



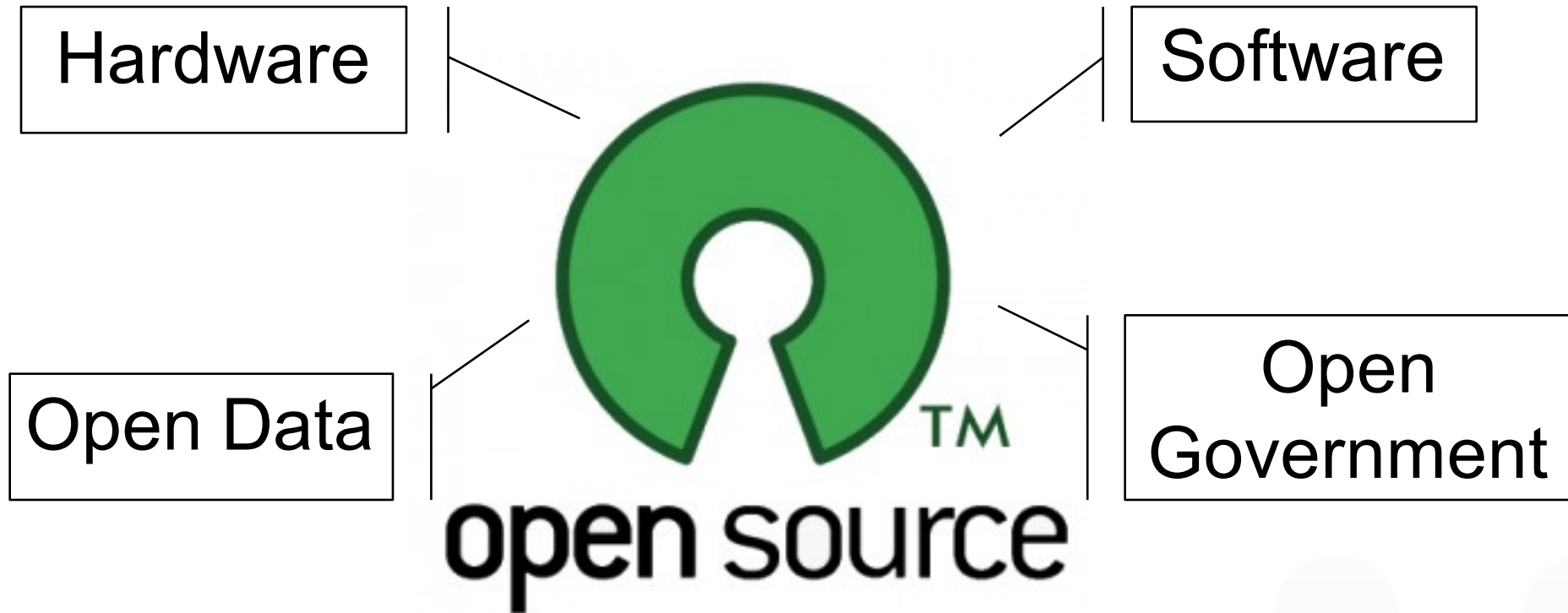
Open Source Business Intelligence

Josep Curto Díaz
Profesor Master Business Intelligence y Big Data, UOC

Índice

- La nueva revolución es open source
- Ventajas y desventajas del OSBI
- Impacto del Open Source
- Open Source y Business Intelligence
- Open Source y Business Analytics
- Open Source y Big Data

La nueva revolución es open source



Ventajas y desventajas del OS

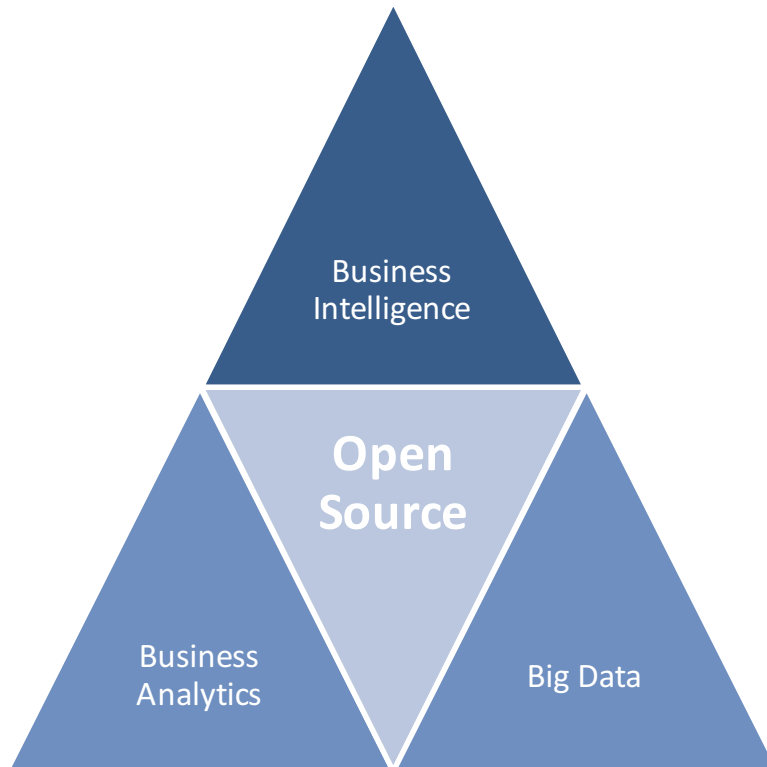
Ventajas

- Reducción del TCO
- Acceso al código fuente
- Soporte de la comunidad
- Independencia de un único proveedor
- Flexibilidad y agilidad

Desventajas

- Difícil soporte a grandes organizaciones
- Falta de funcionalidades específicas
- Versión *community* limitada frente a versión *enterprise*

Impacto del Open Source



- Afecto en todos los ámbitos
- Más y más empresas de todo tamaño están usando open source o algunos de sus módulos por las prestaciones que ofrecen para el análisis de datos.
- En el contexto del BI se ha producido la compra de los principales actores.
- En el contexto del Business Analytics, R y Python se han convertido en los lenguajes de facto para el científico de datos.
- En el contexto de Big Data, las principales tecnologías son open source.
- Además, en el último año más y más empresas de software están adoptando el open source para la creación de ecosistemas de desarrolladores.

Open y Business Intelligence

Pentaho

- Incluye: Reporting, OLAP, ETL, Dashboards y Minería de datos
- Foco en el tratamiento de datos
- Comprada por Hitachi

Actuate

- Incluye: Reporting (BIRT)
- Foco en reporting
- Comprada por Open Text

Jaspersoft

- Incluye: Reporting, OLAP, ETL y Dashboards
- Foco en visualización y BI embebido
- Comprada por Tibco

SpagoBI

- Incluye: Reporting, OLAP, ETL, Dashboards y Minería de datos
- Foco en 100% OS

Open Source y Business Analytics

R

- Entorno y programa para el análisis estadístico
- Basado en librerías
- Soportado por la mayoría de fabricantes

H2O

- Plataforma para analítica predictiva
- Foco en machine learning y big data

Weka

- Colección de algoritmos de minería de datos
- Software basado en java

TensorFlow

- Librería de machine learning
- Creada por Google y liberada en 2015

Open y Big Data

Apache Hadoop

- Ecosistema para el tratamiento y procesamiento de grandes datos
- Implementa el paradigma Map/Reduce

MongoDB

- Base de datos NoSQL
- Orientada a documentos

Apache Spark

- Motor para el procesamiento de datos a gran escala con foco en la rapidez

Neo4j

- Base de datos NoSQL
- Orientada a grafos



Muchas gracias por vuestra atención