



HAL
open science

Visualization of Blockchain Data: A Systematic Review

Natkamon Tovanich, Nicolas Heulot, Jean-Daniel Fekete, Petra Isenberg

► **To cite this version:**

Natkamon Tovanich, Nicolas Heulot, Jean-Daniel Fekete, Petra Isenberg. Visualization of Blockchain Data: A Systematic Review. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 27 (7), pp.3135 - 3152. 10.1109/TVCG.2019.2963018 . hal-02426339

HAL Id: hal-02426339

<https://hal.science/hal-02426339>

Submitted on 2 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Visualization of Blockchain Data: A Systematic Review

Natkamon Tovanich, Nicolas Heulot, Jean-Daniel Fekete, and Petra Isenberg

Abstract—We present a systematic review of visual analytics tools used for the analysis of blockchains-related data. The blockchain concept has recently received considerable attention and spurred applications in a variety of domains. We systematically and quantitatively assessed 76 analytics tools that have been proposed in research as well as online by professionals and blockchain enthusiasts. Our classification of these tools distinguishes (1) target blockchains, (2) blockchain data, (3) target audiences, (4) task domains, and (5) visualization types. Furthermore, we look at which aspects of blockchain data have already been explored and point out areas that deserve more investigation in the future.

Index Terms—Blockchain, Bitcoin, Ethereum, Information Visualization, Visual Analytics, State-of-the-Art Survey.

1 INTRODUCTION

BLOCKCHAIN technology has become popular in the last 10 years after the Bitcoin cryptocurrency was introduced by Satoshi Nakamoto [33]. Since then, the blockchain concept has been used to develop decentralized systems to store and maintain the integrity of time-stamped transaction data across peer-to-peer networks. Bitcoin [33] and Ethereum [48] are popular examples of blockchain use for digital currencies, smart contracts, and decentralized applications. As the technology has revolutionized transactions and exchanges, it found applications in various industries, including energy production, mobility, and logistics [13].

Despite its active use, blockchain is a new technology and its use in practice is still evolving and poorly understood. Broadly, a blockchain is a decentralized system governed by autonomous mechanisms that ensures the data stored among peers is correct and that prevents possible fraudulent activities. Blockchain data is stored and maintained among peers in the network by the consensus of the network majority. As a result, activities on the blockchain are driven by the interaction of its users, in contrast to centralized controlled systems where users are regulated by a central server. In order to adopt blockchain technology in a wider set of domains, we will need to explore and analyze transaction data to better understand emergent user behavior and mechanisms in blockchain systems. As such, visualization and visual analytics (referred to as “VA”) tools can support human analysts in deriving first hypotheses and models of blockchain use.

We contribute a systematic overview of past solutions proposed in the VA community, but also take a close look at how blockchain visualizations are used by practitioners and researchers in economics, computer science, and even public audiences. For our survey, we collected visualizations

from both online websites and academic articles and refer to those articles and websites as “sources” throughout the article. We systematically assessed the motivations and characteristics of each source and defined a classification scheme to group visualizations based on five aspects: target blockchains, blockchain data, task domains, target users, and visualization types. Finally, we provide a summary of blockchain visualizations that have been proposed as well as perspectives on aspects that are rarely explored and would benefit from further exploration.

2 BACKGROUND ON BLOCKCHAIN TECHNOLOGY

The blockchain concept was introduced in the early 1990s as a theoretical system to store a time-stamped digital document that cannot be modified [5] [19]. The articles proposed data structures and algorithms to store non-modifiable data and maintain trust in decentralized systems without central control. In 2008, a person or group of people under the pseudonym Satoshi Nakamoto published the seminal article: “Bitcoin: A Peer-to-Peer Electronic Cash System” [33] that proposed a way to prevent double-spending of digital currency transactions without requiring a trusted third-party [43]. Since then, the blockchain has been implemented in many different domains, including cryptocurrency, smart contracts, and intellectual property management [42].

Among the existing blockchains in the public domain, *Bitcoin* and *Ethereum* are the most well-known public blockchain platforms. Both of them use different instantiations of the blockchain concept. *Bitcoin* has currently the most widely used cryptocurrency blockchain with the highest total market capitalization (~\$140 billion), as of November, 2019 [11]. Bitcoin blockchain data alone contains over 470 million transactions (over 250 GB of raw data) and is constantly growing [O10]. Its currency is the Bitcoin (BTC), valued ~\$7.66k with important fluctuations.

On the other hand, *Ethereum* focuses on the implementation of smart contracts [48]. A smart contract is a piece of computer code that is guaranteed to run in the same way on all peers. Ethereum has a currency unit called Ether

- N. Tovanich and N. Heulot are with IRT SystemX, Paris-Saclay, France. E-mail: {natkamon.tovanich, nicolas.heulot}@irt-systemx.fr
- N. Tovanich, J.-D. Fekete, and P. Isenberg are with Inria, France. E-mail: {natkamon.tovanich, jean-daniel.fekete, petra.isenberg}@inria.fr
- N. Tovanich is also with Université Paris-Saclay, France.

