politècnica de Catalunya [15].

### 3. Blockchain per a una identitat digital

La ciberseguretat és una disciplina bàsica per a protegir la informació personal en el món digital. Diverses companyies utilitzen tècniques d'criptació de dades per protegir els seus clients d'atacs informàtics que puguin posar en perill la seva privacitat. Un dels mètodes més utilitzats actualment i amb potencial de creixement durant els pròxims anys és el blockchain [16].

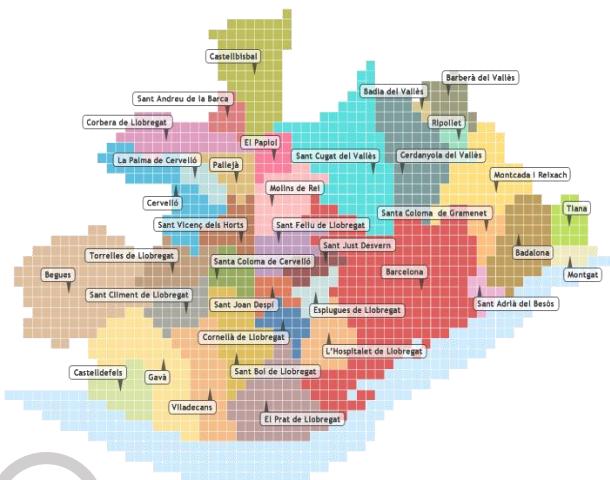
També des de les administracions s'està fent una aposta per l'ús de blockchain. La implementació de perfils d'identitat digital per a realitzar tràmits afavoreix l'exposició pública de dades. És per això que governs, com per exemple el de Catalunya [17], estan treballant en dissenyar plataformes segures i en les quals els ciutadans hi puguin confiar.

## OPEN DATA

Aprofitar la informació disponible gràcies al *Big Data* depèn de l'aparició d'aplicacions i de professionals que tinguin coneixement per gestionar-la. En aquest sentit, oferir dades obertes a la població permet facilitar l'accés a tothom qui ho desitgi i fomenta l'emprenedoria. L'*Open Data* permet crear una comunitat d'analistes i propicia el sorgiment d'iniciatives innovadores.

### 1. IDEAMB, dades geogràfiques obertes

IDEAMB és una plataforma d'infraestructura de dades espacials (IDE) d'accés obert, creada per l'Àrea Metropolitana de Barcelona. La plataforma actua com un sistema d'informació geogràfica (SIG) i permet obtenir tota mena d'informació sobre l'AMB. Aquesta informació pot ser interessant per diferents perfils professionals, especialment per geògrafs, enginyers i arquitectes. La iniciativa de l'AMB ajuda a la millora dels sistemes IDE actuals, tant del català (IDEC), com de l'espanyol (IDEE) i europeu (INSPIRE). Amb aquest projecte es pretén facilitar l'accés a aquest tipus de dades per tal d'incrementar la qualitat dels estudis territorials, així com d'altres projectes emprenedors [18].



### 2. Kaggle, competicions per a analistes de dades

Kaggle és una plataforma de competicions per analistes de dades, tant per a professionals com per a aficionats [19]. La companyia es va fundar el 2010 i el 2017 va ser adquirida per Google. Des de l'inici, Kaggle ha guanyat molta popularitat entre els analistes de dades i ha aconseguit crear una comunitat d'investigadors i experts en aquest àmbit. A la plataforma s'hi poden trobar projectes, espais de debat i bases de dades obertes per a tot



aquell qui desitgi realitzar experiments i recerca en un camp determinat. Ara bé, la popularitat de Kaggle prové dels concursos que organitzen les grans corporacions utilitzant la pàgina web. Des del sector automobilístic, amb empreses com Ford, al sector financer, amb companyies com Two Sigma, són diverses les empreses que han obert dades al públic i han proposat reptes per al públic en general. Els millors models proposats reben compensacions econòmiques, algunes de quantitats molt considerables.

### 3. Open Data per a la transició a smart cities

El concepte d'*smart city* es fonamenta sobre la capacitat d'utilitzar grans volums de dades per a millorar la gestió pública i adaptar-se a les necessitats particulars dels ciutadans. És per això, que moltes administracions estan optant per crear portals de dades obertes. Amb aquesta iniciativa es busca millorar la transparència dels governs i mantenir al ciutadà informat i, alhora, impulsar la innovació a partir del desenvolupament d'aplicacions. Com més accessibles siguin les dades, més fàcil serà desenvolupar tecnologies per a esdevenir una *smart city*. Concretament, Espanya és un país punter en bases de dades obertes. Segons l'informe de l'European Data Portal [20], Espanya és el segon país amb major nombre i qualitat de dades obertes per l'administració pública.

