

# TEMA 0

MÓDULO:  
HERRAMIENTAS DE BIG DATA

## INTRODUCCIÓN AL BLOQUE I

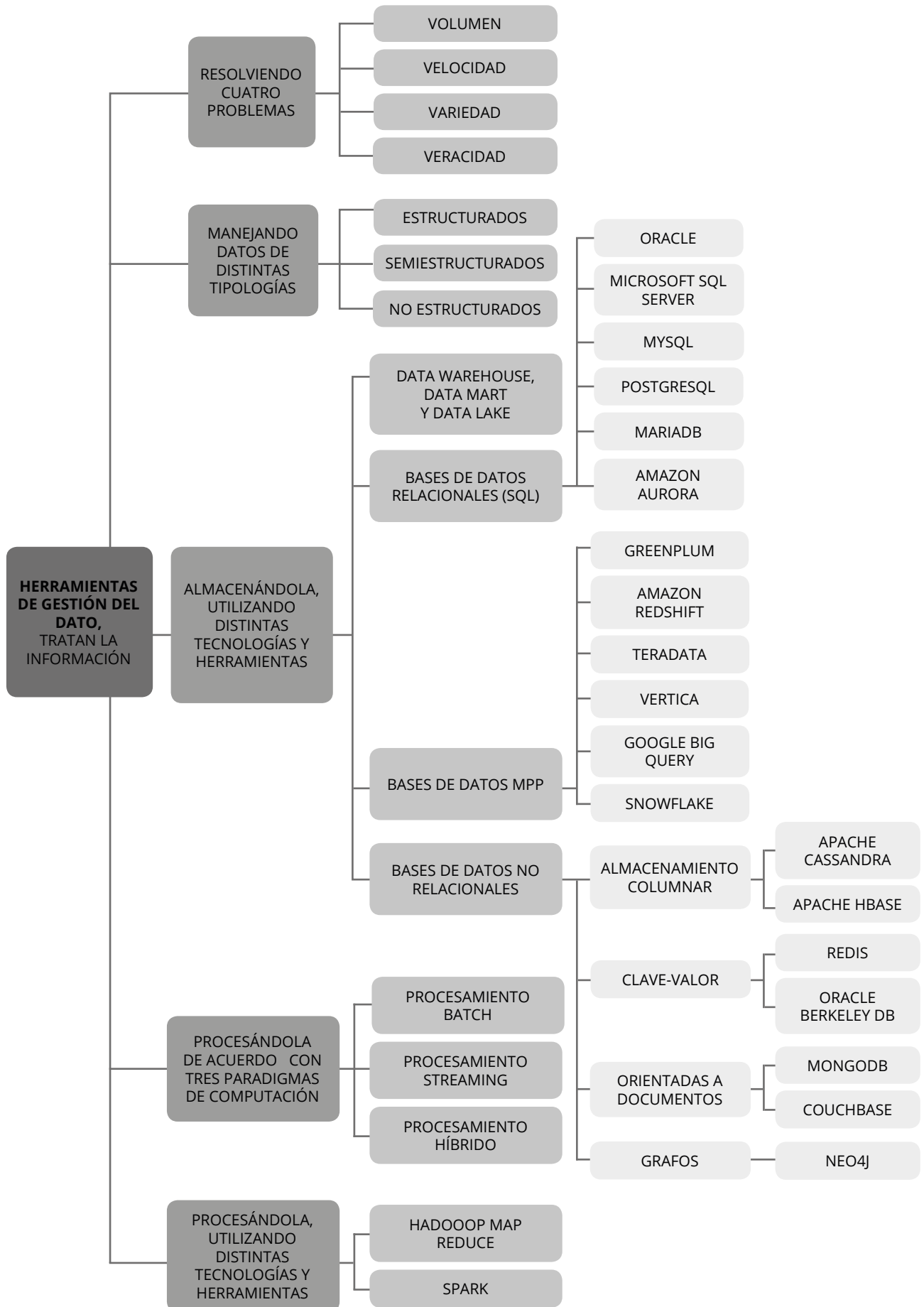
# STAR WARS EPISODE I THE PHANTOM MENACE

