The Semantic Web and Blockchain at a Meeting Point

Klevis Shkembi [1], Petar Kochovski [1], Thanasis G. Papaioannou [2], Anthony Simonet-Boulogne [3], Marco Ciaramella [4], Caroline Barelle [5], Vlado Stankovski [1]

[1]*Faculty of Computer and information Science, University of Ljubljana, Slovenia*
[2]*Department of Informatics, Athens University of Economics and Business (AUEB), Greece*
[3]*iExec, France*
[4]*IntelliSemantic srl, Rivoli, Italy*
[5]*European Dynamic, Luxembourg*
E-mail: klevis.shkembi@fri.uni-lj.si

Abstract

The Internet is becoming more centralized, more asymmetric in terms of information and power distribution, more biased, less secure and less trustworthy. Blockchain technologies already allow the decentralized exchange of digital assets in a secure and fair manner, but its application to information transmission is mostly unexplored. This article outlines our vision for ONTOCHAIN, a semantically enhanced blockchain software ecosystem that enables the creation of secure distributed applications that empower users, ensure their privacy, high quality of service, and ultimately encourage pluralism and democracy. The primary goal of ONTOCHAIN is to achieve trustworthy service exchange and content handling for a variety of disciplines, including health, economy, public services, energy and sustainability, news, media, entertainment, industry 4.0, and tourism, employing advanced knowledge management mechanisms.

1 Introduction

The Internet's success is based on free expression, open innovation, and interoperability. However, there are rising worries that the pursuit of high performance and profit is compromising openness, trustworthiness, privacy, and security. When users connect with internet services, a number of risks have been discovered: the centralization of power (*.e.,* only a few actors have access to information and expertise), undetermined provenance of information (*e.g.,* fake news), anonymity in the service of criminal behavior, invasions of personal privacy, misuse of personal data (*e.g.,* Cambridge Analytica scandal), no equitable incentives for high-quality contributions and people' fundamental rights are being compromised.

Distributed Ledger Technologies (DLT) provide decentralization as a major feature such as blockchain [1]. Blockchains are "trustless", in the sense that the systems in place allow all participants in the system to agree on what the normative truth is without relying on any third party. Therefore many network stakeholders (*e.g.,* developers, miners and users) share authority and confidence, rather than putting them in one company or individual. However, blockchain has so far not covered identity management, trustworthiness assessment of data and entities, data management confidence, smart contracts that understand data semantics, secure data exchange or secure storage. Furthermore, it suffers from the Scalability Trilemma, which refers to the tradeoffs between decentralization, security, and scalability that crypto projects must make when deciding how to optimize the underlying architecture of their own blockchain and was coined by Vitalik Buterin (founder of Ethereum) [4].

This document argues our plan to build a multi-layer, modular blockchain framework that will enable the implementation of a variety of next-generation real-world solutions, including trustworthy web and social media, trustworthy crowdsensing, trustworthy service orchestration, unsupervised, decentralized online social networks, empower practitioners to handle the many issues of the Internet with numerous ledger technologies. The proposed blockchain-based architecture is anticipated to improve performance and scalability by including various business logic, access mechanisms, and governance structures. Our technology framework will serve as a foundation for the Next Generation Internet (NGI), which will support values such as openness, decentralization, inclusiveness, and privacy protection while also returning control to end-users so that they can benefit from democratic, transparent, and trustworthy decision-making mechanisms.

The remainder of this paper is organized as follows: In Section 2 we describe our approach to build a semantically-enriched, trustworthy blockchain ecosystem. In Section 3 we overview the architecture of our framework and describe its main components. In Section 4 we explain, how our technology framework enables promising use-case scenarios. In Section 5 we explain the potential various impact that our proposed ecosystem may have. Finally in Section 6 we conclude our work and outline future work.

2 Strategic objectives

As the Internet of Things (IoT) emerged, we have the ability to achieve new levels of decentralization, but also to collaborate amongst diverse cyber physical systems based on the concepts of the semantic web. Blockchain technologies, with their primary features of decentralisation, traceability, and transparency, are well suited to this agenda and may help to the successful functioning of such smart applications and systems [2]. The objective of this work is to build a single, shared ledger for the administration of shared ontological concepts, including instances of such concepts, using the fundamental fea-

tures of blockchain. ONTOCHAIN will have the ability and feature to interlink off-chain data, information, and AI services with on-chain information that eliminates the need for costly on-chain activities while introducing major new characteristics such as traceability, privacy and democratic processes. Different 'entities' memberships can be anonymous, but they can also be connected to real-world identities when the law requires it. They can be directly "executable" through the use of various semantic reasoners operating directly on blockchain, as well as potentially employing trustworthy off-chain real-world data (e.g., IoT) through the use of Smart Oracles and Decentralized Oracles. Overall, ONTOCHAIN vision is presented in Figure 1.