Una ciutat que més destaca en l'oferta d'*Open Data* actualment és Londres, amb més de 500 bases de dades lliures [21] en aspectes molt diversos. És per això que ja han sortit iniciatives per millorar la sostenibilitat del trànsit o augmentar la seguretat nocturna dels carrers. Nova York és una altra ciutat que està fent esforços per explotar el potencial de les seves dades. La ciutat celebra anualment la New York Open Data Week [22] que premia les millors aplicacions que fan ús de les dades lliures ofertes per la ciutat.

## Les dades a les empreses

L'anàlisi de dades és una tècnica beneficiosa per les empreses a l'hora de prendre decisions i estalviar recursos. Ara bé, també requereix la creació de nous espais de treball i de l'aparició de nous perfils professionals capaços d'afrontar els reptes futurs. Una bona gestió de les dades serà un factor diferencial en l'èxit de les companyies.

### 1. Construcció de data centers en ambients freds

L'augment dels sensors i la captació de dades implica directament un augment de la necessitat de capacitat d'emmagatzematge d'aquestes dades. Això comporta que les dimensions dels centres de dades hagin crescut i, paral·lelament, n'hagi crescut el consum energètic. Algunes companyies, com Google o Facebook [23], estan optant per construir els seus data centers en zones fredes com ara els països nòrdics o en el cercle polar. Darrere d'aquesta decisió hi ha diversos motius. Per una banda, és més fàcil trobar grans espais on construir aquests centres en comparació amb les grans ciutats. A més a més, l'ambient fred ajuda a reduir costos de refrigeració que requereixen aquest tipus d'infraestructura. I paral·lelament, en aquests països l'energia acostuma a ser barata i renovable.

No obstant això, l'elevat consum energètic d'aquests centres pot suposar una amenaça per a la sostenibilitat energètica en aquests indrets [24]. A més a més, la latència en les comunicacions i la distància als centres urbans, pot ser un problema per la gestió de les dades.

### 2. Optimització de processos amb business intelligence

L'interès en les dades per part de les companyies es basa en el seu ús a l'hora de prendre decisions estratègiques, el que es coneix com a business intelligence. La tendència en l'ús d'aquests processos ha crescut amb els anys i nombroses empreses ja basen gran part de les seves decisions comercials en estudis de Big Data. Aquesta pràctica propicia un consum més eficient dels recursos i ajuda a reduir despeses en els processos.

El servei de gestió mèdica del Regne Unit (NHS), per exemple, va obtenir un estalvi de gairebé 18.000 milions d'euros a través de la gestió i l'anàlisi de les dades de la població [25].



### 3. Analista de dades, un perfil amb demanda creixent

La revolució de les dades queda reflectida en l'anàlisi del mercat laboral. Les necessitats d'entendre, gestionar i mostrejar la informació existent, ha comportat l'aparició de nous perfils professionals coneguts com a analistes o científics de dades. La seva demanda és creixent en els darrers anys. Espanya, és el 8è país amb més professionals amb aquest perfil, i es van incrementant a un ritme del 50% anual, amb perspectives que segueixi així durant els pròxims anys [26]. A més a més, és una professió de les més ben pagades i amb major flexibilitat laboral.

El perfil d'un analista de dades és un perfil transversal [27]. Acostuma a ser una persona amb un nivell d'estudis elevat i amb formació en matemàtiques, estadística i programació. A més a més, en l'anàlisi de dades es treballa en diferents softwares i llenguatges, per tant, és necessari un coneixement ampli dels mètodes i tecnologies d'anàlisi.