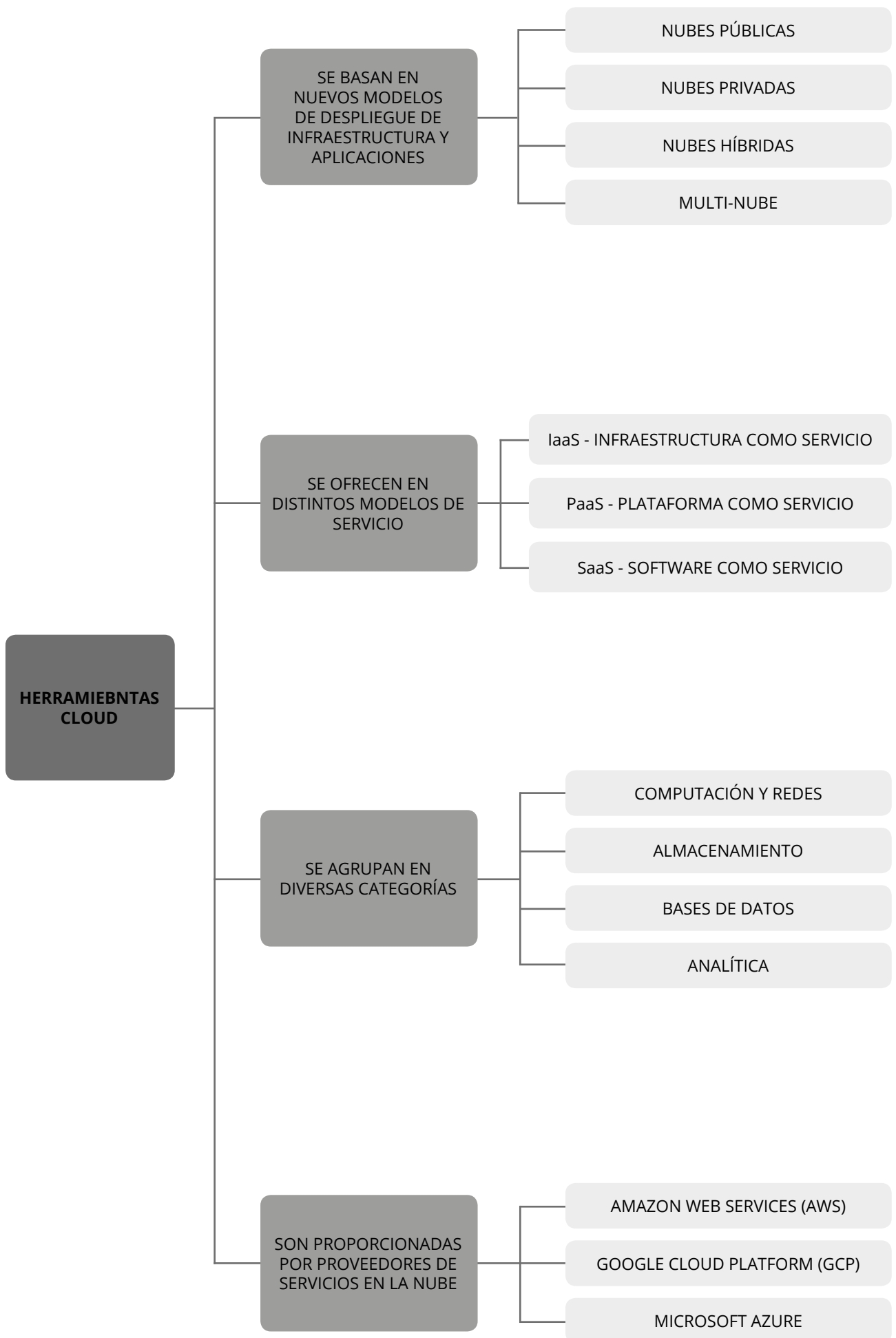


*Institut de Formació Contínua-IL3*  
UNIVERSITAT DE BARCELONA

© de esta edición: Fundació IL3-UB, 2020

# 1. MIND MAP









# OBJETIVOS

- Conocer las distintas herramientas de gestión del dato.
- Saber cuándo y dónde aplicar cada una de las herramientas de gestión del dato.
- Conocer las distintas herramientas Cloud y entender las ventajas del nuevo modelo de computación en la nube.
- Conocer las distintas herramientas de monitorización y control.
- Ser capaz de aplicar los conocimientos adquiridos de forma práctica.



# EVALUACIÓN

**Evaluación continua** del trabajo realizado en clase mediante la resolución de 3 partes:

1. Prueba **teórica**: al superar los **test** con éxito se alcanzará la posición de **Initiate Level**.
2. Prueba **individual**: al superar el **trabajo individual** se logrará la posición **Padawan Level**.
3. Prueba **grupal**: superar el **trabajo colectivo** supondrá conseguir la posición **Knight Level**.

