



Com esdevenir una organització *Data Driven* *Aprofita les dades per ser més competitiu*

White Paper #1

El·laborat per: Socis promotors:



Autors:

Marc Blasi: Cap de projectes Sènior Big Data CoE - Eurecat

Marc Torrent: Director del Big Data CoE Barcelona

ÍNDEX

1. INTRODUCCIÓ	3
2. PERSONES I ORGANITZACIÓ	4
2.1 ORGANITZACIÓ	4
2.1 ROL I RESPONSABILITATS.....	5
3. GOVERN DE LES DADES I PROCESSOS	8
3.1 ESTRATÈGIA DE DADES	8
3.2 GOVERN DE DADES	8
3.3 DADES I PROJECTES BIG DATA	9
3.4 QUADRES DE COMANDAMENT	10
4. TECNOLOGIA	11
4.1 CENTRALITZACIÓ	11
4.2 DEMOCRATITZACIÓ DE LES DADES	11
5. CONCLUSIONS	12
6. REFERÈNCIES	14

1. INTRODUCCIÓ

Des d'un punt de vista tecnològic, el Big Data està format per un conjunt de tecnologies i tècniques que permeten extreure informació valuosa de grans volums de dades, de fonts molts diverses, d'una manera àgil i eficient. Segons defineix Gartner en el seu glossari IT, el Big Data *'is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation'*. El Big Data permet, per tant, aportar nous insights de valor a tot tipus d'organitzacions, i millorar la seva productivitat i competitivitat a través de millors eines per a la presa de decisions, l'optimització de processos i un millor coneixement dels clients i del mercat. En conseqüència, en els darrers anys hem estat testimonis de com el terme **Big Data ha deixat de ser considerat únicament un terme associat al domini tecnològic i està sent adoptat amb una visió més global i transversal en aquelles organitzacions que estan incorporant les dades com un actiu estratègic**. Aquestes organitzacions estan redefint els seus processos per a orientar-los a la presa de decisió més estrictes basats en l'anàlisi rigorosa de les dades.

Per definició, una organització 'Data Driven' és conscient del gran valor que pot aportar a l'organització el processament rigorós de la informació interna i externa i la implantació d'una estratègia de dades que impliqui a tots els departaments i les persones de l'organització, millorant la presa de decisions i afavorint la innovació.

Existeix, a més, un subgrup que cal emfatitzar (normalment grans corporacions o petites empreses innovadores) amb un propòsit clar i un objectiu d'innovar utilitzant l'analítica avançada de les dades i les eines i tècniques associades. Aquestes organitzacions no només tracten d'aplicar les solucions existents d'analítica sinó que també dissenyen solucions concretes, sovint innovadores, per l'anàlisi i millora dels seus processos de negoci. Aquest tipus d'organitzacions es denominen *'innovation aimed data-driven organizations'*.

La cultura de dades és la base per esdevenir una organització *'data driven'*, i per tal de promoure-la cal centrar-se en els següents dimensions:

1. **Persones i organització**, per tal de disposar de les habilitats i responsabilitats necessàries.
2. **Govern de les dades i processos** associats per que defineixen com gestionar-les i optimitzar la seva explotació.
3. **Tecnologia** per facilitar l'accés i l'ús de les dades.

La resta del document, tot i que esperem porti informació útil per qualsevol tipus d'organització i professional de les dades, se centra principalment en les que hem anomenat *'innovation aimed data driven organizations'*, i no en aquelles organitzacions que simplement emmagatzemen la informació i utilitzen eines estàndards per a la seva explotació.

