Data labeling for data security in data lifecycle: A state of the art and issues

Kenza Chaoui*, Nadia Kabachi**
Nouria Harbi**, Hassan Badir *

*IDS Team ENSAT, Abdelmalek Essaadi University Tangier
Kchaoui1994@gmail.com
hbadir@uae.ac.ma
**ERIC Laboratory,Lumière Lyon 1 University Lyon, France
nadia.kabachi@univ-lyon1.fr
nouria.harbi@univ-lyon2.fr

Résumé. One of the most serious issues with cloud computing is data security, As businesses start on digital transformation, there is a clear requirement for privacy and data protection. Organizations today have more data, applications, and websites than they have ever had before. Data security has risen to the top of the priority list for cloud computing security. Despite the fact that a variety of solutions have been proposed, the majority of them only address one stage of the data life cycle, such as storage, which is insufficient to handle the cloud data security challenge because threats appear at all stages of the data life cycle. During the data life cycle process, any stage's security breaches could affect data security. Therefore, data security must be considered throughout the data lifecycle. The main contribution of this article is a new perspective on data security solutions based on the data lifecycle, which is crucial and can be used as a guide to create a complete security solution.A literature review on the entire data life cycle is carried out and a research gap of unresolved issues that may be research questions for our future work is presented.Also a proposed solution on data labeling used for data tracking to secure data in all stages in data life cycle is presented.

1 Introduction

Cloud computing is a virtualized system that allows users to access compute, storage, and software resources as well as servers from a single platform. Data management services are currently provided in the user's local environment, however CSP Cloud Service Providers provide them remotely. Users may not know where, when, how, why, or by whom their data is seen or modified in a cloud environment because services are supplied in abstract form. Cloud computing, on the other hand, has several security risks. CSPs are also more vulnerable to adversaries and hackers who can take advantage of these benefits. The cloud is vulnerable from a security and data privacy perspective, as sensitive user data is stored in a third-party CSP. In Michelin et al. (2018) all these weaknesses and mistrust build a common element which is the issue of trust and safety between providers and consumers, because Cloud Computing requires