# Profiting from Kitties on Ethereum: Leveraging Blockchain RDF Data with SANSA

Damien Graux[1], Gezim Sejdiu[2], Hajira Jabeen[2], Jens Lehmann[1,2], Danning Sui[3], Dominik Muhs[3] and Johannes Pfeffer[3]

[1] Fraunhofer IAIS, Germany
{damien.graux,jens.lehmann}@iais.fraunhofer.de
[2] Smart Data Analytics, University of Bonn, Germany
{sejdiu,jabeen,jens.lehmann}@cs.uni-bonn.de
[3] Alethio
dmuhs@protonmail.ch
{danning.sui,johannes.pfeffer}@consensys.net

**Abstract.** In this poster, we will present attendees how the recent state-of-the-art Semantic Web tool SANSA could be used to tackle blockchain specific challenges. In particular, the poster will focus on the use case of CryptoKitties: a popular Ethereum-based online game where users are able to trade virtual kitty pets in a secure way.

## 1 Introduction

During the recent years, the interest of the research community for blockchain technologies has raised up since the Bitcoin white paper published by Nakamoto [4]. Indeed, several applications either practical or theoretical have been made, for instance with crypto-currencies or with secured exchanges of data. More specifically, several blockchain systems have been designed, each one having their own set of properties, for example Ethereum [6] –further described in Section2– which provides access to so-called *smart-contracts*. Recently, Mukhopadhyay *et al.* surveyed major crypto-currencies in [3]. In parallel to ongoing blockchain development, the research community has focused on the possibilities offered by the Semantic Web such as designing efficient data management systems, dealing with large and distributed knowledge RDF graphs. As a consequence, recently, some studies have started to connect the Semantic Web world with the blockchain one, see for example the study of English *et al.* presented in [1].

In this poster, we will offer attendees the possibility of discovering the blockchain world, with the Semantic Web framework SANSA applied in the CryptoKitties use case.

## 2 Preliminaries

*Ethereum* is a **decentralized platform, which allows smart contracts** (e.g. codes written in Solidity language) to operate on it. Anyone can write their own

smart contracts and deploy them on the Ethereum network, to create applications. Moreover, anyone can also interact with them; this is done by calling contract functions, which in turn trigger events logged to a public record. Costly interactions requiring computational resources are provided with a fee given to the Ethereum block miners. Via smart contracts, applications are supposed to run exactly as defined without any possibility of downtime, censorship, fraud or any third-party interferences [6]. Despite the ecosystem's security measurements and audits, there are some cases where the smart contracts are attacked, with e.g., the DAO attack.[1] With hundred-thousands of activities happening every day on the Ethereum network, the need for tracing, as well as monitoring has risen over time. Although all the blockchain data is stored publicly and transparently, it is hard for non-technical participants to access or understand it.

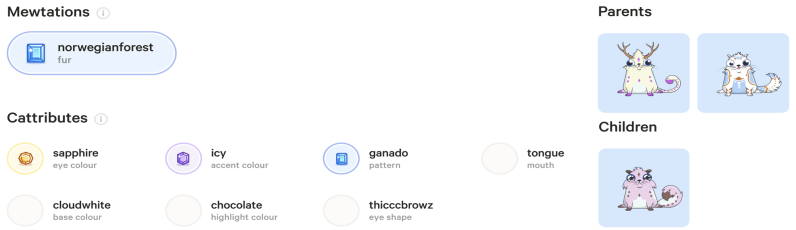
Alethio[2] is a spin-off working with Ethereum technologies. It is an advanced analytics platform making Ethereum more accessible and digestible for everyone. It could be seen as a "blockchain archeology" of the Ethereum main network, providing not only all the original transaction data and log messages, but also analytic results with specific metrics. Their extensive data set (currently encompassing 36 billion rows of records) contains large-scale blockchain transaction data modelled in RDF according to the structure of the Ethereum ontology. *EthOn* (The Ethereum Ontology) [5] formalizes all the concepts and terms of the Ethereum ecosystem in OWL. EthOn describes all the Ethereum objects as classes in the ontology. Also, it depicts all the interactions and attributes of objects, as properties from the ontology perspective of view.

*The SANSA stack* [2] is an open source framework[3] that allows RDF processing at scale. It provides a set of libraries for distributed reading, executing SPARQL queries, performing inference as well as analytics over large-scale knowledge graphs.

## 3 SemanticWeb & Ethereum: the CryptoKitties Use Case

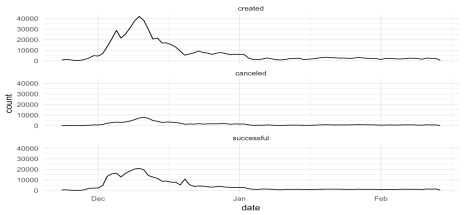
The Ethereum ecosystem generates a large amount of data, including but not limited to protocol-level data (e.g. average block time, gas prices), as well as application-level data (e.g. account interactions, smart contract deployments). To efficiently handle this volume of data, Alethio has investigated different tools and frameworks with one focus: the infrastructure should be resilient, load-bearing, and most importantly, scalable. And so, for that reason to overcome the variety of the different data sources, Alethio introduces semantification of Ethereum network and uses SANSA as an underlying engine for large scale distributed RDF based querying, reasoning, and machine learning on top of these RDF datasets. To show the joint effort between SANSA and Alethio, we describe a use case on how SANSA can be used to analyze Ethereum at new scales.