2. PERSONES I ORGANITZACIÓ

2.1 ORGANITZACIÓ

Segons Gartner¹, definir una organització al voltant de les dades i el Big Data està relacionat amb el concepte de 'TI bimodal' que permet desenvolupar i subministrar serveis de TI tenint en compte el següents dos modes de treball:

- **El Mode 1 - Tradicional**, que corresponen a projectes TI guiats i amb temps d'implementació relativament llargs. Aquests estan basats principalment en els requeriments, la seguretat, la mitigació de riscos, l'eficiència i la precisió. S'engloben en aquest mode les implementacions de DWH i de Business Intelligence tradicionals, i els desplegaments Big Data en producció (un cop validat el seu valor per l'organització).
- **El Mode 2 - Exploratori**, basat principalment en la incertesa, l'agilitat i la velocitat, on els desenvolupaments són iteratius i es basen en la descoberta d'insights de valor de negoci a través de l'anàlisi de dades. Aquest mode es pot dur a terme identificant les necessitats de negoci (enfocament deductiu) o a partir de l'anàlisi de les dades en brut per tal d'identificar patrons per identificar insights rellevants (enfocament inductiu).

Per organitzacions on el mode exploratori hi té un pes específic rellevant, és molt important tenir en compte quan s'està definit l'organització, determinar a on s'ha d'ubicar l'equip capaç d'extreure nous insights de les dades, l'equip de ciència de dades i Big Data. En aquest sentit, Gartner¹ identifica els següents escenaris:

Escenaris

¹ 'Organizing for Big Data Through Better Process and Governance'. Gartner. 2015.

On allotjar el Laboratori Big Data a la teva Organització

Dintre de les unitats de negoci

S'aplica en organitzacions altament descentralitzades o quan hi ha una alta demanda dintre d'una àrea de negoci.

Dintre de la unitat de TI

S'aplica quan hi ha una demanda en analítica avançada / ciència de dades altament estesa en diferents unitats de negoci i on hi ha una forta centralització dels serveis de TI.

Com una funció separada

S'aplica quan hi ha una forta necessitat d'innovació i l'estructura organitzativa existent no permet la incorporació de forma fàcil a un gran laboratori de dades. Aquest escenari funciona en el cas d'organitzacions centralitzades àgils.

Parcialment dintre d'una gran unitat de negoci i dintre de la TI o en un equip separat

S'aplica quan l'equip de ciència de dades té un bon finançament per separat, i pot treballar d'una manera multi-funcional.

Com una estructura virtual

Consisteix en establir un equip de ciència d'analítica de dades independent sobre una base purament virtual, deixant als científics de dades dispersos en tota l'organització. Això pot semblar la solució més simple, però li faltarà la direcció, l'enfocament i el poder d'execució.

Taula 1. Escenaris a on ubicar l'equip de ciència de dades i Big Data. Font: Gartner¹.

2.1 ROL I RESPONSABILITATS

Els rols fonamentals a incorporar en una organització *data driven* depèn de la seva mida, model definit, i la naturalesa de la seva activitat. Tot i així, a continuació es descriuen els rols i responsabilitats principals que les organitzacions estan incorporant en els seus equips d'analítica de dades:

- **Chief data Officer (CDO):** és el responsable d'assegurar que l'organització és *data driven*. Lidera la gestió de dades i analítica associada pel negoci i, per tant, és responsable dels diferents equips especialitats en dades.
- **Data Scientists (científic de les dades):** són els membres clau de l'equip de ciència de dades. Permeten extreure coneixement i informació valuosa de les dades. Tenen visió general del procés d'extrem a extrem i poden resoldre problemes de ciències de

dades, la construcció de models analítics i algoritmes. Combinen diverses habilitats relacionades amb les matemàtiques, l'estadística, la programació i visualització, però també han de tenir habilitats comunicatives, per poder explicar els resultats obtinguts a l'organització.

