A middleware for systems consumes Ethereum data in soft real-time: a Semantic Web approach

1st Célio Márcio Soares Ferreira*, 2nd Charles Tim Batista Garrocho*, 3rd Carlos Frederico Marcelo da Cunha Cavalcanti*, 4th Jorge Sá Silva† and 5th Ricardo Augusto Rabelo Oliveira*

*DECOM (Computer Department), UFOP (Federal University of Ouro Preto), Ouro Preto, Brazil
†Department of Electrical and Computer Engineering (DEEC), University of Coimbra, Coimbra, Portugal
celio@linuxplace.com.br, charles.garrocho@ifmg.edu.br, cfredufop@gmail.com, sasilva@deec.uc.pt, rrabelo@gmail.com

Abstract—Blockchain is already advancing in journeys beyond cryptocurrency applications, and Ethereum, already called the world’s computer, is going on a path that intends to reinvent the internet or Web 3.0, currently the leading platform for deploying so-called distributed applications (DApp). The growth of these Dapps in different spheres of society, as Smart Cities and Industry 4.0 IoT applications demand new proposals and models that integrate Ethereum and its ecosystem data with existing datasets on the traditional Web. This work presents our efforts to apply ontologies representing an Ethereum network and ecosystem using the Semantic Web model to extract and link its data. We show EthExtras a new ontology that extends and simplifies the EthOn, and as proof of concept and sample of use, we design a middleware web that exposed as RDF graphs the Ethereum data in soft real-time.

Index Terms—Blockchain, Ethereum, Semantic Web, Smart Contract

I. INTRODUCTION

Ethereum is the second most popular Blockchain platform, and the ability to develop complete decentralized applications using its Smart Contract support already makes it called World Computer or Web3.0.

Ethereum is famous for managing new cryptocurrencies propositions or Tokens and Non-Fungible Tokens (NFT), using Smart Contracts to automate their transactions.

Ethereum has been the preferred development platform for a decentralized developer community coding Smart Contracts in the decentralized paradigm and using all ecosystems of decentralized feature services like the InterPlanetary File System (IPFS) to generate decentralized applications (DApp).

The motivation goes far beyond the security and reliability network provided by Blockchain [1]. The objective is to free the dependencies and control of big tech enterprises, centralized internet services, or governments.

All this growth in the use of Ethereum in different scenarios and domains of problems justifies the relevance in investigating new strategies of extracting data from its Blockchain network. The link with this data with datasets already found in the traditional Web enriches the queries generating new insights [2] [3].

One proposal to standardize models and integrate data is the Semantic Web. Proposed in 2001 by Tim Berners-Lee, Semantic Web mainly offers data models that can extend the World Wide Web, allowing machines and humans to work together. It uses ontology models to allow query data from the Web using SPARQL endpoints, linking datasets. Semantic Web is the proposition to simplify the consumption of Web data, allowing integration and cooperation man-machine. The Semantic Web is currently the W3C recommendation for integrating WEB data sources. Using Semantic Web technics enables standard access, integration, and query, expanding the possibilities of insights between different application domains.

Ontologies representing Ethereum give access to its Blockchain entities, such as accounts, Blocks, transactions, receipts, contracts, Tokens, NFT. EthOn [4] is an of ontologies found in the literature to represent the main Ethereum entities using RDF and Web Ontology Language (OWL).

EthOn formalizes the concepts and terms of the Ethereum, Blockchain in OWL, describing the Ethereum objects as classes in ontology. Its covers the major Blockchain and State Transition concepts as Blocks, Accounts, Transactions, Contract Messages, States, State Transitions; and the network concepts, Blockchain, Node, Protocol Variant, forking, Network.

This work presents our efforts to apply Semantic Web technics to extract and link data to an Ethereum Blockchain ecosystem. It proposes data access standardization for applications that use Ethereum and Smart Contracts as a tool, as DApps. This proposition allows the link with pre-existing Web datasets and Ethereum data.

The main contributions are:

- The EthExtras ontology that adds components to EthOn proposing some auxiliary classes that facilitate the understanding and relationship between Ethereum entities.
- a middleware developed in Python Flask that uses EthOn and EthExtras ontologies as proof of concept and exposes URI representing the entities of Ethereum as RDF. The URI is a SPARQL endpoint and can be used for visualization and query.

In Session II, we show the technologies and concepts necessary to an understanding of our work. In Session III we relation works. Session IV speak about the ontologies EthON, and we proposition EthExtras, given details about its modeling.
Session V, we show the middleware to consume Ethereum. Finally, in VI we discuss the semantic web and Blockchain in IoT App scenarios and finally the conclusion in Session VII.