[1] `https://www.coindesk.com/understanding-dao-hack-journalists/`
[2] `https://aleth.io/` (last accessed June 11th 2018)
[3] `https://github.com/SANSA-Stack`

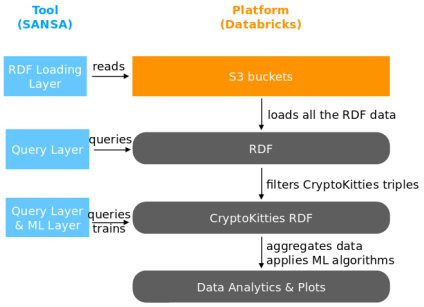



(a) The unique attributes of a Kitty.

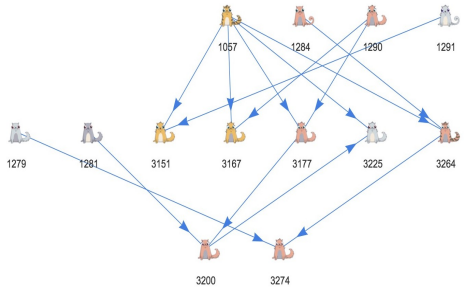
(b) An instance of a Kitty.




(c) History of three types of auction events.




(d) The process pipeline.




(e) An illustration of a small family tree.

*CryptoKitties*[4] is one of the first games to be built on blockchain technology. In particular, CryptoKitties initiated and released the first generation virtual kitties, with delicately designed icons and genes sequences. All the kitties are virtual with some biological feature settings. Shown in Figure 1b is a kitty with its specific biological attributes displayed in Figure 1a. The attributes are stored in a sequence, succeeded from its parents' gene sequences, with possibility of *mewtations*. An owner can sell, breed or gift it to other user. When users sell or breed it, they will send transactions to the CryptoKitties smart contracts, which will complete the execution of either transferring ownership between users, or generating a new kitty. Based on that, game users can trade or breed kitties like traditional collectibles, while having the guarantee that the blockchain will track ownership securely. Moreover, one can breed two kitties to create a brand-new, genetically unique offspring.

[4] `https://www.cryptokitties.co/` (last accessed June 11$^{th}$ 2018)

*Data Challenges.* Alethio has been exploring efficient means of processing large RDF data sets. SANSA empowers Alethio to read and query the data at scale as described in Figure 1d. Indeed, once the complete RDF data set is loaded, SANSA filters it to retain only the CryptoKitties triples –transactions, contract messages and log information– before performing more specific analyses.

Practically, the challenges tackled with SANSA can be divided into two groups: game performance and customer behaviors. The first one focuses on time series metrics: throughput time, the event volume, number of active users and amount of spent Ether, which can jointly estimate the trend of popularity for the game. In Figure 1c, the history of CryptoKitties auctions events shows clearly that there was a peak of traffic in December after the game was launched for around one month. By this time series, we can estimate the popularity of the game throughout history. The second one requires machine learning algorithms to detect correlations between indicators (e.g. to determine whether richer owners have the tendency to collect special/rare kitties which are more expensive). and topology from a network view. In Figure 1e, we present a small subset of the kitty family tree, where incest happened during the reproduction: kitty 1057 is the secondary-degree relative (grandparent) of kitty 3200, while later it bred with kitty 3200 and gave birth to kitty 3225.

## 4 Conclusion

During this poster session, attendees will be able to discover the blockchain world and the specificities of the Ethereum through the CryptoKitties use case. In addition, we will present how the SANSA stack –and the Semantic Web standards– has been used to tackle some specific problems dealing with blockchain analyses.

## References


1. English, M., Auer, S., Domingue, J.: Block chain technologies & the semantic web: A framework for symbiotic development. In: Computer Science Conference for University of Bonn Students, J. Lehmann, H. Thakkar, L. Halilaj, and R. Asmat, Eds. pp. 47–61 (2016)
2. Lehmann, J., Sejdiu, G., Bühmann, L., Westphal, P., Stadler, C., Ermilov, I., Bin, S., Chakraborty, N., Saleem, M., Ngonga, A.C.N., Jabeen, H.: Distributed semantic analytics using the SANSA stack. In: Proceedings of 16th International Semantic Web Conference - Resources Track (ISWC'2017) (2017), `http://svn.aksw.org/papers/2017/ISWC_SANSA_SoftwareFramework/public.pdf`
3. Mukhopadhyay, U., Skjellum, A., Hambolu, O., Oakley, J., Yu, L., Brooks, R.: A brief survey of cryptocurrency systems. In: Privacy, Security and Trust (PST), 2016 14th Annual Conference on. pp. 745–752. IEEE (2016)
4. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system (2008)
5. Pfeffer, J., Beregszazi, A., Detrio, C., Junge, H., Chow, J., Oancea, M., Pietrzak, M., Khatchadourian, S., Bertolo, S.: Ethon - An Ethereum ontology (2016)
6. Wood, G.: Ethereum: A secure decentralised generalised transaction ledger. Ethereum project yellow paper **151**, 1–32 (2014)