- **Citizen Data scientist:** és la persona dintre de l'organització que típicament no ha està formada específicament per ser *data scientist*, però que pot extreure valor, a través de la seva experiència, explorant les dades, des de les unitats de negoci. Poden executar una sèrie simple de tasques analítiques utilitzant eines de descoberta de dades.
- **Data Engineer (enginyer de dades):** S'encarrega de proporcionar les dades d'una manera accessible i apropiada als usuaris i *data scientist*. És un perfil especialitzat en infraestructura Big Data. Desenvolupa i explota tècniques, processos, eines i mètodes que han de servir per al desenvolupament d'aplicacions Big Data. Té un gran coneixement en gestió de bases de dades, arquitectures de clústers, llenguatges de programació i sistemes de processament de dades.
- **Data Steward (administrador de dades):** és responsable de mantenir la qualitat, disponibilitat i seguretat de les dades. Persegueix millorar l'emmagatzemament i presentació de les dades en tota l'empresa. Té coneixements dels processos de negoci i de com les dades s'utilitzen dins d'aquests processos.
- **Business Data Analyst (analista de dades):** participa en les iniciatives i projectes d'anàlisi de dades. És la persona que recull les necessitats dels usuaris de negoci pel *data scientist* i presenta el resultats obtinguts.
- **Data Artist** són experts en Business Analytics i són els responsables de crear els gràfics, infografies i altres eines visuals per ajudar a les diferents persones de l'organització a comprendre dades complexes.
- **Usuaris** són els responsables d'utilitzar i posar en pràctica els resultats obtinguts dels diferents models analítics realitzats.

Actualment, moltes organitzacions estan formant (o contractant) a personal en els perfils i tasques identificades com a necessàries i estan creant equips multidisciplinaris per tal de disposar de les diferents competències clau: 'La comprensió del domini', 'Ciència de Dades' i 'Habilitats de TI'. És de vital importància per una organització que aquests equips multidisciplinaris puguin comunicar i presentar adequadament els resultats al negoci i, per tant, han d'incloure experts de domini (els usuaris, analistes de dades i/o Data Artists) que poden determinar quines són les preguntes que s'han de resoldre i identificar els *insights* de valor a través de l'anàlisi de les dades a presentar a l'organització.

Per fomentar la cultura basada en les dades és necessari evangelitzar i informar a tots els empleats, a través del CDO, que aquest tipus de treball basat en les dades és necessari i crític per a l'èxit i té un benefici en totes les àrees de l'organització.

Per aconseguir aquest canvi cap a una cultura basada en dades, s'ha de desenvolupar un plantejament holístic integral que tingui una orientació a negoci i a l'analítica. Com en grans projectes que intenten canviar la filosofia d'una empresa, s'han de dur a terme els següents passos a l'hora de fomentar aquesta cultura en l'organització:

- **Determinar i comunicar una visió comuna i objectiu per l'analítica de les dades** per l'Organització liderat per les diferents àrees usuaris conjuntament amb el departament de TI.
- **Definir la metodologia de gestió del canvi** a implementar en tota l'organització. Aquesta ha de tenir en compte les necessitats analítiques específiques. És important implicar i comunicar als responsables de negoci l'evolució del programa de formació definit.
- **Definir un pla de formació relacionat amb la cultura de les dades** a on s'estableix el calendari previst de formacions. Aquest pla de formació ha de tenir en compte les diferents capacitats i rols de les persones de l'Organització.

3. GOVERN DE LES DADES I PROCESSOS

Per definició, una organització *data driven* analitza dades internes, i en molt casos juntament amb dades externes, en els seus processos de negoci clau, per facilitar la presa de decisions. Per aquest motiu, aquestes organitzacions estan adaptant els seus processos i sistemes per tal de proveir accés immediat a informació rellevant per a l'organització. Per a fer-ho, estan definint un model de govern de les dades i polítiques per l'ús de la informació.

3.1 ESTRATÈGIA DE DADES

L'estratègia de dades és l'element clau per qualsevol negoci que vol seguir essent competitiu en el seu mercat, però aquesta és encara més important per aquelles empreses que volen crear disruptió a través de la generació de nous productes i serveis (majoritàriament les que anomenem *'innovation aimed data-driven organizations'*) a través de l'anàlisi de les dades.

