

Arquitecturas y tecnologías para el *big data*. Actividades



Marta Zorrilla - Diego García-Saiz
Enero 2017

Este material se ofrece con licencia: [Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



Actividad #1

- Objetivo:
 - Diseñar e implementar un *streaming pipeline* con una fuente de datos real o simulada.
- Tareas:
 - Determinar el problema a resolver (contar palabras de twitts, monitorizar eventos en un sitio web, etc.).
 - Seleccionar la arquitectura y tecnologías.
 - Instalar en local o bien hacer uso de servicios *cloud* para su implementación.
 - Programarlo y testearlo.

Actividad #2

- Objetivo:
 - Establecer las diferencias entre la gestión de datos transaccional y datos en *streaming*.
- Tareas:
 - Caracterizar, de acuerdo a los aspectos señalados en la siguiente tabla, la gestión de datos transaccional vs la gestión de datos en *streaming*.

Actividad #2

	Transaccional	Streaming
Data		
Access		
Queries		
Order		
Support (persistence)		
Update rate		
Temporal requirements		
Accuracy		
Heterogeneity		
Algorithms		