# Exposé

## Enabling Compute-to-Data through verifiable Off-chain computations in Blockchain-based Data Trading Platforms

Kevin Hertwig, 388430

Technische Universität Berlin, Straße des 17.Juni 135, 10623 Berlin, Germany
pressestelle@tu-berlin.de

**Keywords:** Data Marketplace · Blockchain · zkSNARKs · ZoKrates

## 1 Context

With the proliferation of new technologies in recent years, the amount of data increases with a potential to revolutionize essential aspects in our world. In fact, the worldwide interaction with data follows an exponentially growing trend, with a volume of 79 Zettabytes (ZB) in the last year, and an expected volume of 181 ZB in 2025 [17]. This giant pool of data can help companies to develop new business models, make better decisions through analytics and build smarter applications with machine learning models. For example, access to data can enhance the treatment of diseases in the healthcare sector [18,33,29] and increase productivity and efficiency in the agriculture sector [13]. More and more connected Internet of Things (IoT) devices accelerate the collection of valuable data [24,20]. However, without open and censorship free access to data, it cannot be used. Consequently, the economic value of data promotes the need for online data marketplaces [7,19].

A conventional data marketplace is a two-party trading platform, where data owners, referred as sellers $S$, monetize their valuable data to get compensated for data access by interested parties, referred as buyers $B$ [2]. Such data marketplaces come in different variations. For example, there are paid subscription models for an interface to dynamic real-time data in contrast to a one-time-purchase model for access to a static resource. In addition, data marketplaces can be classified as Business-to-Business (B2B) or Business-to-Consumer (B2C) platforms. However, all conventional two-party data marketplaces have one thing in common – it is not possible to guarantee fair exchange, i.e. receiving legitimate data as $B$ while receiving the agreed-upon payment as $S$, without a Trusted Third Party (TTP) [2]. Hence, the two-party model is typically extended by a centralized trading platform, where $S$ uploads and advertises his or her data, and the platform sells the data on behalf of $S$ [2,8,29].

Unfortunately, centralized data trading platforms suffer from a variety of issues. A dishonest $S$ may be tempted to refuse data access or return manipulated data after receiving the payment, harming data availability and integrity [29,20]. On the opposite, a dishonest $B$ may be tempted to never pay the price after receiving the data [20]. This results into a trust problem between $S$ and $B$ that a centralized trading platform should solve as a middleman. However, a malicious platform might take advantage of its monopoly to advertise products and distort rankings for their own profit [26]. Even worse, a malicious platform has access to advertised data, breaching privacy, [2] and might resell the data without $S$'s knowledge [27,29,8]. All in all, the centralized platform is a Single Point of Failure (SPOF) [8] that is not able to satisfy fairness, security, privacy and non-discrimination, among others [2]. A decentralized infrastructure might help to reach all desirable properties [26]. Blockchain technology, first introduced in 2008 by Satoshi Nakamoto [22], provides a promising approach to that.

## 2 Problem & Research Question

Any digital data marketplace, whether centralized or decentralized, needs some specific key components and features. This includes: (i.) *Fairness*; (ii.) *Transparency*, *Privacy* and *Security*; (iii.) *Regulation*; as well as (iv.) *Efficiency*, according to Banerjee and Ruj [2]. Ramachandran et al. [26] complements this by functional requirements such as: (v.) *Posting and Discovery*; (vi.) *Data*

*Transfer and Payments*; (vii.) *Metadata Organization*; (viii.) *Data Quality - Buyer and Seller Ratings*; (ix.) *Data Quality - Curation and Recommendations*; and (x.) *Identity - and Access Control Management (IAM)*. These requirements are surrounded by the buyer and seller, as depicted in Figure 1.