L'estratègia de dades ha de definir com s'han d'utilitzar les dades internes i externes per generar nou valor en l'organització. Per aquest motiu, aquesta s'ha d'elaborat tenint en compte tots els agents clau de l'organització i per tant, ha de ser revisada i aprovada. Ha d'incorporar els objectius analítics definits tenint en compte els objectius de negoci així com els elements clau i prioritats del negoci. Incorpora també el full de ruta definit indicant com s'utilitzaran les dades internes i externes en l'organització i els plans d'accions a dur a terme per assegurar la consecució dels objectius analítics definits. Aquesta estratègia ha d'incorporar les diferents preguntes de negoci que es volen respondre a través de l'analítica de dades.

Per altra banda, una estratègia de dades ha de ser flexible. Ha de ser un document viu que es pugui ajustar a mesura que les prioritats del negoci canvien o la tecnologia evoluciona.

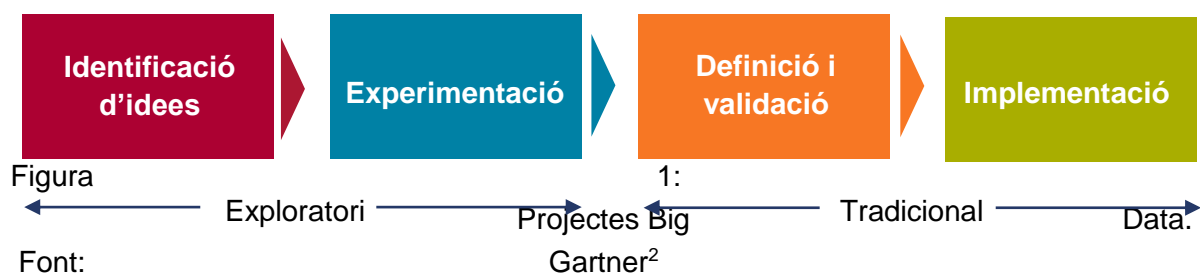
3.2 GOVERN DE DADES

El govern de dades és crític per tota organització. En el govern de les dades es defineix els processos definits per a la ingesta, emmagatzematge i protecció de les dades en l'Organització. Per aquest motiu, disposar d'un model de govern de les dades és un element clau per a tota organització *data driven*, per tal d'assegurar que aquestes es gestionen d'acord amb les normes i directrius definides.

Gartner² identifica els següents tres nivells a l'hora de definir el model de govern de les dades:

1. **Professional:** les diferents persones que treballen amb les dades han de ser conscients de la naturalesa i sensibilitat del treball que realitzen, i han de conèixer els elements ètics relacionats amb privacitat, confidencialitat, transparència i identitat definits en la organització.
2. **Procés:** Les dades han de ser protegides. En aquest sentit, s'ha de definir un procés d'aprovació quan s'obtinguin nous *insights* pel negoci a on es revisin, entre d'altres, aspectes legals, operacionals i de màrketng. L'objectiu d'aquests processos d'aprovació és assegurar que s'estan complint la normatives existents.
3. **Tècnic:** S'han de prendre totes les mesures necessàries per protegir la privacitat de les dades i assegurar-ne la seva seguretat. En aquest sentit, s'ha de poder implementar les eines necessàries que assegurin que els anàlisi realitzats tenen en compte els aspectes de privacitat i seguretat definits en l'organització.

3.3 DADES I PROJECTES BIG DATA



Tal i com s'indicava en l'apartat '2.1. Organització', les organitzacions *data driven* parteixen d'una organització de TI Bimodal. La relació entre les diferents fases dels projectes Big Data i els modes de funcionament és la següent:

- Les fases 1 i 2 representen el mode d'exploració Big Data (**Mode 1 'exploratori'**).
- Les fases 3 i 4 representen l'explotació Big Data i, per tant, es basa en el mode més tradicional (**Mode 2 'tradicional'**).

² 'Organizing for Big Data Through Better Process and Governance'. Gartner. 2015.

