

Web 3.0 Meets Web3: Exploring the Convergence of Semantic Web and Blockchain Technologies

Oshani Seneviratne^{1,*,\dagger}, Deborah L. McGuinness¹

¹*Rensselaer Polytechnic Institute, Troy, NY, USA*

Abstract

We outline the synergistic convergence of semantic web technologies, which have driven the advent of Web 3.0, and blockchain technologies, which have catalyzed the flourishing Web3 ecosystem. The integration of these technologies holds immense potential for transforming data representation, interoperability, and trust within decentralized knowledge graphs. The utilization of semantic web technologies enables the creation of machine-readable data formats, facilitating seamless understanding and exchange across heterogeneous systems. Complementing this, blockchain technologies provide an immutable and tamper-proof ledger, offering the foundation for establishing trust in decentralized knowledge graphs. We discuss the adoption of standardized vocabularies and smart contract powered schema alignment to enhance data exchange and integration with a focus on semantic interoperability, trustworthiness in semantic reasoning processes, and ownable and traceable resources.

Keywords

Semantic Web, Decentralized Knowledge Graphs, Blockchain, Smart Contracts

1. Introduction

Web 3.0, often referred to as the next generation of the World Wide Web, encompasses advancements in technologies such as artificial intelligence and the semantic web, whereas Web3 specifically represents the decentralized web with “ownable” resources built upon blockchain technologies.

A Brief History: The initial era of the World Wide Web (1990-2005) featured decentralized and community-governed open protocols, whereby the primary value was attributed to the network’s edges encompassing servers run by various individuals [1]. Conversely, the subsequent Web 2.0 epoch (2005-2020), popularized by Tim O’reilly [2], witnessed a shift towards siloed and centralized services operated by corporations, leading to the majority of value accruing to a few select entities such as Google, Apple, Amazon, and Facebook. Then, with the advent of the semantic web [3], the use of standardized vocabularies and ontologies to encode the

Trusting Decentralised KGs on the Web (TrustDeKW’23) at the Extended Semantic Web Conference (ESWC’23), May 28, 2023, Hersonissos, Greece


*Corresponding author.

✉ senevo@rpi.edu (O. Seneviratne); dlm@cs.rpi.edu (D. L. McGuinness)

🌐 <https://faculty.rpi.edu/oshani-seneviratne> (O. Seneviratne); <https://faculty.rpi.edu/deborah-mcguinness> (D. L. McGuinness)

🆔 0000-0001-8518-917X (O. Seneviratne); 0000-0001-7037-4567 (D. L. McGuinness)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

connections between the nodes popularized the notion of Web 3.0 [4]. Now, at the dawn of the Web3 [5], which is driven primarily by the use of blockchain technologies, the decentralized and community-governed principles of the original World Wide Web are being amalgamated with the sophisticated and contemporary functionalities of Web 2.0. This new era, characterized by token orchestration and user identities, is owned by both the users and builders and decentralizes all facets previously centralized by Web 2.0. Essentially, Web3 offers the potential to attain the richness of Web 2.0 but in a decentralized framework that was a cornerstone of Web 3.0 but never reached its full potential.

Re-decentralizing the Web: Efforts such as Solid (Social Linked Data) [6] aim to provide a framework for building decentralized applications by placing data ownership in the hands of the users, thereby increasing transparency, privacy, and security and offering an alternative to the centralized data models of Web 2.0. Some Solid pod providers allow users to rent storage space¹, and even governments may provide personal data storage services to their citizens.

New Advancements: Many decentralized applications, i.e., ‘Web3 dApps’, built on popular blockchains provide similar user experiences to the typical Web 2.0 applications. One of the main advantages of blockchain technology is that it provides a tamper-proof decentralized ledger that can be used to track and verify the authenticity of data. This process can enhance the trustworthiness of data by providing a secure and immutable record of data ownership and usage. Smart contracts, in particular, can automate the governance and management of decentralized data, enabling stakeholders to define and enforce rules and conditions for data access, sharing, and usage. The underlying incentive mechanism of the blockchain would ensure the system’s sustained longevity as it hits a critical mass. Moreover, blockchain technology can also provide a means for incentivizing participation and contributions to data at the application level, which can help to overcome the issue of limited adoption and engagement. By implementing tokenomics, users can be rewarded for their contributions, creating a self-sustaining ecosystem that incentivizes participation and promotes network growth.