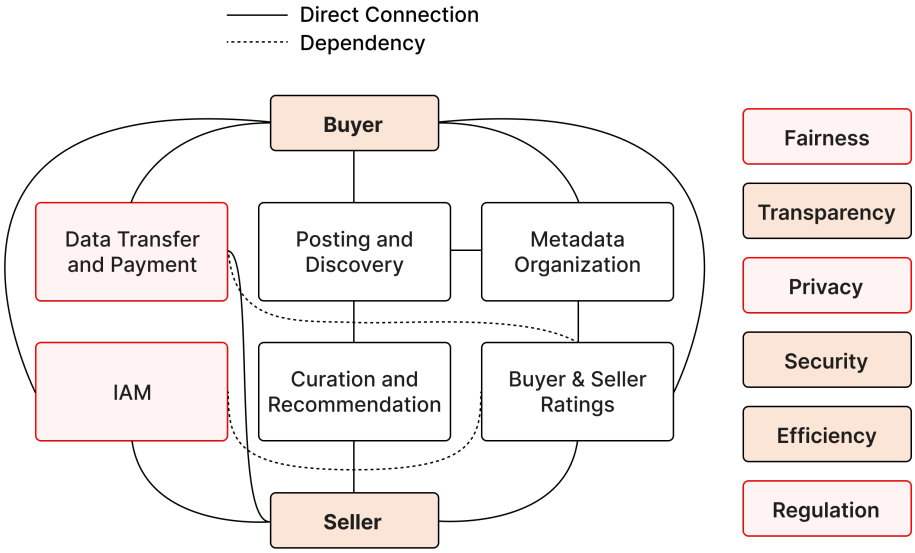


**Fig. 1.** Key components and features of a digital data trading platform. All components in red show the aspired enhancements by the contributions of this thesis.

"Blockchain-based application architectures benefit from a set of unique properties including immutability and transparency of cryptographically-secured and peer-recorded transactions, which have been agreed upon by network consensus" [11]. According to that, it is trivial to implement sufficient *Transparency* and *Security* in blockchain-based data marketplaces. Furthermore, it turns out the non-trivial *Fairness* problem in a two-party trading relationship seems to be solved by Blockchain technology in one atomic swap [9,21]. However, *Privacy*, *Regulation* and *Efficiency* remains a bigger problem for blockchain-based data marketplaces due to limitations of public permission-less Blockchains such as Bitcoin [22] and Ethereum [6].

Public permission-less Blockchains are peer-to-peer networks which inherently validate and process transaction data at every node, in order to guarantee network consensus. By virtue of its design, all data is available at each node that makes this system purposely public in favour of transparency properties. However, data sharing often incorporates Personal Identifiable Information (PII) and confidential data. Consequently, storing and computing private data on public permission-less Blockchains is conflicting with Privacy requirements, enforced by regulations such as the European General Data Protection Regulation (GDPR) [14]. Nevertheless, in any case it is not advisable to store large amounts of data in Blockchains due to block size limitations, Blockchain bloating and extremely high associated storage costs.

Off-chain storage solutions are suggested to overcome this limitation. In particular the Content-Addressable Storage Pattern [11] seems to be a reasonable solution to store large amounts of data while keeping key properties of Blockchains such as immutability. Decentralized peer-to-peer networks such as IPFS [3], Filecoin [25] and SWARM [32] provide a promising approach to that. However, these networks suffer from the same Privacy issues as Blockchains due to open access and data replication across the network.

Another problem seems to be the enforcement of access and usage regulations in a distrusted setting without a TTP. A seller loses full control over his data when the buyer receives it. Hence, he or she is not able to enforce geographical usage restrictions or purpose limitation for example. This especially violates the principle of purpose limitation codified in Art. 5 of the GDPR [14, Art. 5 (1 b)]. Given a malicious buyer, he or she might even resell the purchased dataset without the seller's knowledge.

I observe, that most of the problems occur when data leaves the boundaries of the seller, i.e. a raw dataset is moved to a storage platform and/or moved to the buyer, where it subsequently is processed. While storage obstacles might be addressed by data encryption, *Regulation* and *Privacy* still remains a problem, when the buyer decrypts the raw dataset. According to that, I suggest to flip the strategy, i.e. raw datasets never leave the boundaries of the seller. Hence, the seller only moves the computed result of statistical queries and aggregations to the buyer, thereby protecting PII and confidential data. This paradigm is referred as *Compute-to-Data*. However, a problem of this paradigm is that, given a malicious seller, the buyer cannot verify if the result has been computed correctly. This leads to the following research questions:

1. What components are mandatory for a blockchain-based data trading platform and how to apply the Compute-to-Data paradigm on them?
2. How can the Compute-to-Data paradigm be made verifiable without sacrificing for privacy?
3. How practical is the proposed solution?

## 3 Related Work

Blockchain-based data trading platforms have been researched for a long time. There are a variety of different approaches where Blockchains are primarily used in almost every implementation (i.) as a transparent tamper-proof log of transactions; (ii.) as an enforcement point for access control decisions; and (iii.) as a medium to ensure fair exchange. Off-chain computation is then leveraged in conjunction to enhance privacy, ownership and regulation of datasets.

In the most simple implementation in [24,26,2], sellers encrypt their dataset before publishing it to the Blockchain or to a different storage system. Decryption keys are then distributed securely to the buyer after purchasing a dataset. However, when the buyer decrypts the dataset, regulation and privacy can't be guaranteed anymore.

