|  |  |
| --- | --- |
| TALLINN UNIVERSITY OF TECHNOLOGY | |
| School of Information Technologies | |
|  | |
|  | |
| Ioane Sharvadze 166798IVSM | |
| **Framework for Peer-to-Peer Data Sharing over Web Browsers** | |
| Master’s Thesis | |
| Supervisor: | Vishwajeet Pattanaik |
|  | PhD |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author:

Abstract

In the era of Web 2.0 we are using numerous applications that store and manage user data. Changing one application with another means losing data. Recent data leaking scandals [1] raise a question of data privacy and ownership.

While there are numerous approaches to solve this problem, this thesis provides a new approach that enables developers to create applications that empower users with data ownership capabilities with ease.

The thesis presents a platform enable application to share user data via peer to peer connections and save it locally on user’s own device.

While the framework can be extended to work with different kinds of devices, thesis will only focus on chrome browser extension and web application solution, since the idea of a generic platform came from specific application.

**Keywords**: Liquid democracy, Peer-to-peer network, Data Ownership, Security, WebRTC

Thesis is written in English and is pages long, including chapters, figures and tables.

List of abbreviations and terms

|  |  |
| --- | --- |
| WebRTC | *Web Real Time Communication* |
| TUT | Tallinn University of Technology |
| P2P | Peer to Peer |
|  |  |

Table of contents

[Author’s declaration of originality 2](#_Toc512694538)

[Abstract 3](#_Toc512694539)

[List of abbreviations and terms 4](#_Toc512694540)

[Table of contents 5](#_Toc512694541)

[Table of contents 5](#_Toc512694542)

[List of figures 7](#_Toc512694543)

[List of tables 8](#_Toc512694544)

[1 Introduction 9](#_Toc512694545)

[1.1 Relevance of the Topic 9](#_Toc512694546)

[1.2 Background and literature review 10](#_Toc512694547)

[1.2.1 Linked Data, Solid 10](#_Toc512694548)

[2 11](#_Toc512694549)

[2.1 11](#_Toc512694550)

[3 Summary 12](#_Toc512694551)

[4 References 13](#_Toc512694552)

[Appendix 1 – 15](#_Toc512694553)

List of figures

No table of figures entries found.

List of tables

**No table of figures entries found.**

# Introduction

This thesis is about enabling Peer to Peer data sharing between users and provides a solution for a Chrome Extension and application developers. The goal is to remove centralized server as a data store and empower users with real data ownership. Ultimately, we would like to remove server at all and make users able to communicate with each other without any middleware, but later I will explain that this is impossible.

The framework goals came from a real-life application, called WebWeaver, but it solves a problem for other applications too, because built platform does not need to have a knowledge of data itself, and thus it can be used for any other applications, that need to share data between users.

WebWeaver is a tool that enables users to weave data into specific web element without knowledge and support of the website. It is currently being developed as Chrome Extension, that adds this feature to all websites. This thesis will create client library and web services, that enables application to store data on user’s device and directly share it with specific users. As later described for establishing communication between users, server is necessary, but actual communication and data transfer happens on Peer to Peer network.

As stated above the thesis provides a generic solution, that is why later I will demonstrate its usage for a different application, that will help to understand and generalize problem.

## Relevance of the Topic

This section will explain the importance of this research.

There are two major problems associated with the centralized data management. Data Ownership and privacy. Whenever using any well-known applications such as Facebook, Google Plus, Twitter etc. we see that they own and manage our data and it is impossible to continue using other application without losing account data [1]. Imagine after years of using Facebook, user wanted to use another application, that would have similar features, but different interface, or provide extended/better solution. User would not be able to transfer its data to the new website.

Data privacy is another concern for users. While people send messages to each other, their data is stored by application owner. Especially after recent user information leak [2], users start to worry about their private or limited (shared only with friends) data.

However, there are numerous proposals to tackle this problem, yet none of them are widely adopted [1].

