Chapter 1
What Is Open Source Software (OSS) and What Is Big Data?

Richard S. Segall
Arkansas State University, USA

ABSTRACT
This chapter discusses what Open Source Software is and its relationship to Big Data and how it differs from other types of software and its software development cycle. Open source software (OSS) is a type of computer software in which source code is released under a license in which the copyright holder grants users the rights to study, change, and distribute the software to anyone and for any purpose. Big Data are data sets that are so voluminous and complex that traditional data processing application software are inadequate to deal with them. Big data can be discrete or a continuous stream data and is accessible using many types of computing devices ranging from supercomputers and personal workstations to mobile devices and tablets. It is discussed how fog computing can be performed with cloud computing for visualization of Big Data. This chapter also presents a summary of additional web-based Big Data visualization software.

INTRODUCTION: HOW OPEN SOURCE SOFTWARE, FREE SOFTWARE, AND FREEWARE DIFFER

Open Source Software (OSS)
Open-Source Software (OSS) is a type of computer software in which source code is released under a license in which the copyright holder grants users the rights to study, change, and distribute the software to anyone and for any purpose. (Wikipedia (2019a))
DOI: 10.4018/978-1-7998-2768-9.ch001
For software to be considered “Open Source”, it must meet ten conditions as defined by the Open Source Initiative (OSI). Of these ten conditions, it’s the first three that are really at the core of Open Source and differentiates it from other software. These three conditions are according to the Open Source Initiative (2007):

1. **Free Redistribution**: The software can be freely given away or sold.
2. **Source Code**: The source code must either be included or freely obtainable.
3. **Derived Works**: Redistribution of modifications must be allowed.

The other conditions are: (Open Source Initiative (2007))

4. **Integrity of The Author’s Source Code**: Licenses may require that modifications are redistributed only as patches.
5. **No Discrimination against Persons or Groups**: no one can be locked out.
6. **No Discrimination against Fields of Endeavor**: commercial users cannot be excluded.
7. **Distribution of License**: The rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.
8. **License Must Not Be Specific to a Product**: the program cannot be licensed only as part of a larger distribution.
9. **License Must Not Restrict Other Software**: the license cannot insist that any other software it is distributed with must also be open source.
10. **License Must Be Technology Neutral**: no click-wrap licenses or other medium-specific ways of accepting the license must be required.

Macaulay (2017) discussed benefits of open source software that are summarized in Figure 1 below.

*Figure 1. Benefits of Open Source Software (OSS) (Derived from Macaulay (2017))*
Related Content

What Is Open Source Software (OSS) and What Is Big Data?
[www.igi-global.com/chapter/what-is-open-source-software-oss-and-what-is-big-data/248872?camid=4v1a](www.igi-global.com/chapter/what-is-open-source-software-oss-and-what-is-big-data/248872?camid=4v1a)

Introduction to the Popular Open Source Statistical Software (OSSS)
[www.igi-global.com/chapter/introduction-to-the-popular-open-source-statistical-software-osss/248874?camid=4v1a](www.igi-global.com/chapter/introduction-to-the-popular-open-source-statistical-software-osss/248874?camid=4v1a)

Generalized Linear Model for Automobile Fatality Rate Prediction in R
[www.igi-global.com/chapter/generalized-linear-model-for-automobile-fatality-rate-prediction-in-r/248876?camid=4v1a](www.igi-global.com/chapter/generalized-linear-model-for-automobile-fatality-rate-prediction-in-r/248876?camid=4v1a)
A Comparison of Machine Learning Algorithms of Big Data for Time Series Forecasting Using Python
www.igi-global.com/chapter/a-comparison-of-machine-learning-algorithms-of-big-data-for-time-series-forecasting-using-python/248878?camid=4v1a