Truong et al. rely on encrypted datasets as well, however with a more sophisticated key-distribution scheme [30]. According to that, prefix encryption, as a form of Hierarchical Identity-Based Encryption (HIBE), links the decryption key to an identity, and therefore simplifies key distribution and allows fine-grained access control. Nevertheless, privacy and regulation remains a problem after decrypting the dataset.

Serrano and Cuenca [27] enhance data ownership, privacy and misusage against regulations with Homomorphic Encryption (HE). According to that, the buyer can perform arbitrary computations on encrypted datasets, while the result is decrypted by the seller. Hence, the dataset never leaves the boundaries of the seller.

A different approach is followed by Enigma [1] which uses Secure Multi-Party Computation (sMPC or MPC) for data queries to build a platform with guaranteed privacy by design. With MPC, every party in the protocol has only access to a meaningless piece of data, whereas only the buyer finally receives the result of the computation.

Dai et al. proposes a secure data trading ecosystem (SDTE) [8] where buyers pay for the analysis of the seller's dataset in the form of statistical queries and aggregations. The processing is protected in a Trusted Execution Environment (TEE), specifically using Intel's Software Guard Extension (SGX) enclaves. Hynes et al. [16] as well as Xiao et al. [34] follow similar approaches, while the latter adds (i.) a novel way to commit the computation result onto the Blockchain; and (ii.) a verifiable proof to show compliance to data usage policies as a consumer. However, TEE's unfortunately have been shown to be susceptible to side-channel attacks [5,4].

The most privacy preserving approach is followed in [15] by Fotiou et al. with a computation result based on Differential Privacy (DP). "DP addresses the paradox of learning nothing about an individual, while learning useful information about a population" [31]. This implicitly introduces an error margin into the computation result. However, this paper does not implement output verifiability. Albeit, [23,31,18] show that output verifiability with DP is possible by the use of Zero-knowledge proofs (ZKP).

## 4 Solution Approach

This thesis presents the implementation of a practical real world data marketplace for verifiable statistical computations on private datasets. It addresses *Privacy*, *Fairness* and *Regulation* problems

of blockchain-based data trading platforms, while the focus is on *Privacy. Furthermore, I only focus on static tabular datasets with very infrequent changes. Hence, I do* ***not*** *focus on real-time streaming data and training of machine-learning models.* My proposed implementation implicitly targets the *Data Transfer* and *Payment* process as well as *IAM* – some of the fundamental functional requirements, as already depicted in Figure 1. Specifically, I construct a secure blockchain-based data trading ecosystem, using Blockchain as a medium to (i.) prevent single-point of failure; (ii.) define data usage policies; (iii.) create a transparent, non-repudiable and tamper-proof log of transactions; and (iv.) enforce a fair data exchange protocol. However, the *focus* of this thesis is rather on creating a generic mechanism to make a variety of computations on arbitrary private datasets verifiable, and Blockchain is a necessary part of the puzzle. This thesis uses Verifiable Off-chain Computation (VOC) to address this problem [10,11].

VOC is a derivative of Verifiable Computation and aims to secure the integrity of computations performed by untrusted parties off the Blockchain. The result of the computation is then published to the Blockchain and verified on-chain with a cryptographic proof, attesting its correctness. Offloading the computation has multiple benefits – (i.) it increases the scalability by avoiding complex redundant computations on each node; (ii.) it reduces on-chain transaction costs by significantly lowering the size of transactions; and (iii.) it improves privacy by hiding PII and confidential data from the public ledger. [10,12,28,35]

According to [10], a reasonable VC scheme for off-chain computations needs to fulfill the following requirements: (i.) non-interactivity; (ii.) cheap verification; (iii.) weak security assumptions; and (iv.) zero-knowledge. ZkSNARKs, ZkSTARKs and Bulletproofs provide a valid approach to the aforementioned requirements. ZkSNARKs are a type of Zero-knowledge proof (ZKP) that are *non-interactive* and *succinct*. *Non-interactivity* defines the possibility to convince a verifier of a particular statement with only *one* message [10,12,28]. *Succinct* defines a proof that is small in size, compared to ZkSTARKs and Bulletproofs, and can be verified cheaply and quickly, typically within a few milliseconds [28]. This thesis uses a ZKP with ZkSNARKs for verifiable computations on private datasets.

The proposed blockchain-based data trading platform is designed for two party relationships between a buyer $B$ and seller $S$. The current high-level protocol for a trade between $B$ and $S$ on the proposed platform is described by the following exemplary use case and shown in the subsequent Figure 2, without technical details.