**40% de professionals amb estudis de màster o doctorat**

**700.000 vacants laborals a l'estat l'any 2019**



**28% de creixement en la demanda de data analyst l'any 2020**

**59% de millora en la presa de decisions empresarials**



Característiques i perfil professional de l'analista de dades [28] [29]

# 4. PRINCIPALS REPTES

**REpte 1: DECISIONS BASADES EN LES DADES I L'EVIDÈNCIA**

**REpte 2: PASSAR DELS BENEFICIS PRIVATS AL VALOR SOCIAL DE LES DADES**

**REpte 3: SOBIRANIA EN LA IDENTITAT DIGITAL**

**REpte 4: PRINCIPALS REPTES PER A LES ORGANITZACIONS**

**REpte 5: DEMANDA DE NOUS PEFILS PROFESSIONALS**

## REpte 1: DECISIONS BASADES EN LES DADES I L'EVIDÈNCIA

El pluralisme a les nostres societats ha anat augmentant en els darrers anys. I la presa de decisions basada en la investigació tradicional disminueix en una realitat on les qüestions i els contextos estan cada vegada més interconnectats, canviants, desordenats, complexos i desconeguts. Sovint es prenen decisions precipitades, sense aprofundir en la diagnosi, per compensar la falta de competència en un tema o una falta de temps per dedicar a entendre'l realment.

*En la realitat actual les coses no són familiars, són incertes i són desconegudes. I, al mateix temps, la capacitat d'obtenir habilitats d'interpretació sòlides requereix un coneixement precís*

També s'està vivint en una realitat de novetats contínues, amb coses que mai s'han vist i que es desconeix el paper que jugaran. En aquesta realitat, la política, que per naturalesa és una pràctica orientada al futur, té com a objectiu donar forma a un futur que es considera desitjable.

L'aplicació de la previsió -l'avaluació del que pot passar o es necessitarà en el futur- és essencial. D'aquesta manera es podran dissenyar estratègies i decisions robustes i resistents davant de la incertesa per poder construir un futur comú desitjat. A priori, la previsió i el futur no contenen dades, únicament permeten utilitzar hipòtesis, expectatives, possibilitats i preferències.

*En aquesta nova realitat, es pot esperar que l'evidència esdevingui més diversificada i esbiaixada*

Això pot contribuir al fet que cada vegada més es prenguin decisions basades en valors i no tant en l'evidència. Les dades poden esdevenir una eina potent que permetin qüestionar supòsits, i a estar més oberts a aprendre de comparar i contrastar punts de vista del futur i perfeccionar les habilitats de pensament crític. En aquest context, l'increment de l'ús de les dades empíriques en el disseny de polítiques i en el procés de presa de decisions esdevé essencial [30].

## REpte 2: PASSAR DELS BENEFICIS PRIVATS AL VALOR SOCIAL DE LES DADES

Les administracions públiques són conscients de la importància de fer avançar l'economia de les dades a l'hora de prioritzar determinades estratègies. Per un costat perquè és poc probable que les forces del mercat per si soles condueixin a un ús socialment òptim de les dades. I per l'altre perquè la manca de transparència i de regulació en la gestió de les dades, genera una percepció d'inseguretat a empreses i ciutadans quan atempten a la protecció de la seva privadesa [31].

*Els generadors de dades i els clients sovint no poden apreciar i anticipar fins a quin punt estan exposats a danys potencials a causa de la informació que proporcionen (possiblement sense adonar-se'n).*

En aquest sentit, les entitats que promouen les polítiques d'open data des de les administracions proposen que, per exemple, les dades que s'han recopilat per a finalitats de màrqueting puguin ser reutilitzades per millorar la investigació en atenció sanitària, sense cap cost addicional. Consideren que sense tenir en compte aquests efectes derivats potencials, l'anàlisi de dades produïda a

l'economia, en el millor dels casos, farà incrementar els beneficis privats de les empreses i no el valor social [32].

Els reguladors han de buscar l'equilibri entre fer evolucionar l'economia cap a una societat digital i alhora garantir les normes de protecció de dades. La primera normativa aprovada a escala europea en aquest sentit és el Reglament general de protecció de dades (GDPR) aprovat el passat 25 de maig de 2018.

A Barcelona existeix la iniciativa Open Data Institute (ODI) que és una associació que neix per promocionar les dades obertes a l'entorn de Barcelona. Aquesta iniciativa pretén crear un Hub de coneixement entorn de l'Open Data, per transformar les dades obertes en valor a l'àmbit de Barcelona. Es basa en 5 eixos de treball: xarxa, divulgació, formació, investigació i empreudoria amb dades obertes [33].



## REpte 3: SOBIRANIA EN LA IDENTITAT DIGITAL

Actualment ja existeix la tecnologia i els coneixements per tal que no sigui que institucions centralitzades, com Facebook, Google o les autoritats públiques, registrin les nostres dades personals per separat a les seves bases de dades, que, per cert, són una manera ineficient i vulnerable de fer-ho. Aquesta vulnerabilitat particular del model centralitzat, queda il·lustrada amb la gran quantitat de fuites de dades i casos de mal ús de dades [1].