Area of decentralized data management remains in active research. Therefore, this topic should be considered as important.

## Background and literature review

This chapter will discuss some of the solutions currently available and it will explain the concerns that made it necessary to create a new solution.

### Linked Data, Solid

Solid is very interesting platform that was built specifically to target data ownership. It uses Web ID to identify users. User should choose the service that will host its data. Service can be third party or self-hosted, but it should implement Solid interface to support all the features. It uses RDF-based resources to link data [1].

The reason why we don’t want to go with Solid is that it still stores data on non-user device. It can be hard for user to set up service or find any free hosting service. So, it could be better if without any configuration, user could start using application, but also have its own data on a personal computer.

# 

## 

# Summary

# References

|  |  |
| --- | --- |
| [1] | “Cambridge Analytica Scandal,” [Online]. Available: https://www.theverge.com/2018/4/10/17165130/facebook-cambridge-analytica-scandal. |
| [2] | E. Mansour, A. V. Sambra, S. Hawke, T. Berners-Lee, M. Zereba, S. Capadisli, A. Ghanem and A. Aboulnaga, “A Demonstration of the Solid Platform for Social Web Applications”. |
| [3] | R. Thomas, “Representational State Transfer (REST),” 2000. [Online]. Available: http://www.ics.uci.edu/~fielding/pubs/dissertation/rest\_arch\_style.htm. |
| [4] | J. Pouwelse, P. Garbacki, D. Epema and H. Sips, “The Bittorrent P2P File-Sharing System: Measurements and Analysis”. |
| [5] | M. Moore, “The semantic web: an introduction for information professionals,” Thomson Reuters, 2011. |
| [6] | C. Holmberg, S. Hakansson and G. Eriksson, “Web Real-Time Communication Use Cases and Requirements,” March 2015. [Online]. Available: https://tools.ietf.org/html/rfc7478. |
| [7] | B. Ford, P. Srisuresh and D. Kegel, “Peer-­to­-Peer Communication Across Network Address Translators”. |
| [8] | D. Draheim, M. Felderer and V. Pekar, “Weaving Social Software Features Into Enterprise Resource Planning Systems”. |
| [9] | B. Dodson, I. Vo, T. J. Purtell, A. Cannon and M. S. Lam, “Musubi: Disintermediated Interactive Social Feeds for Mobile Devices”. |
| [10] | H. Alvestrand, “Google Releases WebRTC source code,” [Online]. Available: http://lists.w3.org/Archives/Public/public-webrtc/2011May/0022.html. |
| [11] | “WebPack,” [Online]. Available: https://webpack.js.org/. |
| [12] | “Polling,” [Online]. Available: https://en.wikipedia.org/wiki/Polling\_(computer\_science). |
| [13] | “Most Popular Programming Languages and Frameworks,” [Online]. Available: https://insights.stackoverflow.com/survey/2017#technology. |
| [14] | “Mongoose,” [Online]. Available: http://mongoosejs.com/. |
| [15] | “Mlab,” [Online]. Available: https://mlab.com/. |
| [16] | “Making Web Annotations Dynamically Robust and Semantically Rich”. |
| [17] | “IndexedDb,” [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/IndexedDB\_API. |
| [18] | “hypothes.is,” [Online]. Available: https://web.hypothes.is/. |
| [19] | “Heroku Cloud Application Platform,” [Online]. Available: Heroku.com. |
| [20] | “genius.com,” [Online]. Available: https://genius.com/web-annotator. |
| [21] | “Express Fast, unopinionated, minimalist web framework for Node.js,” [Online]. Available: https://expressjs.com/. |
| [22] | “Chrome Message Passing,” [Online]. Available: https://developer.chrome.com/extensions/messaging. |
| [23] | “Chrome Background Pages,” [Online]. Available: https://developer.chrome.com/extensions/background\_pages. |
| [24] | “ Real-time bidirectional event-based communication.,” Socket.io, [Online]. Available: Socket.io. |

# Appendix 1 –