**Example** S advertises a dataset $x$ about health information of a population on the marketplace. The dataset $x$ is described by metadata, i.e. the time interval, data category, amount of columns and rows as well as the header of each column, among other metadata. The actual content of the dataset is hidden from the buyer. A potential $B$ enters the data marketplace and is interested in receiving the average age of cancer diagnosed patients in 2021. He or she discovers a promising dataset $x$ of $S$ in the health category, containing all cancer diagnosed patients worldwide. $B$ chooses to purchase the desired computation $\phi(x)$ for the *age* column. He or she commits the purchase by an on-chain transaction, including the given price by $S$. All coins are now temporarily locked in a smart contract. $S$ subsequently gets notified by the purchase and calculates the average on a private node of $S$, protecting privacy and confidentiality of $S$'s dataset. $S$ transfers the result $z = Enc_{pb(B)}(y = \phi(x))$ to the Blockchain, encrypted with the public key of $B$. $B$ gets notified and decrypts the result with his or her private key $Dec_{sk(B)}(z)$. $B$ is now uncertain about the correctness of the result and therefore constructs a program $\pi$ for the seller to prove (i.) the correctness of the computation; (ii.) the result originates from the advertised dataset; and (iii.) the proof is given by the seller. He or she then publishes a smart contract for the verification of the proof. $S$ is now responsible to generate such a proof for the verifier smart contract, to transparently proof all the requirements of $\pi$. According to that, the execution of the protocol terminates in the following situations:

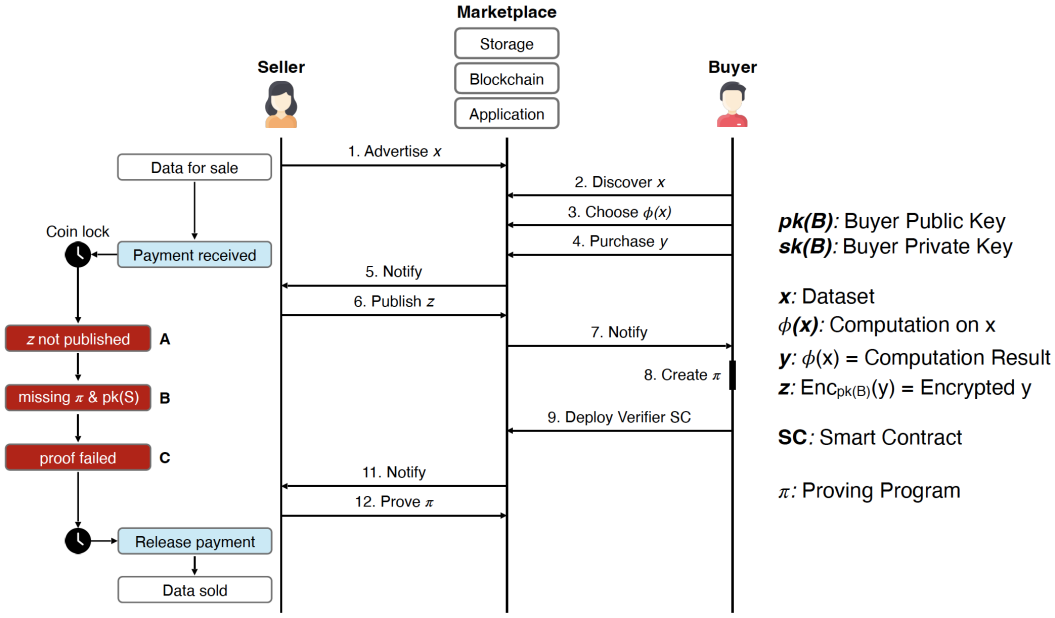


**Fig. 2.** Proposed protocol for a blockchain-based data trading platform with Compute-to-Data and Verifiable Off-chain Computation for verifiable statistical queries on private datasets. The marketplace between a buyer $B$ and seller $S$ is highly abstracted into storage, Blockchain and application components, which are necessary for the entire data marketplace ecosystem.

1. $S$ finally manages to submit a valid proof to the verifier contract. In this case, the payment is released to $S$.
2. $S$ does not publish a computation result. In this case, $B$ can withdraw his payment after a timeout. **(Scenario A)**
3. $B$ does not construct a proving program and does not publish the proving key to $S$. In this case, $S$ can withdraw the payment after a timeout. **(Scenario B)**
4. $S$ is not able to submit a valid proof to the verifier contract. In this case, $B$ can withdraw his payment after a timeout. **(Scenario C)**

**Note**: This protocol is only a first idea and the final protocol may highly differ. At the time of writing this exposé I already noticed that there is one major weakness in the current protocol. A malicious $B$ could send a fake proving key to $S$, so that it is impossible for $S$ to construct a valid proof. According to that, $B$ would always be able to withdraw his payment, if he has no honest interest in actually verifying the correctness of his purchased values. While this is not the focus of this thesis, I will try to create a completely trustless protocol between $B$ and $S$.