Per això, és necessari avançar en els avantatges distribuïts que aporta la tecnologia del blockchain i la criptografia avançada. Aquesta tecnologia permetrà facilitar únicament les dades específiques necessàries per aquells que han de validar-les en un context determinat per fer ús dels seus serveis. Amb aquest sistema no serà necessari proporcionar més dades de les imprescindibles pel tipus de producte o servei que es vulgui adquirir.



Per exemple, Amazon recull una gran quantitat de dades personals, com ara dades de pagament, adreça i altres dades personals, abans de poder fer una comanda. Això els permet enllaçar totes les altres dades amb la identitat del client. Però per vendre un llibre, Amazon únicament necessita saber que el client és una persona real i que és capaç de pagar. Amb una identitat autosobirana, el sistema pot garantir que el client és una persona real i que ha dipositat el preu del producte a un servei de pagament. Per altra banda, l'adreça postal del client se li facilitarà únicament al missatger, ja que aquesta informació és innecessària i irrellevant pel servei que ofereix Amazon.

## REpte 4: PRINCIPALS REPTES PER A LES ORGANITZACIONS

A partir de la tendència actual, es preveu que en els pròxims anys el volum de dades que es generaran a tot el planeta incrementarà de forma exponencial. Per tant, el tractament intel·ligent de les dades per part de les organitzacions serà clau per millorar l'eficiència dels seus serveis, i per poder oferir una major aportació de valor a les persones.

Actualment, disposar d'una eina de gestió i anàlisi de dades esdevé ja una necessitat per a qualsevol organització. La complexitat d'aquesta eina variarà en funció de la seva mida i del seu sector [34]. Més enllà de les eines de gestió i anàlisi de dades, un dels principals reptes en els pròxims anys, és

veure quines accions s'han de posar en marxa perquè empreses i institucions obtinguin el màxim valor de l'ingent volum de dades que es generaran.

El primer que cal és emmagatzemar-les i, posteriorment, és necessari disposar del coneixement i la capacitat per poder processar-les amb certa eficiència, de tal manera que puguin ser analitzades.

Per aconseguir això, cal que les organitzacions s'adaptin a aquest nou entorn digital i es transformin internament, el que comporta abordar els següents punts [2]:



### TRANSFORMACIÓ CULTURAL EN ELS EQUIPS

La resistència al canvi acostuma a ser el gran taló d'Aquil·les de moltes organitzacions. Fomentar l'adopció de noves capacitats digitals per als integrants de l'organització és un element clau per aconseguir l'èxit en la transformació digital.



### INVERSIÓ EN TECNOLOGIA

L'era del Big Data proporciona l'oportunitat d'afrontar amb certes garanties el repte d'un nou entorn. Amb més volum de dades, major varietat en la seva tipologia i major velocitat tant en la generació com en la necessitat de processament.



### ADOPCIÓ DE NOUS PERFILS

Per poder abordar amb garanties la cadena de valor de les dades, resulta imprescindible comptar amb nous perfils dins de l'organització, perfils que es veuran amb més detall a continuació.

## REpte 5: DEMANDA DE NOUS PEFILS PROFESSIONALS

La revolució digital que suposa el desenvolupament de l'Economia de les Dades ha obert el debat sobre l'impacte que exercirà en l'ocupació. En la mesura que les màquines puguin reemplaçar part o la totalitat d'algunes tasques desenvolupades actualment per treballadors, gràcies al processament d'informació de manera instantània (machine learning) o als nous avenços en la producció industrial (com la impressora 3D o el vehicle autònom), resulta inevitable pensar que bona part dels llocs de treball d'avui desapareixeran o, com a mínim, es transformaran [35].

No obstant això, la revolució digital també obre un ventall divers d'oportunitats laborals que és convenient conèixer i aprofitar. Entorn de les noves àrees de coneixement digital i l'explotació de les potencialitats que ofereix, emergeixen noves professions digitals, cada vegada més especialitzades. Aquestes professions es poden classificar segons les següents àrees de treball [36]:

La demanda d'aquests perfils digitals no només ha crescut de forma exponencial en els últims anys, sinó que exercirà un paper rellevant en la nova estructura del mercat de treball. El 2016, segons les estimacions d'IDC i Open Evidence de la Comissió Europea, hi havia 356.000 treballadors a Espanya exercint tasques associades a l'Economia de les Dades en pràcticament tots els sectors econòmics. La seva penetració representava el 2,3% en termes relatius al total de treballadors de cada sector, valor inferior a la mitjana de la UE-28 (3,1%) [37].