2. Smart Contracts for Semantic Interoperability

The Semantic Interoperability Problem: Suppose there is a knowledge ecosystem that is being used to share data between different organizations in the healthcare sector. The knowledge graph contains patient medical records, drug interactions, and treatment protocols, and each organization has its internal data schema and ontology. However, sharing and combining data between these different organizations can only be easy with semantic interoperability. For example, an organization may use the term “heart attack” to describe a medical condition, while another organization may use “myocardial infarction.” If the two organizations were to share their data without semantic interoperability, it would be difficult to automatically identify that these terms refer to the same medical condition. Semantic interoperability techniques such as well-known ontologies and vocabularies make it possible to represent data in a standardized, machine-readable format that different systems can easily understand. This process helps ensure data is shared and processed correctly across the network, increasing trust and reducing the risk of errors or misunderstandings. In the healthcare example above, if everyone used a

¹<https://solidproject.org/users/get-a-pod>

shared medical ontology, it would be possible to ensure that medical terms and conditions are represented in a standardized way. This helps ensure that data is shared and processed correctly across different organizations, increasing trust in the data and reducing the risk of errors or misunderstandings. However, many organizations have their own terminology and schema, and entity alignment is easier said than done.

Smart Contracts Enabling Semantic Interoperability: Smart contracts can play a crucial role in facilitating schema alignment by providing transparency, trust, and automation to the process. For instance, smart contracts can embed alignment rules and logic, allowing the alignment process to be executed automatically when specific conditions are met. This automation reduces manual effort and minimizes potential errors in the alignment process. They enable multiple participants to verify and agree upon the alignment results, mitigating the need for centralized trust by incorporating consensus mechanisms. By introducing token-based incentives, smart contracts promote collaboration and engagement from a broader range of stakeholders, leading to more comprehensive and accurate schema alignments. Smart contracts can facilitate dynamic and decentralized schema alignment, where changes in one schema trigger corresponding adjustments in interconnected schema through predefined rules. This allows for real-time updates and adaptability to evolving data requirements and ensures consistent alignment across the network. By embedding validation rules within the smart contract, potential inconsistencies can be identified and addressed, ensuring the accuracy and integrity of the aligned schema.

3. Trustworthy Knowledge Inference and Data Consistency

Semantic Reasoning: Semantic reasoning refers to the ability of machines to infer new knowledge from existing data using logical rules and inference engines. By incorporating semantic reasoning, it is possible to automatically check the consistency of the data, identify errors and inconsistencies, and generate new insights from the data. This process can help increase trust in the data by ensuring it is accurate and up-to-date. However, interoperability and compatibility issues can arise when using different reasoning systems or integrating reasoning capabilities into existing infrastructures. Different reasoning engines may have variations in their underlying logic, rule languages, and inference algorithms, which can create challenges in aligning and harmonizing reasoning approaches, hindering the seamless integration of semantic reasoners in heterogeneous environments.

Semantic Reasoners with Smart Contracts: One potential approach to implementing semantic reasoners using smart contracts is breaking down the reasoning process into smaller, more manageable steps. Each step could be implemented as a separate smart contract, which could then be chained together to form a larger reasoning algorithm. Another approach could involve using off-chain computation to perform the majority of the reasoning process and then using a smart contract to verify the results and enforce the terms of the agreement or contract. There are, however, some scalability issues that need to be addressed. In the case of public blockchains like Ethereum, the need for every node to execute every contract invocation can lead to performance bottlenecks, as the network must reach a consensus on each transaction's execution. As the number of transactions and smart contracts increases, it can strain the network and result in slower transaction processing times. Various approaches such as layer 2

solutions, sharding, off-chain computation [7], and consensus mechanism optimizations are being explored to mitigate these scalability limitations.

Usecase: Consider a scenario where retailers share product information within a decentralized knowledge graph. The graph contains data about various products, including their specifications, pricing, and availability. Each retailer has their internal data schema and may use different terminology or classifications for similar products. In this case, semantic reasoners can be employed to identify inconsistencies in the product information. For instance, let's assume that one retailer lists a particular electronic device as a "smartphone," while another categorizes the same device as a "mobile phone." By inputting the data from multiple retailers into a semantic reasoner, it can detect such discrepancies and highlight the need for alignment. Semantic reasoning can infer that the device in question belongs to the same category, regardless of the different terms the retailers use. This inference can help ensure consistent and accurate product classification within the knowledge graph. If subsequent statements from different retailers introduce conflicting information, such as varying prices or contradictory specifications for the same device, the reasoner can flag these inconsistencies as errors.