## 5 Implementation

The final outcome of this thesis will consist of multiple software artifacts, implementing a fully functional data marketplace, as described in the use case of the previous chapter 4. The combination of all software artifacts builds the blockchain-based data trading platform. According to that, the platform is composed of a public permissionless Blockchain, a Blockchain Indexer, a Marketplace User Interface, a decentralized Off-chain Storage Node, a private Compute Node and a Zero-knowledge Proof Service.

**Blockchain** The Blockchain is the single source of truth for all advertised datasets on the the marketplace user interface. It provides non-repudiable tamper-proof logs of transactions, everything visible via the user interface. All necessary smart contracts will be written using the Solidity[1]

[1] https://github.com/ethereum/solidity

programming language and deployed to an Ethereum Virtual Machine (EVM) compliant public permissionless Blockchain such as Ethereum itself, or a layer 2 network such as Polygon.

**Blockchain Indexer** Unfortunately, the Blockchain is a time-ordered append-only data structure, which makes it highly inefficient to query, i.e. filter, search, paginate and aggregate. However, for a practical data marketplace this is absolutely necessary. Blockchain Indexers typically provide an efficient protocol, to extract raw data from the Blockchain, process it and store it into some kind of database, as soon as a new block is written to the ledger. Additionally, Blockchain Indexers expose an API, to provide highly efficient and fast access to Blockchain data with all desired querying capabilities. The Blockchain indexer will be implemented with the Graph[2].

**User Interface** The user interface is the entry point to the blockchain-based data trading platform. A seller can advertise and monetize arbitrary datasets whereas a buyer can transparently browse the marketplace and buy interesting datasets. It will provide most of the functional requirements of Figure 1 and will be implemented as a Vue[3] 3 single page application (SPA) with commonly used libraries such as Ethers.js[4] and Hardhat[5] to efficiently build modern decentralized applications (dApps). To securely interact with smart contracts, I will use Metamask[6] as a wallet.

**Off-chain Storage** The off-chain storage node is not used to store the advertised raw datasets. It is rather used to store the dataset's metadata according to the Content-Addressable Storage Pattern[11], including a detailed description of the dataset and its hash. As a decentralized peer-to-peer storage network, the InterPlanetary File System[7] (IPFS) will provide highly available and immutable access to metadata. A content identifier (CID) is stored with every advertised dataset on-chain.

**Compute Node** The compute node is a private node of the seller to fulfill the purchase order of the buyer, i.e. computing a statistical query on a private dataset. This compute node will be implemented as a Node.js[8] application.

**Zero-knowledge Proof Service** The ZKP service is the focus of this thesis and will automatically generate proving programs based on the individual purchase of a buyer. It compiles high-level ZoKrates[9] code into an executable constraint system (ECS) using the Intermediate Representation (ZIR) of ZoKrates. Furthermore, it automatically generates an evidence key pair for proof creation and verification, also referred as the one-time setup. Finally, a buyer can send the proving key to the seller and deploys a smart contract for proof verification. The seller uses the proving key and the generated proving program to deliver such a proof.

## 6 Evaluation

The outcome of this thesis will be evaluated according to predetermined functional and non-functional requirements. As already described in the previous sections, this thesis aims to enhance *Privacy*, *Fairness* and *Regulation* in blockchain-based data trading platforms by combining Compute-to-Data and Verifiable Off-chain Computation. These 3 requirements are hard to measure numerically, but the evaluation will include an in-depth discussion about the pros and cons of the proposed system design accordingly, and compared to related work.

[2] https://thegraph.com/en/
[3] https://vuejs.org/
[4] https://github.com/ethers-io/ethers.js/
[5] https://github.com/NomicFoundation/hardhat
[6] https://metamask.io/
[7] https://docs.ipfs.tech/
[8] https://nodejs.org/en/
[9] https://github.com/Zokrates/ZoKrates

The second part of the evaluation will include multiple benchmarks with regards to the practicality, also referred as efficiency, of the proposed system design. Interestingly, *Efficiency* is one of the key features of any digital data trading platform as depicted in chapter 2. Therefore, I will analyze efficiency of on-chain and off-chain components individually, according to costs, scalability and computation time.