En aquest sentit, altres estudis constaten una manca important de professionals qualificats que satisfacin les necessitats de les empreses. Això es produeix especialment en el cas dels desenvolupadors d'aplicacions mòbils, solucions de Big Data i especialistes en ciberseguretat. Aquesta manca de professionals exigeix un reforçament del sistema educatiu i de les polítiques actives d'ocupació, sobre la base de les habilitats requerides per aquests nous llocs de treball [38].



### MÀRQUETING DIGITAL

- Responsable de màrqueting digital
- Expert en màrqueting
- Expert en SEM i SEO
- Responsable de relació digital amb clients
- Responsable de conversió i captació Programàtica
- Responsable de trànsit web



### BIG DATA I BUSINESS ANALYTICS

- Director de dades.
- Científic de dades.
- Enginyer de dades
- Responsable de visualització de dades de negoci
- Analista d'intel·ligència de clients



### SOCIAL MEDIA

- Responsable de gestió de comunitats i xarxes socials
- Responsable de l'estratègia de mitjans socials



### ESTRATÈGIA DIGITAL

- Responsable d'estratègia digital
- Responsable de comunicació digital
- Cap de projectes digitals



### CONTINGUT DIGITAL

- Expert en contingut digital de marca
- Responsable de contingut digital
- Expert en redacció publicitària



### E-COMMERCE

- Responsable de comerç electrònic
- Expert en anàlítica per comerç electrònic



### RECURSOS HUMANS

- Responsable de reclutament digital
- Analista de recursos humans



### DIGITAL SALES

- Executiu de comptes digitals



### CIBERSEGURETAT

- Director de seguretat de la informació
- Delegat de protecció de dades.