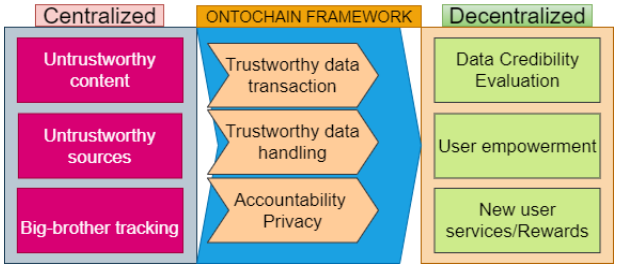

Figure 1: Illustration of the ONTOCHAIN concept.

3 Concept and architecture

The goal of this innovation action is to engage the broader research and innovation community in disciplines such as semantic web, artificial intelligence, security, privacy, distributed systems, game theory, and economics. Like

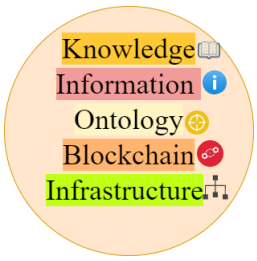

Figure 2: Achieving trust with ONTOCHAIN's technology.

blockchain, the Semantic Web is designed to build trust in highly decentralized, semantically sophisticated, and dynamic environments. We intend to build an innovative new technology stack that functions similarly to Semantic Web technologies and seeks to improve trust in semantics, data, knowledge, and information on the internet today (see Figure 2).

This framework will enable the implementation of a number of different next-generation real-world solutions applications. The various protocols will be used to create use-cases as shown in Figure 3. ONTOCHAIN Application and Core protocols will integrate interactions across various blockchain frameworks while masking them from use-cases to enable seamless inter-service process collaboration (see Figure 4). Furthermore, data saved on separate chains may be connected. This connection is stored in new ONTOCHAIN chains. We use a modular architecture to provide scalability, accessibility, and great performance. Each layer's modules and functionality are based on the functionality provided by the lower levels. The functionality of the modules at each layer, as well as their dependencies, are explained in a top-down approach below.

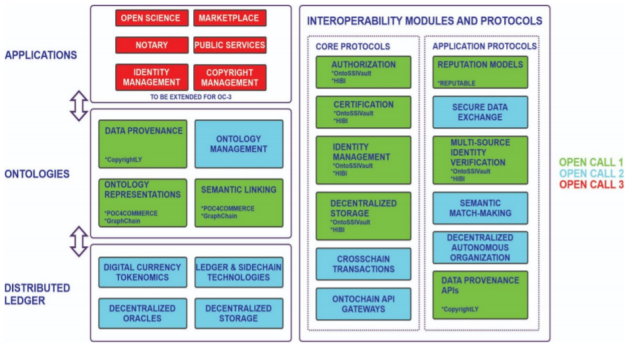

Figure 3: ONTOCHAIN architectural framework.

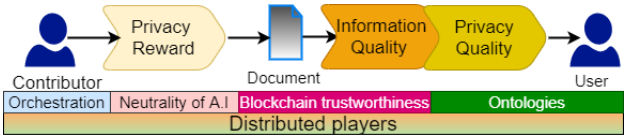

Figure 4: ONTOCHAIN's vision and approach.

3.1 Use-case layer

Trustworthy Information Exchange: To assess the trustworthiness of information, this use case employs and integrates data provenance techniques, decentralized oracles, and user trustworthiness. Secure data exchange methods are used to transmit data safely across transacting parties using cryptographic processes, while decentralized reputation models are used to judge the trustworthiness of data sources and the data itself.

Trustworthy and Transactional Content Handling: This use case provides safe and privacy-aware data querying by combining any of the following: permitted data access/handling, data credibility evaluation, copyright implementation, and secure and privacy-aware data querying. This use case also involves the safe transfer of any financial assets between parties participating in a data exchange. It will establish and create tools and processes that would allow regulatory, judicial, and law enforcement authorities to examine and otherwise impact data transactions in clearly specified situations as envisioned by the legislator.