II. BACKGROUND

For better understanding, our contribution will describe the components of the Blockchain and Semantic Web. These are the basis of this research, and in this session, we will cover the characteristics and technologies related to them and used in our work.

A. Blockchain

The Blockchain is a data structure in which a chain of blocks linked by the hash of the previous block and has a set of transactions submitted by accounts. Multiple nodes share this data structure in a Peer-to-Peer (P2P) architecture, and any modifications that occur are public to all nodes. All transactions use encryption and are signed and submitted to the Blockchain. It is not possible without private keys, the consensus mechanism, and the validation of the network nodes. The main entities of Blockchain are Blocks, Accounts and Transactions Figure 1.

![Fig. 1: Blockchain Blocks](image)

B. Ethereum

The Ethereum project is an Open Source Blockchain created to be the global decentralized computing platform for executing Smart Contracts programs. It uses Blockchain sync mechanisms to manage state changes, using its default cryptocurrency, the Ether (ETH).

Ethereum follows the initial revolutionary idea of the internet as a free and collaborative network in which information could also be free, but which has moved towards becoming a network controlled and regulated by a few. The control of the internet by a few leads to some problems that we currently experience, such as fake news, interference in elections in the country. In addition, data leakage problems and privacy and targeting of advertisements, videos, and information based on profiles collected without authorization and artificial intelligence are often used for personal insights. And it is in this expectation of being the new internet for a new pool of necessities of a local node setup.

A popular service in Ethereum and DApp ecosystem is Infura [6] is used to connect with Ethereum and InterPlanetary File System (IPFS) public and test networks. Infura provides a Blockchain development suite and API and avoids the necessity of a local node setup.

C. Semantic Web

The Semantic Web is currently the W3C recommendation for integrating WEB data sources. Proposed in 2001 by Tim Berners-Lee, Semantic Web mainly works on offering data models that can be an extension of the World Wide Web, allowing machines and humans to work together. For example, it links the meaning of words through ontologies.

Ontologies are data models representing a set of concepts within a domain and the relationships between them related to content published on the Internet.

The model of an ontology is an Ontology Web Language (OWL), and the currently recommended format for storage and query data is the Resource Description Framework (RDF); its format representing a triple graph consists of:

- subject, expressions that RDF uses to describe the resource
- predicate, a specific description of a resource can be an attribute or a relation between subject and object
- object, a named property, and its value.

To retrieve information from an RDF graph, we have SPARQL; this is the semantic query language recommended by W3C; it recovers and manipulates data in RDF format, returning an RDF Graph as a result.

Today we already observe interconnected Semantic Web communities such as the Linked Open Data (LOD), and this has over a thousand connected and open-access datasets.

III. RELATED WORK

The works with approach Semantic Web, Blockchain are not extensive covered and only recently found few relevant investigations. In this session, we present the main results found related to the focus of our research.

The article [7] goes deeper into the benefits of the Semantic Web and Blockchain, providing an overview of scenarios that benefit from the union of these approaches and analyzing its advantages and disadvantages.

The work [8] shows the results of a project aiming at developing the conceptual schema of Ethereum using UML.
The motivation for this was that most Ethereum literature is found from a technical or an economic perspective. Developers, researchers, and students need a simple model to understand the deep foundations of Ethereum, in contraposition of having that search details of this technology and its objects relationships in a large number of books, papers, and web references. This research differs from our work because it does not use Semantic Web for modeling but UML.

The research [9] describes the BLONDIE ontology, representing the semantics of structures related to relevant Blockchain projects: Bitcoin, Ethereum, and Hyperledger. This ontology focuses on integrating the standard data formats of the different Blockchain platforms, separate from our proposition that uses EthOn to deepen searches in Ethereum.

The article [10] presents the creation of a linked data index for educational platform Open Blockchain [11], implemented on Ethereum. This work aims to efficiently support efficient data access and Smart Contracts through the Semantic Web technology stack. The work implements a semantic index for the platform and exposes the data as Linked Data, indexed at block and transaction-level according to the BLONDIE ontology. Unlike our work that applies the EthOn ontology, this work focuses its results on the BLONDIE ontology.