A continuació es detallen les diferents fases a dur a terme en la identificació i execució de projectes Big Data:

1. **Identificació d'idees:** L'objectiu d'aquesta fase és identificar i agrupar idees amb la finalitat de dur a terme experiments (per exemple, a través de workshops). En aquest sentit, el CDO conjuntament amb els usuaris de negoci han de determinar prioritats (tenint en compte l'impacte i l'aplicabilitat) per tal de poder abordar-les.
2. **Experimentació:** Les idees que es duren a terme necessiten ser traduïdes a hipòtesis de negoci que puguin ser provades. Un cop validat l'experiment a dur a terme, s'ha de donar accés a les dades necessàries per tal de provar els models analítics i predictius i validar la hipòtesi. La durada d'aquests experiments ha de ser curta per tal d'assegurar un ràpid retorn per les organitzacions.
3. **Definició i validació:** Aquesta fase consisteix en la definició del Business Case del projecte o prova de concepte a dur a terme indicant clarament el pressupost estimat i el ROI esperat. En aquest punt, s'ha de revisar si la solució a adoptar compleix la normativa vigent (legal) i interna, retorna valor al client o a la companyia.
4. **Implementació:** Correspon al desplegament dels experiments validats en fases anterior. En aquest punt, la metodologia de desenvolupament ha de tenir en compte totes les fases del cicle de vida de les dades i, per tant, ha de tenir en compte la seva ingesta, la gestió i emmagatzematge, el seu anàlisi i procés i, finalment, la seva visualització.

3.4 QUADRES DE COMANDAMENT

Un element clau en les organitzacions *data driven* són els **quadres de comandament** per tal de revisar la informació al voltant de les mètriques més rellevants per a la organització (i cada departament) que requereixen d'anàlisi diari. De cara a construir aquests quadres de comandament s'han de tenir en compte els següents elements:

- **Latència.** En aquest sentit, per tal d'assegurar la consulta diària de les dades, s'ha de preveure que poden haver canvis en les visualitzacions, i tenir clares les accions que es faran quan canviïn.
- **Valor.** S'ha de gestionar els diferents informes i quadres de comandament disponibles per tal de validar periòdicament que aquest retornen el valor esperat. S'ha de considerar, per tant, que aquest informes i quadres de comandament són una eina viva que ha d'ajudar a gestionar l'organització.
- Comprensió **visual** i, per tant, revisar que la informació es visualitzi correctament.
- Possible **fatiga** dels usuaris per tal d'assegurar que no es programen massa alertes que els puguin dessensibilitzant.

4. TECNOLOGIA

Aprendre a entendre el problema, formular un experiment, fer bones preguntes, emmagatzemar dades, calcular estadístiques, implementar un algoritme i verificar l'exactitud del resultat són elements clau dels equips de ciència de dades elements clau. Aquests són, en molts casos, més importants que les eines i tecnologies a utilitzar. Tot i així, hi ha dos aspectes a tenir en compte en organitzacions *data driven* relacionats amb la tecnologia: la centralització i democratització de les dades.

4.1 CENTRALITZACIÓ

El CDO o l'equip d'analítica de dades són els responsables d'ingerir les dades de diferents sistemes i bases de dades i emmagatzemar-les en un repositori central – en algun format específic o estandarditzat – per què aquestes puguin ser analitzades i utilitzades. Es persegueix centralitzar la informació de l'organització en un únic repositori que pot ser enriquit per fonts de dades externes.

Per tant, aquest repositori central incorpora no només informació i dades de la pròpia organització sinó que també pot emmagatzemar dades externes d'APIs obertes, venedors, o altres fonts. Aquestes es poden correlar amb les dades internes per tal d'enriquir la informació amb la finalitat d'ajudar a respondre a les preguntes que són importants per al negoci.

4.2 DEMOCRATITZACIÓ DE LES DADES

La democratització de les dades consisteix en facilitar i donar accés als usuaris de l'organització al repositori central de dades. La democratització presenta diferents reptes que s'han d'abordar a l'hora de definir aquest repositori de dades: assegurar que les dades disponibles són correctes i ajudar als usuaris a identificar les dades de valor. En molts desenvolupaments, preparar les dades pels usuaris (i per tant, assegurar que siguin correctes) correspon al 80% de la feina d'un *data scientist*. Per aquest motiu i per tal d'ajudar als empleats a fer el millor ús de les dades, moltes organitzacions estan incorporant un nou perfil clau: el *data Stewart*. Aquest té per missió assegurar la coherència i la qualitat de les dades mitjançant la incorporació de les eines i processos necessaris.