One of the most suitable measurements for on-chain components are the costs for each transaction, with regards to gas fees. This is important because the proposed system design is only feasible as a real-world system when costs are reasonably low for the buyer and seller. Since the Blockchain is primarily used to secure the protocol, costs can be compared to conventional buyer and seller protection systems from Ebay or PayPal for example. The most important on-chain transaction will probably be the verification of the zero-knowledge proof, which has to be cheaper than the on-chain execution in the first place. Benchmarks will vary in the size of the dataset and the computation algorithm.

The second analysis includes the evaluation of off-chain components. According to that, the ZKP, which is composed of different phases, is the most important evaluation. I will benchmark especially the time for the one-time setup as well as the proof generation time. The proof generation time will probably have the most impact on the practicality of the proposed system design. Benchmarks will again vary in the size of the dataset and the computation algorithm. Furthermore, I will take the underlying hardware with regards to the computational power into account for this analysis.

The entire system can be measured in the overall time and amount of exchanged messages until the trade between buyer and seller is completely fulfilled. All benchmarks can be compared to conventional data marketplaces as well as related work, and used to construct suggestions to improve the system design in future work.

## 7 Timeline

The Gantt chart in Figure 3 shows my expected workload for each individual chapter of the master thesis. Since scientific research is usually not a sequential process, the chart should rather serve as an indicator for the volume I dedicate to each chapter. Nevertheless, I will still try to stick to the timeline as best as possible, to stay focused and to have a feeling about the current progress.

Throughout the entire time period of six months, ongoing literature reviews as well as thesis writing is planned. In the beginning, at least six weeks are dedicated to extensive research and writing for the Introduction, Background, Related Work and Requirements chapter. This serves as a knowledge foundation to come up with a feasible concept and design for the implementation of a Proof of Concept (PoC). The implementation phase is split up into four smaller agile development cycles with each having a length of one month. Every development cycle produces a usable outcome, starting with a Minimum Viable Product (MVP). The MVP has only minimal functionality and might not already implement the core of the thesis. However, the following PoC should already implement the vast majority of the desired features. With two more releases, there should be enough time to refine and finalize the PoC.

Towards the end, I will evaluate and conclude my thesis and gather ideas for future work. Throughout the entire processing time, I will often find myself rewriting parts as a result of new insights. Accordingly, there is plenty of time planned for refinements.