The SANSA Semantic Web tool in the [12] poster is used in an Ethereum DApp, the CryptoKitties game, an online game based on Smart Contracts that allows players to trade characters securely, virtual pets. The Project uses a tool called Alethio [13] as an Ethereum analytics tool, providing transactions and logs in RDF modeled on the EthOn ontology, in addition to consolidating some results. Just like our work,

---

### Classes used in EthOn

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>PlatformNet</td>
<td>A network of a Platform, Ropsten, for instance, is an Ethereum develop Network</td>
<td>hasGenesisBlock, hasName, netName</td>
</tr>
<tr>
<td>Platform</td>
<td>A Blockchain Platform, in this research is Ethereum</td>
<td>hasCurrency, hasExternalReference, hasName, hasNet</td>
</tr>
<tr>
<td>Abi</td>
<td>The ABI of a Smart Contract</td>
<td></td>
</tr>
<tr>
<td>Blockchain</td>
<td>The Blockchain</td>
<td></td>
</tr>
<tr>
<td>Currency</td>
<td>The cryptocurrency of Platform in Ethereum is the Ether</td>
<td>hasCurrencyName, hasCurrencyPrefix</td>
</tr>
<tr>
<td>dbpedia:Blockchain</td>
<td>The dbpedia reference to Blockchain</td>
<td></td>
</tr>
<tr>
<td>dbpedia:Ethereum</td>
<td>The dbpedia reference to Ethereum</td>
<td></td>
</tr>
<tr>
<td>ethon:Account</td>
<td>An Ethereum Account</td>
<td>balance, hasOwner</td>
</tr>
<tr>
<td>ethon:Block</td>
<td>A Ethereum Block</td>
<td>miner, mixHash, receiptsRoot, stateRoot, transactionsRoot</td>
</tr>
<tr>
<td>ethon:ContractAccount</td>
<td>The External Ethereum Account used as a reference to a Smart Contract calls</td>
<td>balance, hasABI</td>
</tr>
<tr>
<td>ethon:Tx</td>
<td>A Ethereum transaction</td>
<td>inBlock, belonstToBlock</td>
</tr>
<tr>
<td>ethon:TxReceipt</td>
<td>The receipt of a ethereum transaction</td>
<td>blockHash, contract, cumulativeGasUsed, from, gasUsed, inBlock, logsBloom, type, to, transactionHash, transactionIndex</td>
</tr>
<tr>
<td>ethon:Uncle</td>
<td>The uncle block it happens when more with one child block is created from a parent block</td>
<td>miner</td>
</tr>
<tr>
<td>foaf:Person</td>
<td>A abstract representation of a user</td>
<td></td>
</tr>
</tbody>
</table>


**TABLE I: Classes**
this research focuses on the EthOn ontology its results in order to make Ethereum data more digestible for end users.

The work [14] investigates item description and discovery in a Blockchain to the Supply Chains using approaches of Semantic Web proposing a framework that provides an object discovery layer in resume object discovery, registration, and selection. The core of this proposition is to use Hyperledger Blockchain with an ontology of object and product discovery. Unlike our approach concentrate on Ethereum, it focuses its approach on Blockchain Hyperledger technology [15].

The research [16] presents futuristic scenarios and ideas and industrial scenarios, using Blockchain networks for data
feed and Semantic Web for data interconnection. The work presents the DeSCA prototype that records the interactions of the participants of a Supply Chain system in BlockChain and replicates this data in RDF format using the BLONDiE ontology as a basis. As well as our approach, the motivation of this research is to show that the Semantic Web principles can be used in a decentralized internet using Blockchain as Background to compose futuristic DApps scenarios.

An empirical analysis of store RDF triples in Blockchain compared with JSON storage are presented in [17]. The data in the proof of concept is produced by the IoT devices and the Ethereum Gas costs of these operations and its effectiveness in querying the database using SPARQL. Our proposal, different from this work, uses a web middleware that, through Web3.py [18], the Python library to speak with Ethereum BlockChain and on the fly generates the triple RDF of the Ethereum entity called via its URI link.

The work [1] presents a distributed database compatible with Blockchain named GraphChain that exposes the data with RDF graphs in a semantic model. Its graphs can be published as web-accessible linked data objects using HTTP and can also be queried by SPARQL. GraphChain, to define its semantics, uses its own OWL ontology to define structural entities. Some prototypes using Java, C, and JavaScript are used to demonstrate the dynamic of this Blockchain. This work uses the RDF graphs as part of the structure of an unprecedented Blockchain, being a different proposal from our middleware proposal that uses the current Blockchain Ethereum networks to extract the graphs exposing this data in a semantic model and using it with the available web ontologies and we EthOn extension the EthExtras ontology.

IV. ETHEREUM ONTOLOGY

In this session, we will cover the main points of our approach to using and extending the EthOn ontology with the EthExtras ontology. Figure 2 shows a vision of our proposition.