Tot i així, la democratització de les dades pot requerir en molts casos la revisió i redefinició de les pràctiques i eines de gestió i explotació de dades de l'organització. Es persegueix que les dades estiguin emmagatzemades de tal forma que els usuaris les puguin utilitzar de forma intuïtiva enlloc d'implementar sistemes complexos per a la seva explotació.

5. CONCLUSIONS

Tenir èxit amb les dades requereix un canvi cultural real que afecta a tota la organització. Requereix facilitar l'accés a les dades, no només als *data scientists* (tot i que són claus en la cultura de dades), sinó també a tots els membres de l'organització i facilitar-ne el seu aprenentatge. Es requereix, per tant, que les dades siguin fiables i de qualitat, que estiguin disponibles quan es necessiten, que la manera de veure els resultats de l'anàlisi aplicada sigui intuïtiva i que les diferents persones amb accés, aprenguin a revisar i a interpretar-les per obtenir-ne nous *insights* que permetin anticipar els problemes, ampliar el coneixement del negoci, millorar la productivitat i mantenir-se en l'avantguarda

L'establiment de la cultura de dades no és fàcil. Requereix de persistència i paciència. És més probable que es tingui èxit si es comença desenvolupant-la a través de l'execució de petits projectes, que permetin analitzar els resultats obtinguts i validar els beneficis envers la organització, abans de començar implementant un projecte ambiciós. Però, independentment de l'aproximació que s'esculli, fomentar la cultura de dades en l'organització és la clau de l'èxit en el segle XXI.

Per aquest motiu i tenint en compte les oportunitats que es presenten amb l'ús del Big Data, és de vital importància que les empreses identifiquin i abordin els obstacles que es presenten que els puguin impedir aprofitar aquestes oportunitats. En aquest sentit, segons BSI & Circle research³ s'identifiquen reptes tècnics, culturals i de percepció per a la seva adopció:

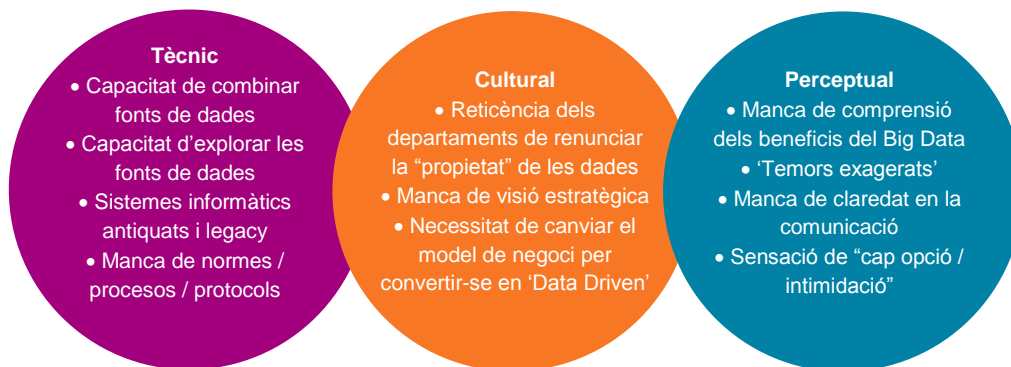


Figura 2: Reptes en l'adopció del Big Data. Font: BSI & Circle research³.

Per tal d'esdevenir una organització *data driven* es plantegen els següents passos i directius:

³ 'Big Data and standards market research'. British Standards Institution and Circle Research. 2016.