4. Ownable and Traceable Resources

Web3 Resources: Unlike the traditional Web 2.0 model, where centralized platforms exert control over user data and resources, Web3 introduces decentralized systems that empower individuals to own and manage their digital resources directly. Such ownable resources refer to various digital entities, such as non-fungible tokens (NFTs), digital collectibles, and other digital representations of value or ownership. This shift towards ownable resources in Web3 promotes greater autonomy, security, and transparency, empowering users to have full control over their digital assets and fostering a more equitable and decentralized digital ecosystem.

Provenance Assertions: The provenance ontology [8] defines concepts, relationships, and properties to represent various aspects of provenance, such as entities, activities, agents, and the relationships between them. However, it does not provide a verifiable means for asserting the ownership and transfer of resources. By representing semantic web triple assertions as ownable resources on a blockchain, it allows for increased data transparency, traceability, and verifiability. Each triple assertion can be associated with a unique identifier and ownership record, ensuring that the origin and history of the assertion can be reliably traced back to its creator.

Resource Ownership Transfers with Smart Contracts: Smart contracts, as self-executing and tamper-proof agreements, can facilitate the implementation of usage control mechanisms. By incorporating smart contracts within the broader system architecture, it becomes possible to define and enforce rules and conditions for accessing and using data, including adherence to the data provider's wishes or requirements. However, usage control is a complex and multifaceted topic, and there are indeed significant challenges to be addressed. Effective enforcement of data providers' wishes and requirements requires a comprehensive approach that encompasses technical solutions and legal, ethical, and governance considerations.

5. Conclusion

In conclusion, the integration of semantic web (Web 3.0) technologies and blockchain technologies (Web3) holds great promise for advancing the capabilities and utility of the Web. Using standardized vocabularies and ontologies enables interoperability among nodes, promoting trust and minimizing errors in knowledge sharing. Despite the benefits, several challenges and limitations need to be addressed. Implementing semantic web reasoners using smart contracts requires substantial modifications to existing implementations, increased deployment costs, and smart contract language constraints. Future work should focus on overcoming semantic interoperability, scalability, and governance challenges while exploring synergistic integration with other emerging technologies to enhance functionality in our rapidly evolving data-driven world. The inherent transparency and immutability of blockchain technology offer valuable benefits, particularly in contexts where trust and auditability are critical. However, in domains like healthcare, where privacy and data protection are paramount, it is essential to establish strict control over sensitive information exposure. It is not necessary to store all data directly on the blockchain. Instead, off-chain storage solutions can securely store sensitive information while storing only necessary metadata or references (to the encrypted or anonymized data stored off-chain) on the blockchain.

The convergence of the Semantic Web and blockchain technology has significant potential to transform the Web into a more useful and trustworthy space, and further research can pave the way for addressing the challenges associated with scalability, interoperability, and governance while exploring the integration of other emerging technologies to enhance functionality and effectiveness in a rapidly evolving data-driven world.

References

- [1] T. Berners-Lee, R. Cailliau, A. Luotonen, H. F. Nielsen, A. Secret, The world-wide web, *Communications of the ACM* 37 (1994) 76–82.
- [2] T. O'reilly, What is web 2.0, " O'Reilly Media, Inc.", 2009.
- [3] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, *Scientific american* 284 (2001) 34–43.
- [4] J. Hendler, Web 3.0 emerging, *Computer* 42 (2009) 111–113.
- [5] S. Voshmgir, Token economy: How the Web3 reinvents the internet, volume 2, Token Kitchen, 2020.
- [6] A. V. Sambra, E. Mansour, S. Hawke, M. Zereba, N. Greco, A. Ghanem, D. Zagidulin, A. Abounnaga, T. Berners-Lee, Solid: a platform for decentralized social applications based on linked data, MIT CSAIL & Qatar Computing Research Institute, Tech. Rep. (2016).
- [7] M. Shukla, J. Lin, O. Seneviratne, Blockiot-RETEL: Blockchain and IOT based Read-Execute-Transact-Erase-Loop Environment for Integrating Personal Health Data, in: 2021 IEEE International Conference on Blockchain (Blockchain), IEEE, 2021, pp. 237–243.
- [8] T. Lebo, S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, J. Zhao, Prov-o: The prov ontology, *W3C recommendation* 30 (2013).