A. EthOn Ontology

EthOn [4] is an ontology to be a model semantical to represent Ethereum BlockChain and ecosystem. It has classes, object, datatype properties, and annotation. Supports multiple inheritances; for instance, Block is a child class of State Transition and Block.

The complete specification that covering the concepts hierarchy classes and object properties of EthOn can be found in [19].

EthOn taxonomy is in constant evolution, and its potential has not been explored yet, could new restrictions, classes, and subclasses are added to contribute or extend. To this task, we design a new ontology called EthExtras, extending or simplifying the EthOn ontology with extras Class and objects.

B. EthExtras Ontology

EthExtras is the name of we new ontology proposition to extends the EthOn ontology. Its format is in OWL and published in [20].

As with EthOn, the domain and scope of EthExtras covers the environment of Blockchain in Ethereum technologies, already proposing some links with external classes and datasets, being an example of an external link, and relationship with these. The Table I list the classes used in EthExtras.

For the design of the OWL format of EthExtras, we used Protégé [21], a project to develop ontologies from Stanford University. The Table II give details about the EthExtras class properties.

The representation of Blockchain Ethereum through these two ontologies makes it possible to generate a helpful model for consuming and providing data and usage in conjunction with other external data models exposing web endpoints. The Figure 2 illustrates the use of the EthExtras and EthOn classes and properties.

V. CONSUMING ETHEREUM DATA USING SEMANTIC WEB

We built as a proof of concept of we Semantic Web proposition a middleware that uses EthOn ontology and EthExtras capable of generating soft real-time RDF graphs of Ethereum data, and we based on the semantic adapter code from the IoT-Framework Engine project [22]. To Development, was used python with Flask, and the main libraries were Rdflib for generating the graphs and Web3.py for communicating with Ethereum. Our middleware code is found in [23].

The middleware is responsible for connecting to Ethereum and extracting its data through Web3.py. The extracted data is modeled in RDF graphs by Rdflib and using REST API in Flask to create endpoints SPARQL in XML, JSON-LD, N3, and Turtle formats.

A version to proof of concept using Infura to communicate with Ethereum Networks was deployed in the Heroku platform in address [24].

The REST routes available in Middleware are in Table III. The route returns an RDF capable of being used for SPARQL queries linked to other available datasets.

VI. WHAT CAN THE NEW IOT APPS BENEFIT FROM SEMANTIC WEB?

Scenarios of Blockchain applications different from the Fintech domain has recent spotlights. Due to the current expected demand for IoT, we can place the management of the expected high volume of data coming from these devices in SmartCities (SC) and I4.0 Apps as possible challenges in the coming years. Blockchain’s security features and the power and simplicity of data binding make the duo called Web3, BlockChain, and Semantic Web a potential solution.

Using some Semantic Web concepts can solve issues related to the standardization of access to IoT devices and data produced by them that we do not have today due to the heterogeneity of interfaces and API. This data is consumed, identified, recorded, filtered, and discovered by a standard, and this makes IoT information integrate more collaboratively with other datasets, creating a WoT (Web of Things) [25]. WOT allows standardized services and solutions elaborated and combined with other RDF datasets as the endpoints RDF of
this work, the Ethereum Semantic Web middleware, including filters and data aggregators. A data model in RDF using Semantic Sensor Network (SSN) ontology [26] produced from IoT devices can be linked to endpoints that provide data modeled in EthOn and Eth Extras, enabling SPARQL queries to generate new soft-realtime insights. SSN ontology describes actuators and sensors, covering their observations, procedures, characteristics, and observed properties. Some propositions of data extract and model in SSN are found in the literature. One of them is OpenIoT using XGSN [27], which implements a virtual sensor layer to give data visibility using SSN. The virtual sensor annotates the data and its storage in a graph database, the Open Virtuoos, allowing SPARQL queries. 

VII. CONCLUSION

This research contributes to some problems that are still open in the literature, mainly when considering the few works that combine Blockchain and Semantic Web. We proposition, the EthExtras ontology extension, complete the EthOn and create a model that can be use of base to consume data of Ethereum simply. We middleware using this combination allows consumes Ethereum data in soft-realtime and is motivated to propose integrating data from a Blockchain as Ethereum with external world applications in a standardized way. This middleware and its endpoints RDF representing the objects of Ethereum can be used to use and link with external datasets, giving power and flexibility to developers create queries in a option of traditional API as Etherscan [28]. In future works, we are adding the NFT, Tokens, ENS, IPFS and some famous Oracle Services to EthExtras and implement this in the middleware as proof of concept.

ACKNOWLEDGMENT

Thanks to FAPEMIG, CNPq, CAPES and UFOP (Federal University of Ouro Preto).

REFERENCES