1. **Identificar oportunitats d'alt valor.** Identificar projectes que aportin valor pel negoci i impacte positiu en l'organització. Aquests projectes han de ser la base per construir i desplegar els fonaments per dur a terme noves oportunitats Big Data i en analítica de dades. En aquest sentit, és necessari començar en projectes petits i, en paral·lel, definir projectes més ambiciosos (*'start small and think big'*).
2. **Demosttar el valor als líders de negoci mitjançant programes pilot.** Definir els objectius que es persegueixen i executar proves de concepte que validin l'estratègia definida. En aquest sentit, els projectes i arquitectura Big Data s'han de construir a partir de les resultats obtinguts.
3. **Contractar i/o formar al personal clau per establir l'equip de ciències de dades.** Identificar les habilitats necessàries per construir un equip de ciències de dades, i buscar aquests perfils dins de l'organització. Capacitar les persones o contractar-ne de noves quan sigui necessari tenint en compte els objectius de negoci. Aquest pas, a més, es pot fer amb el suport d'una empresa col·laboradora que porti perfils experts i referències de valor per accelerar el procés.
4. **Ampliar mitjançant l'execució de nous casos d'us.** Si s'escau, reutilitzar les tecnologies utilitzades o incorporar-ne de noves. Aplicar les lliçons apreses en l'execució de noves oportunitats.
5. **Definir els processos clau: estratègia, model de govern i processos.** Començar a treballar amb les dades requereix que tots els processos de les dades s'hagin definit i implementat a l'organització. Per tant, s'ha de definir l'estratègia de dades que ha de descriure la forma en què les dades internes i externes s'utilitzaran per generar nou valor, el model de govern de dades per recopilar, emmagatzemar i protegir les dades i els processos necessaris.
6. **Democratitzar les dades.** Compartir les dades a tots els empleats i permetre'ls experimentar per tal de descobrir nous *insights* que pugui aportar valor a l'organització.

Cal notar que també es pot esdevenir una organització data driven, arribant al pas 6 sense un equip de data scientists intern, o ni tant sols tenint el suport d'un d'extern. En aquest cas però, la capacitat d'innovar amb la explotació de les dades serà significativament menor, i per tant ja no parlariem de *'innovation aimed data driven organizations'*.

6. REFERÈNCIES

- 2016. 'Big Data and standards market research'. British Standards Institution and Circle Research.
- DJ Patil and Hilary Mason. 2015. 'Data Driven. Creating a Data Culture'. O'Reilly.
- Julie Steele. 2015. 'Understanding the Chief Data Officer. How Leading Businesses Are Transforming Themselves with Data'. O'Reilly.
- Frank Buytendijk and Thomas W. Oestereich. 2015. 'Organizing for Big Data Through Better Process and Governance'. Gartner.
- Alexander Linden, Lisa Kart, Lakshmi Randall, Mark A. Beyer and Alan D. Duncan. 2015. 'Staffing Data Science Teams'. Gartner.
- Alan D. Duncan and Frank Buytendijk. 2015. 'How to a Data-Driven Culture in the Digital Workplace'. Gartner.
- Daniel Yuen, Helen Poitevin and Alan D. Duncan. 2015. 'How to Design Effective Analytics Training to Improve Your Business Analytics Maturity'. Gartner.
- Josep Lluís Cano Giner. 2014. 'Big data: nuevos perfiles profesionales'. Revista Equipos & Talento
- Hasan Bakhshi, Albert Bravo-Biosca and Juan Mateos-Garcia. 2014. 'Inside the Datavores'. Nesta
- Harvard Business Review. 2014. 'Embedding Analytics for Growth: Creating a Data-Driven Culture'. <https://hbr.org/2014/12/embedding-analytics-for-growth-creating-a-data-driven-culture>
- Douglas Laney, Frank Buytendijk and Alexander Linden. 2013. 'Predicts 2014: Innovating With Information Will Demand New Data, Organizations and Ideas'. Gartner.
- Jim Sterne. 2013. 'From Data Scientist to Data Artist'. Anametrix
- Hasan Bakhshi and Juan Mateos-Garcia. 2012. 'Rise of the Datavores. How UK Business Analyse and Use Online Data'. Nesta
- Economist Intelligence Unit. 2010. 'Levelling the playing field. How companies use data for competitive advantage'. The Economist.