

Abstract der fertigen Arbeit

Das Fachgebiet der Daten-Provenienz beschäftigt sich mit der Herkunft von Daten. Es erlaubt uns, Vertrauen in Daten aufzubauen. Dafür müssen die Provenienz-Daten selber zuverlässig sein und gewisse Sicherheitsstandards erfüllen. Das zu erreichen erweist sich als schwierig, da Daten-Provenienz oft sehr Domänen-abhängig ist, wodurch eine starke Fragmentierung des Fachgebietes entstanden ist.

Diese Fragmentierung macht es schwierig, einheitliche Sicherheitsstandards zu implementieren. Deswegen haben sich auch für die Sicherheitsanforderungen Domänen-spezifische Lösungen entwickelt. Außerdem erschwert die Fragmentierung die Zusammenarbeit zwischen den verschiedenen Domänen. Dies wiederum erschwert es, die Herkunft von Daten nachvollziehbar zu machen was das eigentliche Ziel von Daten-Provenienz ist.

In dieser Arbeit stellen wir eine Blockchain-basierte Lösung vor um Domänen-unabhängige Suchgebiete für Daten-Provenienz zu erstellen. Dies erlaubt es uns, auch Sicherheitseigenschaften der Blockchain auf eine einheitliche Art und Weise auf Daten-Provenienz zu übertragen.

Mithilfe unserer Lösung können wir die Zusammenarbeit zwischen Domänen ermöglichen, ohne diesen einheitliche Provenienz-Eigenschaften aufzwingen zu müssen. Dies erlaubt es, die Herkunft von Daten über Domänengrenzen hinaus nachvollziehbar zu machen und ein vollständigeres Bild der Datenherkunft zu erstellen. Gleichzeitig kreieren wir eine Plattform auf der Domänen-unabhängige Sicherheitsanforderungen einheitlich umgesetzt werden können.

Abstract of the Thesis

Data provenance allows to reproduce what has happened to data during its lifecycle, i.e., it allows to build trust in data and decisions. To achieve this, provenance data has to be reliable itself and to fulfill certain security requirements. This proves to be a difficult problem since data provenance often has many domain-specific properties. Addressing these domain specificities has led to a fragmentation of the field of data provenance.

This fragmentation is a problem since it makes it hard to implement common solutions for security requirements. Instead, the field of secure provenance often focuses on providing domain-specific solutions. Furthermore, this fragmentation makes collaboration between domains difficult, hindering the overall goal of making the history of data products reproducible.

In this thesis, we present a blockchain-based approach for creating a domain-independent search space for data provenance. At the same time, we are able to utilize this search space to map strong blockchain-based security properties, in a domain-independent way, to the field of data provenance.

With our solution, we can enable collaboration between domains without enforcing common provenance properties to the domains. This would allow us to track data products across domain borders and create more complete provenance views while providing a platform for common security properties.