# PRINCIPALS REFERÈNCIES BIBLIOGRÀFIQUES

- [1] SCENARIO. Fuglsang, S. CAN WE TAKE BACK CONTROL OF OUR OWN DATA? (Març 2019) [en línia]. Disponible a: <http://www.scenariomagazine.com/can-we-take-back-control-of-our-own-data/>
- [2] FUNDACIÓN TELEFÓNICA. Economía de los Datos Riqueza 4.0 (2017) [en línia]. Disponible a: <https://www.fundacioncarolina.es/wp-content/uploads/2018/11/Libro-Economia-de-los-Datos-Ontiveros.pdf>
- [3] WILEY & SAS BUSINESS SERIES. Dean, J. Big Data, Data Mining, and Machine Learning (2014) [en línia]. Disponible a: [https://doc.lagout.org/Others/Data%20Mining/Big%20Data%2C%20Data%20Mining%2C%20and%20Machine%20Learning\\_%20Value%20Creation%20for%20Business%20Leaders%20and%20Practitioners%20%5BDean%202014-05-27%5D.pdf](https://doc.lagout.org/Others/Data%20Mining/Big%20Data%2C%20Data%20Mining%2C%20and%20Machine%20Learning_%20Value%20Creation%20for%20Business%20Leaders%20and%20Practitioners%20%5BDean%202014-05-27%5D.pdf)
- [4] IDC. The Digital Unvers in 2020: Big Data, Bigger Digital, and Biggest Growth in the Far East (2012) [en línia]. Disponible a: <https://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf>
- [5] ERICSSON MOBILITY REPORT. On the pulse of the networked society (Juny 2016) [en línia]. Disponible a: [https://www.abc.es/gestordocumental/uploads/internacional/EMR\\_June\\_2016\\_D5%201.pdf](https://www.abc.es/gestordocumental/uploads/internacional/EMR_June_2016_D5%201.pdf)
- [6] RACONTEUR. Data Economy (2016) [en línia]. Disponible a: <https://raconteur.uberflip.com/i/734244-data-economy-2016>
- [7] AJUNTAMENT DE BARCELONA. Servei de dades obertes de l'Ajuntament de Barcelona, Open data BCN, estadístiques (2019) [en línia]. Disponible a: <https://opendata-ajuntament.barcelona.cat/ca/analytics>
- [8] MIT CENTER FOR DIGITAL BUSINESS. Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? (2011) [en línia]. Disponible a: [http://ebusiness.mit.edu/research/papers/2011.12\\_Brynjolfsson\\_Hitt\\_Kim\\_Strength%20in%20Numbers\\_302.pdf](http://ebusiness.mit.edu/research/papers/2011.12_Brynjolfsson_Hitt_Kim_Strength%20in%20Numbers_302.pdf)
- [9] SAS. The Value of Big Data and the Internet of Things to the UK economy (febrer 2016) [en línia]. Disponible a: [https://www.sas.com/content/dam/SAS/en\\_gb/doc/analystreport/cebr-value-of-big-data.pdf](https://www.sas.com/content/dam/SAS/en_gb/doc/analystreport/cebr-value-of-big-data.pdf)
- [10] DATA CENTER MAP. Cloud Centers [en línia]. Disponible a: <https://www.datacentermap.com/cloud.html>
- [11] MCKINSEY GLOBAL INSTITUTE ANALYSIS. Digital Globalization: The new era of global flows (2016) [en línia]. Disponible a: <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Digital%20globalization%20The%20new%20era%20of%20global%20flows/MGI-Digital-globalization-Full-report.aspx>
- [12] EUROPEAN COMMISSION. Enter the Data Economy, EU policies for a thriving data ecosystem (2016) [en línia]. Disponible a: [https://ec.europa.eu/epsc/sites/epsc/files/strategic\\_note\\_issue\\_21.pdf](https://ec.europa.eu/epsc/sites/epsc/files/strategic_note_issue_21.pdf)
- [13] VISUAL CAPITALIST. Desjardins, J. How much data is generated each day (15/4/2019) [en línia]. Disponible a: <https://www.visualcapitalist.com/wp-content/uploads/2019/04/data-generated-each-day-wide.html>
- [14] WIBSON. Homepage (2019) [en línia]. Disponible a: <https://wibson.org/about-spanish/?lang=es>
- [15] TOWARDS DATA SCIENCE. Tirthajyoti Sarkar. Synthetic data generation – a must-have skill for new data scientists (12/2019) [en línia]. Disponible a: <https://towardsdatascience.com/synthetic-data-generation-a-must-have-skill-for-new-data-scientists-915896c0c1ae>
- [16] DISRUPTOR DAILY. Mire, S. Blockchain in cybersecurity: 11 Startups to watch in 2019 (8/2/2019) [en línia]. Disponible a: <https://www.disruptordaily.com/blockchain-startups-cyber-security/>
- [17] GENERALITAT DE CATALUNYA. Catalunya aposta per la tecnologia 'blockchain' (16/10/2018) [en línia]. Disponible a: [https://web.gencat.cat/ca/actualitat/detall/20181016\\_Catalunya-aposta-per-la-tecnologia-blockchain](https://web.gencat.cat/ca/actualitat/detall/20181016_Catalunya-aposta-per-la-tecnologia-blockchain)
- [18] CONFERENCIA ESRI. González González, JC. IDEAMB – Implementación de ArcGIS para la publicación e inventariado de datos geográficos (25/10/2018) [en línia]. Disponible a: <https://conferencia.esri.es/ponencia/ideamb-implementacion-arcgis-la-publicacion-e-inventariado-datos-geograficos/>