## CRITERIOS MÍNIMOS

El *alumno/a Padawan* para alcanzar el nivel debe superar con éxito los siguientes hitos:

- **Initiate Level:** prueba de asentamiento de conceptos teóricos, para superar esta parte deberás obtener una calificación superior a 5.

**Nota:** Las preguntas que no se contesten de forma correcta restará puntos (indicado en cada actividad).

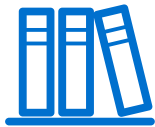
- **Padawan Level:** Realizar, al menos una práctica individual, defendiéndola y justificándola adecuadamente.

- **Knight Level:** Realizar al menos una práctica colectiva (participación activa en reuniones y discusiones de grupo, así como en la elaboración de informes, etc. ), defendiéndola y justificándola adecuadamente.

Los porcentajes de cada hito estarán reflejados en el plan docente y en cada actividad.

Para aprobar el módulo, la media de todos los hitos debe ser superior al 5.

Recuerda que es evaluación continua por lo que cuantas más prácticas realices más posibilidades tendrás de alcanzar el máximo nivel Padawan.



# BIBLIOGRAFÍA Y RECURSOS EN INTERNET

## MATERIAL PARA PROFUNDIZAR EN HERRAMIENTAS DE GESTIÓN DEL DATO

- Shannon Bradshaw, Eoin Brazil, Kristina Chodorow. MongoDB: The Definitive Guide, 3rd Edition (2019). O'Reilly Media, Inc.

Disponible en: <https://www.oreilly.com/library/view/mongodb-the-definitive/9781491954454/>

Libro para ampliar los conocimientos de MongoDB.

- Tom White. Hadoop: The Definitive Guide, 4th Edition (2015). O'Reilly Media, Inc.

Disponible en: <https://www.oreilly.com/library/view/hadoop-the-definitive/9781491901687/>

Libro para ampliar los conocimientos sobre Hadoop.

- Bill Chambers, Matei Zaharia (2018) Spark: The Definitive Guide. O'Reilly Media, Inc.

Disponible en: <https://www.oreilly.com/library/view/spark-the-definitive/9781491912201/>

Libro recomendable para tener unos conocimientos sólidos de Spark.

- James Warren, Nathan Marz (2015) Big Data: Principles and best practices of scalable real-time data systems. O'Reilly Media, Inc.

Disponible en: <https://www.oreilly.com/library/view/big-data/9781617290343/>

Libro para conocer los principios y buenas prácticas en los sistemas en tiempo real. Nathan Marz es el creador de Apache Storm y el concepto de Arquitectura Lambda

- Martin Kleppmann (2017) Designing Data-Intensive Applications. O'Reilly Media, Inc.

Disponible en:

<https://www.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/>

Martin Kleppmann trabajó en LinkedIn y nos da las claves de diseño de aplicaciones con uso intensivo de datos.

- GOESSLING, S.; JACKSON, K.L. (2018) Architecting Cloud Computing Solutions. Birmingham: Packt Publishing.

Disponible en:

<https://www.oreilly.com/library/view/architecting-cloud-computing/9781788472425>

Libro para conocer cómo diseñar e implementar soluciones en entornos Cloud de manera efectiva y eficiente.

- MAXIM, B.; HEISEL, M.; ALI, N. y cols. (2017) Software Architecture for Big Data and the cloud. Burlington: Morgan Kaufmann.

Disponible en: <https://www.oreilly.com/library/view/software-architecture-for/9780128093382/>

Libro para conocer cómo crear arquitecturas en entornos Cloud que den soporte a nuestras aplicaciones Big Data.

## MATERIAL PARA PROFUNDIZAR EN MONITORIZACIÓN Y CONTROL

- Scott Chacon and Ben Straub. Pro Git 2nd Edition (2014). Apress.

Disponible en: <https://git-scm.com/book/en/v2>

Libro oficial para conocer Git

## MATERIAL PARA PROFUNDIZAR EN CLOUD

- Kevin L. Jackson, Scott Goessling. Architecting Cloud Computing Solutions [Internet]: Packt Publishing; 2018 [revisado ;citado 2020 Sep 24]

- Disponible en:  
<https://www.oreilly.com/library/view/architecting-cloud-computing/9781788472425>

- Ivan Mistrik, Rami Bahsoon, Nour Ali, Maritta Heisel, Bruce Maxim. Software Architecture for Big Data and the Cloud [Internet]: Morgan Kaufmann; 2017 [revisado ;citado 2020 Sep 24]

- Disponible en:  
<https://www.oreilly.com/library/view/software-architecture-for/9780128093382/>