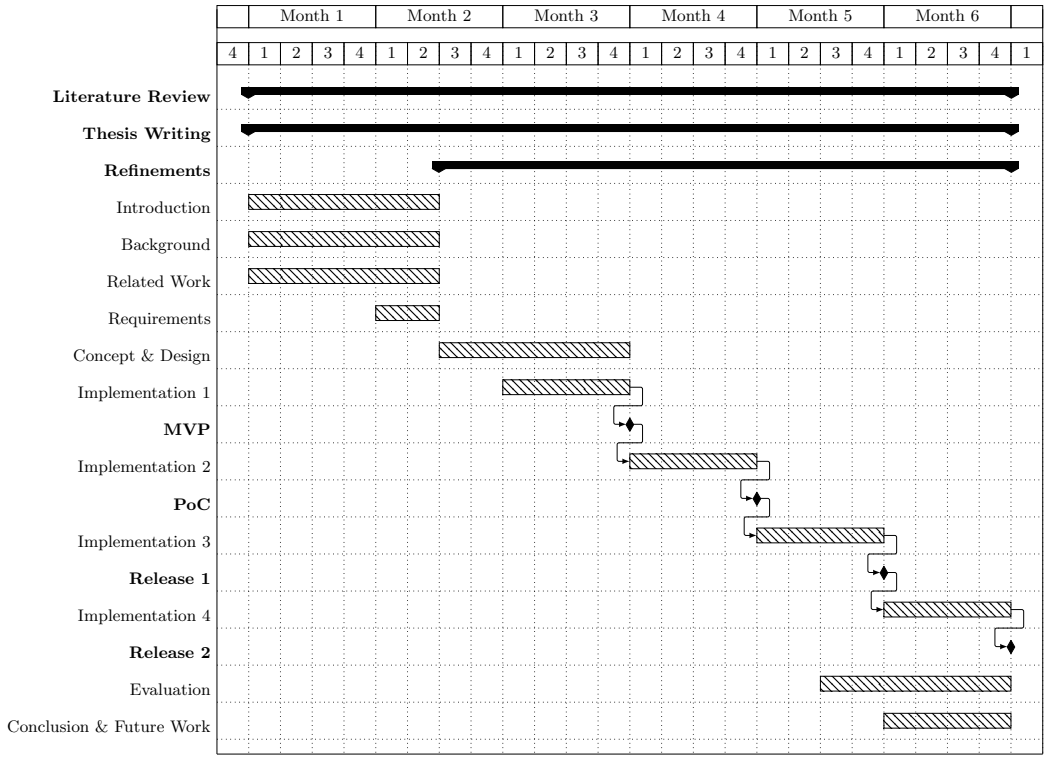


**Fig. 3.** Gantt Chart to visualize the expected Workload of individual Tasks in Scope of the Master Thesis

## References


1. Enigma: Decentralized Computation Platform with Guaranteed Privacy. In: Shrobe, H., Shrier, D.L., Pentland, A. (eds.) New Solutions for Cybersecurity. The MIT Press (2018). https://doi.org/10.7551/mitpress/11636.003.0018
2. Banerjee, P., Ruj, S.: Blockchain Enabled Data Marketplace – Design and Challenges (Sep 2019)
3. Benet, J.: IPFS - Content Addressed, Versioned, P2P File System (Jul 2014)
4. Biondo, A., Conti, M., Davi, L., Frassetto, T., Sadeghi, A.R.: The Guard's Dilemma: Efficient Code-Reuse Attacks Against Intel SGX p. 16
5. Brasser, F., Muller, U., Dmitrienko, A., Kostiainen, K., Capkun, S., Sadeghi, A.R.: Software Grand Exposure: SGX Cache Attacks Are Practical p. 12
6. Buterin, V.: A NEXT GENERATION SMART CONTRACT & DECENTRALIZED APPLICATION PLATFORM p. 36
7. Dageville, B., Cruanes, T., Zukowski, M., Antonov, V., Avanes, A., Bock, J., Claybaugh, J., Engovatov, D., Hentschel, M., Huang, J., Lee, A.W., Motivala, A., Munir, A.Q., Pelley, S., Povinec, P., Rahn, G., Triantafyllis, S., Unterbrunner, P.: The Snowflake Elastic Data Warehouse. In: Proceedings of the 2016 International Conference on Management of Data. pp. 215–226. ACM, San Francisco California USA (Jun 2016). https://doi.org/10.1145/2882903.2903741
8. Dai, W., Dai, C., Choo, K.K.R., Cui, C., Zou, D., Jin, H.: SDTE: A Secure Blockchain-Based Data Trading Ecosystem. IEEE Transactions on Information Forensics and Security **15**, 725–737 (2020). https://doi.org/10.1109/TIFS.2019.2928256
9. Dziembowski, S., Eckey, L., Faust, S.: FairSwap: How To Fairly Exchange Digital Goods. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. pp. 967–984. ACM, Toronto Canada (Oct 2018). https://doi.org/10.1145/3243734.3243857
10. Eberhardt, J., Heiss, J.: Off-chaining Models and Approaches to Off-chain Computations. In: Proceedings of the 2nd Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers. pp. 7–12. ACM, Rennes France (Dec 2018). https://doi.org/10.1145/3284764.3284766
11. Eberhardt, J., Tai, S.: On or Off the Blockchain? Insights on Off-Chaining Computation and Data. In: De Paoli, F., Schulte, S., Broch Johnsen, E. (eds.) Service-Oriented and Cloud Computing, vol. 10465, pp. 3–15. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-67262-5$_1$
12. Eberhardt, J., Tai, S.: ZoKrates - Scalable Privacy-Preserving Off-Chain Computations. In: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing

and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). pp. 1084–1091. IEEE, Halifax, NS, Canada (Jul 2018). https://doi.org/10.1109/Cybermatics$_2$018.2018.00199
13. Elijah, O., Rahman, T.A., Orikumhi, I., Leow, C.Y., Hindia, M.N.: An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges. IEEE Internet of Things Journal **5**(5), 3758–3773 (Oct 2018). https://doi.org/10.1109/JIOT.2018.2844296
14. European Commission: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) (2016), `https://eur-lex.europa.eu/eli/reg/2016/679/oj`
15. Fotiou, N., Pittaras, I., Siris, V.A., Polyzos, G.C., Anton, P.: A privacy-preserving statistics marketplace using local differential privacy and blockchain: An application to smart-grid measurements sharing. Blockchain: Research and Applications **2**(1), 100022 (Mar 2021). https://doi.org/10.1016/j.bcra.2021.100022
16. Hynes, N., Dao, D., Yan, D., Cheng, R., Song, D.: A demonstration of sterling: A privacy-preserving data marketplace. Proceedings of the VLDB Endowment **11**(12), 2086–2089 (Aug 2018). https://doi.org/10.14778/3229863.3236266
17. IDC, Statista: Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025 (Jun 2021), `https://www.statista.com/statistics/871513/worldwide-data-created/`
18. Koutsos, V., Papadopoulos, D., Chatzopoulos, D., Tarkoma, S., Hui, P.: Agora: A Privacy-Aware Data Marketplace p. 13
19. Krishnamachari, B., Power, J., Kim, S.H., Shahabi, C.: I3: An IoT Marketplace for Smart Communities. In: Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services. pp. 498–499. ACM, Munich Germany (Jun 2018). https://doi.org/10.1145/3210240.3223573
20. Lawrenz, S., Sharma, P., Rausch, A.: Blockchain Technology as an Approach for Data Marketplaces. In: Proceedings of the 2019 International Conference on Blockchain Technology. pp. 55–59. ACM, Honolulu HI USA (Mar 2019). https://doi.org/10.1145/3320154.3320165
21. Li, Y., Ye, C., Hu, Y., Morpheus, I., Guo, Y., Zhang, C., Zhang, Y., Sun, Z., Lu, Y., Wang, H.: ZKCPlus: Optimized Fair-exchange Protocol Supporting Practical and Flexible Data Exchange. In: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security. pp. 3002–3021. ACM, Virtual Event Republic of Korea (Nov 2021). https://doi.org/10.1145/3460120.3484558
22. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System p. 9
23. Narayan, A., Feldman, A., Papadimitriou, A., Haeberlen, A.: Verifiable differential privacy. In: Proceedings of the Tenth European Conference on Computer Systems. pp. 1–14. ACM, Bordeaux France (Apr 2015). https://doi.org/10.1145/2741948.2741978
24. Özyılmaz, K.R., Doğan, M., Yurdakul, A.: IDMoB: IoT Data Marketplace on Blockchain. In: 2018 Crypto Valley Conference on Blockchain Technology (CVCBT). pp. 11–19 (Jun 2018). https://doi.org/10.1109/CVCBT.2018.00007
25. Protocol Labs: Filecoin: A Decentralized Storage Network p. 35 (Jul 2017), `https://filecoin.io/filecoin.pdf`
26. Ramachandran, G.S., Radhakrishnan, R., Krishnamachari, B.: Towards a Decentralized Data Marketplace for Smart Cities. In: 2018 IEEE International Smart Cities Conference (ISC2). pp. 1–8 (Sep 2018). https://doi.org/10.1109/ISC2.2018.8656952
27. Serrano, N., Cuenca, F.: A Peer-to-Peer Ownership-Preserving Data Marketplace. In: 2021 IEEE International Conference on Blockchain (Blockchain). pp. 394–400. IEEE, Melbourne, Australia (Dec 2021). https://doi.org/10.1109/Blockchain53845.2021.00062
28. Simunic, S., Bernaca, D., Lenac, K.: Verifiable Computing Applications in Blockchain. IEEE Access **9**, 156729–156745 (2021). https://doi.org/10.1109/ACCESS.2021.3129314
29. Su, G., Yang, W., Luo, Z., Zhang, Y., Bai, Z., Zhu, Y.: BDTF: A Blockchain-Based Data Trading Framework with Trusted Execution Environment. In: 2020 16th International Conference on Mobility, Sensing and Networking (MSN). pp. 92–97. IEEE, Tokyo, Japan (Dec 2020). https://doi.org/10.1109/MSN50589.2020.00030
30. Truong, H.T.T., Almeida, M., Karame, G., Soriente, C.: Towards Secure and Decentralized Sharing of IoT Data. In: 2019 IEEE International Conference on Blockchain (Blockchain). pp. 176–183. IEEE, Atlanta, GA, USA (Jul 2019). https://doi.org/10.1109/Blockchain.2019.00031
31. Tsaloli, G., Mitrokotsa, A.: Differential Privacy meets Verifiable Computation: Achieving Strong Privacy and Integrity Guarantees p. 6
32. Viktor Trón, Aron Fischer, Dániel A. Nagy, Zsolt Felföldi, Nick Johnson: Swap, Swear and Swindle - Incentive System for Swarm p. 28 (May 2016), `https://ethersphere.github.io/swarm-home/ethersphere/orange-papers/1/sw%5E3.pdf`

33. Wang, Y., Kung, L., Byrd, T.A.: Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technological Forecasting and Social Change **126**, 3–13 (Jan 2018). https://doi.org/10.1016/j.techfore.2015.12.019
34. Xiao, Y., Zhang, N., Li, J., Lou, W., Hou, Y.T.: PrivacyGuard: Enforcing Private Data Usage Control with Blockchain and Attested Off-chain Contract Execution (Jul 2020)
35. Xu, C., Zhang, C., Xu, J., Pei, J.: SlimChain: Scaling Blockchain Transactions through Off-Chain Storage and Parallel Processing (Technical Report) p. 15