3.2 Application Protocol Layer

Data Provenance: This module will offer graphical interfaces for accessing and displaying provenance data regarding on-chain and off-chain data from ONTOCHAIN. The full chain of transactions that resulted in a record will be included in the provenance information.
Reputation Models: This module will allow users to construct several decentralized reputation models on the blockchain, ensuring that reputation feedback is real, trustworthy, and anonymous.
Decentralized Oracles: Although Smart Contracts can only read and write data that is kept on their blockchain by design, this module will make it easier for them to work with off-chain data. Some solutions (*e.g.,* Substrate, ChainLink) look at a data source with many instances, then perform a consensus protocol on-chain to confirm the outcome. However, this only shifts the centralization point from the Oracle to the data source.
Market Mechanisms: This module facilitates market mechanisms by providing the fundamental support mechanisms for facilitating data/service transactions. It will enable physical asset trading (e.g., tokenization) and price determination (*e.g.,* auctions, negotiation protocols, etc.), as well as billing, customer support, inventory management, and other services.
Secure Data Exchange: This module covers data exchange functions among dispersed parties, confirming data ownership and access rights, the authenticity of the parties involved, the integrity of the information transferred and the secrecy of the data by means of blockchain underlying mechanisms. Off-chain data is often transferred via data transactions, but on-chain data stores public cryptographic keys and access control lists based on which particular transacted parties are granted enhanced data access to distinct sections of data.
Ontology Representation: This module aims to propose new methods for building ontologies using blockchain technology. Similar to the formation of axiomatic assertions, semantic agreements may be widely agreed upon via blockchain-based consensus. This module will also cover any reasoning techniques, tools, or procedures that may be used to derive new information from a sensing IoT-enabled environment.
Multi-Source Identity: This module aims to use newly built ONTOCHAIN services to register and validate individual digital IDs of physical items. Various AI approaches, for example, may be created to operate on sensing data (IoT-based, sensors, cameras, and similar) to determine if a human belongs to a certain ontological notion.

3.3 Distributed ledger

Blockchain consensus engine: Any blockchain relies on consensus mechanisms to operate. ONTOCHAIN will be scalable, open, cost-effective and energy-efficient, and even a significantly improved new consensus engine if possible. In terms of accessibility, ONTOCHAIN intends to be an open distributed ledger that may theoretically be coupled with multiple blockchain settings, rather than a silo blockchain ecosystem.
Decentralised Storage: With the rise of blockchain, we've seen a new generation of participative storage repositories that can help to solve security and privacy concerns, and store any type of data (*eg.g,* StrorJ). In the near future one may envision new storage services which can help to encrypt and decentralize private data, assist to manage reliability replicas and service quality, while balancing the transactions with the storage costs.

4 Use Cases

ONTOCHAIN will enable a variety of future applications in many sectors, ranging from B2B to C2C to G2C, like: art, commerce and trading, education, finance, healthcare, industry, public sector.
P2P e-commerce use case scenario: users can search for products/services supplied by business entities using data semantics. Users obtain the most trustworthy provider's product/service that suits their requirements. The product/service transaction, as well as any changes in ownership or access provision to a service, is then recorded on the blockchain. At the same time, the product/service is withdrawn from the seller's inventory, and an invoice is generated automatically. Finally, users may provide feedback on previous transactions.
Data Trading use case scenario: Data is securely kept, uniquely identifiable, and only authorized organizations have access to it for pre-determined processing in this application. Predetermined algorithms, which are also part of the smart contract, can be used to handle data in safe enclaves without exposing the original data to a third party. The smart contract's General Data Protection Regulation (GDPR) compliance is automatically verified, and any data processing operations are stored in the blockchain. Any data modifications that result in new data are considered as tradeable assets, with their provenance determined by their relationship to the original data and the processing methods.
The ONTOCHAIN ecosystem will be developed on top of Ethereum layer 1 for the time being. Later on, it will be upgraded to a layer 2 blockchain network, with each component having its own shard.

5 Impact

ONTOCHAIN's vision is to contribute to NGI's success and to accelerate Europeans' potential to grow individually and collectively in this new era, as well as to significantly alter our perception of the digital world around us. The following are the key points of ONTOCHAIN's immediate impact:

- Growing the blockchain and next-generation communities, throughout the duration of a three-year project involving numerous vertical and horizontal actions aimed at community development.
- Educating and financing a large number of researchers, innovators, start-ups, SMEs, and use-case holders (practitioners) in order to raise their knowledge of

the benefits of a more trustworthy and decentralized internet.