# PRINCIPALS REFERÈNCIES BIBLIOGRÀFIQUES

- [19] KAGGLE. Competitions [en línia]. Disponible a:  
<https://www.kaggle.com/competitions>
- [20] EUROPEAN DATA PORTAL. Cecconi, G. i Radu, C. Open Data Maturity in Europe (11/2018) [en línia]. Disponible a:  
[https://www.europeandataportal.eu/sites/default/files/edp\\_landscaping\\_insight\\_report\\_n4\\_2018.pdf](https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf)
- [21] THE GUARDIAN, Card, J. Open data is at the centre of London's transition into a smart city (3/8/2015) [en línia]. Disponible a:  
<https://www.theguardian.com/media-network/2015/aug/03/open-data-london-smart-city-privacy>
- [22] NYC OPEN DATA. Homepage (2019) [en línia]. Disponible a:  
<https://www.open-data.nyc/>
- [23] THE GUARDIAN. Gersmann, H. Facebook builds 'green' datacentre in Sweden (27/10/2011) [en línia]. Disponible a:  
<https://www.theguardian.com/environment/2011/oct/27/facebook-green-datacentre-sweden-renewables>
- [24] MIT TECHNOLOGY REVIEW. Adalbjornsson, T. Iceland's data centers are booming – here's why that's a problem (18/6/2019) [en línia]. Disponible a:  
<https://www.technologyreview.com/s/613779/icelands-data-centers-are-booming-heres-why-thats-a-problem/>
- [25] PRESCRIBING ANALYTICS. NHS efficiency savings: the role of prescribing analytics (2011) [en línia]. Disponible a:  
<http://www.prescribinganalytics.com/>
- [26] UOC. Sierra, G. Científico de datos, la tercera profesión más solicitada en España (10/04/2018) [en línia]. Disponible a:  
<https://www.uoc.edu/portal/es/news/actualitat/2018/082-cientifico-datos.html>
- [27] BBVA. Emerging Tech: el perfil de un científico de datos (15/7/2015) [en línia]. Disponible a:  
<https://bbvaopen4u.com/es/actualidad/emerging-tech-el-perfil-de-un-cientifico-de-datos>
- [28] EL ECONOMISTA. Demanda de expertos en análisis de datos crecerá 28% en 2020 (24/4/2019) [en línia]. Disponible a:  
<https://www.eleconomista.es/ecoaula/noticias/9838179/04/19/Demanda-de-expertos-en-analisis-de-datos-crecera-28-en-2020.html>
- [29] VISUALLY. The growing priority and urgency for Big Data Analytics (9/7/2015) [en línia]. Disponible a:  
<https://visual.ly/community/infographic/business/urgency-big-data-analytics>
- [30] EUROPEAN STRATEGY AND POLICY ANALYSIS SYSTEM. Global Trends to 2030, Shaping the future in a fast-changing world (Novembre 2018) [en línia]. Disponible a:  
[https://ec.europa.eu/epsc/sites/epsc/files/espas\\_2018\\_-\\_thinkpieces\\_web.pdf](https://ec.europa.eu/epsc/sites/epsc/files/espas_2018_-_thinkpieces_web.pdf)
- [31] KPMG. Fernández, A. Clientes: el valor de los datos y la confianza (Gener 2019) [en línia]. Disponible a:  
<https://www.tendencias.kpmg.es/2018/12/clientes-datos-compartir-privacidad-confianza/>
- [32] EUROPEAN POLITICAL STRATEGY CENTRE. Enter the Data Economy (Gener 2017) [en línia]. Disponible a:  
[https://ec.europa.eu/epsc/sites/epsc/files/strategic\\_note\\_issue\\_21.pdf](https://ec.europa.eu/epsc/sites/epsc/files/strategic_note_issue_21.pdf)
- [33] BCN INICIATIVA OPEN DATA. Muñoz, L. 50 Raons Open Data (2019) [en línia]. Disponible a:  
<https://iniciativabarcelonaopendata.cat/ca/50-raons-open-data/>
- [34] MCKINSEY&COMPANY. Booth, A. Hart, J. i Sim, S. Building a great data platform (Agost 2018) [en línia]. Disponible a:  
<https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/building-a-great-data-platform>
- [35] FUNDACIÓN TELEFÓNICA. Turégano, A. El auge de las nuevas profesiones (2017) [en línia]. Disponible a:  
<https://www.fundacioncarolina.es/wp-content/uploads/2018/11/Libro-Economia-de-los-Datos-Ontiveros.pdf>
- [36] INESDI. Top 25 profesiones digitales (2017) [en línia]. Disponible a:  
<https://www.inesdi.com/descargas/Top-25-Profesiones-Digitales-2017.pdf>
- [37] IDC i OPEN EVIDENCE. European Data Market SMART (Febrer 2017) [en línia]. Disponible a:  
[https://www.key4biz.it/wp-content/uploads/2018/04/SMART20130063\\_Final-Report\\_030417\\_2.pdf](https://www.key4biz.it/wp-content/uploads/2018/04/SMART20130063_Final-Report_030417_2.pdf)
- [38] INFOJOBS i ESADE. Estado del mercado laboral en España (2018) [en línia]. Disponible a:  
[https://nosotros.infojobs.net/wp-content/uploads/2019/05/Informe\\_Mercado\\_Laboral\\_InfoJobs\\_ESADE\\_2018.pdf](https://nosotros.infojobs.net/wp-content/uploads/2019/05/Informe_Mercado_Laboral_InfoJobs_ESADE_2018.pdf)