- Providing a platform for the development and validation of new technological solutions and innovative decentralized applications, as well as the introduction of new economic schemes oriented toward a human-centric future Internet progression.
- Developing the necessary proof-of-concept and acceptable expertise in order to provide instructions and a roadmap for a more focused and efficient European policy-making process.

5.1 Shape a more human-centric evolution of the internet

ONTOCHAIN seeks to decentralize knowledge and power while preserving people's anonymity. ONTOCHAIN will provide an unbiased, non-institutional trust rating of content people and services. Furthermore, since ONTOCHAIN is a human-centric ecosystem, its success would have an impact on its future adoption by a huge user community.

5.2 European blockchain ecosystem integrating research and innovation communities

ONTOCHAIN activities and events will serve as a starting point for blockchain researchers and innovators. The initiative will promote concept and project co-creation, as well as a high level of integration between research and development communities. Moreover, ONTOCHAIN will contribute to the work of the European Blockchain Observatory Forum, which is taking place at institutional levels across the EU, by providing blockchain technology guidance. ONTOCHAIN will also participate in European and national blockchain events, which will provide opportunities for the project to interact with other local communities and bridge ideas and experience.

5.3 New business opportunities with maximum growth and impact chances

By providing the building blocks for new decentralized end-user apps, ONTOCHAIN aspires to develop a highly business environment. ONTOCHAIN creates ideal business opportunities for new internet companies by addressing core user concerns, such as trust and privacy, as well as facilitating business building blocks such as identity management, social media copyright management, knowledge management, supply chain management, decentralized reputation models, decentralized market mechanisms, data provenance and trustworthy data handling, and low-entry-cost infrastructure.

6 Conclusion

The ONTOCHAIN project is a NGI initiative aimed at improving data, metadata, semantics, knowledge, and information management technologies. This encompasses the activities of generating, systematizing, storing, transmitting, and otherwise managing trusted data. The project aims to combine the Semantic Web as a trust-building technological stack with blockchain technologies as an existing technology that builds trust based on decentralized consensus among strangers. Since both technologies are used to manage knowledge and generate trust, the ONTOCHAIN project believes that their effective marriage is essential.
In this document we presented the architecture and major features of our unique blockchain ecosystem. In addition, we presented how creative and interesting distributed applications are achievable with two example situations. As future work we intend to describe the ONTOCHAIN architecture in further detail, including comprehensive APIs between the various components as well as implementation details. In addition, we will show how the suggested platform may be used to create prototype implementations of example distributed applications.

Acknowledgment

The research and development reported in this paper have received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreements no. 957338 (ONTOCHAIN: Trusted, traceable and transparent ontological knowledge on blockchain).

References

[1] S.Nakamoto: A peer-to-peer electronic cash system. Bitcoin White Paper, 2008, https://bitcoin.org/bitcoin.pdf

[2] Kochovski, P., Gec, S., Stankovski, V., Bajec, M., Drobintsev, P.D.: Trust manage-ment in a blockchain based fog computing platform with trustless smart oracles.Future Generation Computer Systems101, 747–759 (2019)

[3] Casino, Fran, Thomas K. Dasaklis, and Constantinos Patsakis. "A systematic literature review of blockchain-based applications: Current status, classification and open issues." Telematics and informatics 36 (2019): 55-81.

[4] Wood, G., et al.: Ethereum: A secure decentralised generalised transaction ledger.Ethereum project yellow paper151(2014), 1–32 (2014)

[5] Yasin, A., Liu, L.: An online identity and smart contract management system. In:2016 IEEE 40th Annual Computer Software and Applications Conference (COMP-SAC) IEEE (2016). vol. 2, pp. 192–198.

[6] Christidis, K., Devetsikiotis, M.: Blockchains and smart contracts for the internetof things. Ieee Access4, 2292–2303 (2016)

[7] Dennis, R., Owen, G.: Rep on the block: A next generation reputation system basedon the blockchain. In: 2015 10th International Conference for Internet Technologyand Secured Transactions (ICITST). pp. 131–138. IEEE (2015)

[8] Dinh, T.T.A., Wang, J., Chen, G., Liu, R., Ooi, B.C., Tan, K.L.: Blockbench: Aframework for analyzing private blockchains. In: Proceedings of the 2017 ACMInternational Conference on Management of Data. pp. 1085–1100 (2017)

[9] Hamida, E.B., Brousmiche, K.L., Levard, H., Thea, E.: Blockchain for enterprise:overview, opportunities and challenges. In: The Thirteenth International Confer-ence on Wireless and Mobile Communications (ICWMC 2017) (2017)