Manuscript received September 24, 2019; revised December 20, 2019.

which is used to pay for machines executing the code. As of November, 2019, there are more than 590 million transactions amounted to over 200 GB data size in the Ethereum ledger [O21]. It has been increasingly adopted by companies that formed the Ethereum Enterprise Alliance (EEA) in February 2017. Among the founding members were big companies such as, Microsoft, JP Morgan, Accenture, and Intel [15].

2.1 How does the Bitcoin blockchain work?

Since the Bitcoin blockchain is currently the most well-known (popular) and widely-used (active) blockchain in the public domain, we will start by explaining the mechanism behind the Bitcoin system as many of its concepts also similarly apply to other types of blockchains. In this section, we provide a simplified description for general audiences and refer to the book “Mastering Bitcoin: Programming the Open Blockchain” [2] for more technical detail about the Bitcoin blockchain.

The blockchain is a public *ledger* that records a list of *transactions* in the distributed network. For Bitcoin, most of these transactions are of cryptocurrency value transfers, but more complex transactions are possible (e.g., simple smart contracts, multi-signatures, etc.). Just like regular transactions, blockchain transactions need to be validated: the sending and receiving addresses need to be valid, the sending account needs to contain enough value to be transferred (technically called, *unspent transaction outputs (UTXOs)*) and senders need to have the right to spend the value. In traditional banks, these validations are performed by the bank itself which has to be trusted to avoid double spending or stealing.

With blockchain technology, the ledger is public and distributed. The validation is performed through a consensus reached by a pool of people called *miners*. Anyone can decide to become a miner; it only requires very powerful machines and a good network connection. The validation is done by block, so when transactions are issued, they are buffered and pending in the *mempool* (i.e. transactions waiting to be confirmed/included in a new block). Those transactions are collected and verified by the miners. Some transactions can be rejected, and for the valid ones, they are included in a new block. Then, the new block is added to a chain of blocks (the updated ledger) that cannot be changed, hence the term *blockchain*. The blocks are considered valid if they are accepted by the majority of nodes.

The *validation* miners perform involves running computationally expensive methods to verify the validity of the transaction parties and amounts transferred. These methods are based on public-key cryptography. There are also technical differences between a traditional transaction and blockchain transactions. In the Bitcoin blockchain, the transactions are done from *addresses* and not from accounts. An address can be created at any time for free, and is represented as a long string with cryptographic properties to be able to validate its owner. The owner of the address cannot be inferred from the address itself so transactions are almost anonymous, though they can be tracked, hence they are referred to as *pseudonymous*. One *transaction* can involve several input addresses and send value to multiple output

addresses. All the value of input addresses is sent to the output addresses, so the input addresses end-up empty at the end of the transaction, except that the transaction can send change back to the owner, on any of his or her addresses. If the amount transferred is less than the total amount held by the input addresses, the change is left as a *transaction fee* for the *miner* who will validate the block.

Because the blockchain is a decentralized system, a *consensus protocol* is needed to decide which transaction is valid and should be added to the ledger. Many blockchain-based systems, including *Bitcoin* and *Ethereum*, adopt *Proof-of-Work (PoW)* protocols. *Miners* are in charge of maintaining the blockchain ledger and propose a new block to the network. Since this operation is expensive, they need a reward. The miner who successfully proposed a new block can reclaim a *coinbase transaction*: it includes newly generated value and transaction fees from every transaction in a block. Yet, the validation is performed as a competition: multiple miners perform the computations to validate the block which is, in short, trying to find a number that gives a hash with a specific form. The first miner who solves the puzzle can propose a new block to the network and claim the reward if the majority of miners agree to include it in their ledgers. Therefore, miners get their reward but not regularly in proportion to their computational power. The difficulty of mining is decided by the total computation power in the blockchain network (so-called *hash rate*) which often adapts to reach the desired rate of adding a new block every 10 minutes. The mining process ensures that the data in the Bitcoin blockchain is consistent and prevents attacks from malicious users.

Like in Bitcoin, other blockchains share the general idea of a growing, verifiable but immutable list of records stored in blocks that are linked to one another. Cryptographic measures are used to encode links between blocks. In contrast to the bitcoin blockchain, however, other blockchains may implement a different protocol to store transactions and regulate the consensus in their decentralized networks. Examples of alternative consensus protocols are *Proof-of-stake (PoS)*, *Practical Byzantine Fault Tolerance (PBFT)*, *Ripple* and *Tendermint*. We refer readers to Zheng et al.’s survey [50] that describes different blockchain protocols in greater detail.

2.2 Elements of blockchain data

Here, we generalize blockchain data elements and illustrate those data elements by giving concrete examples in Bitcoin. We will refer to these generalized types of blockchain data in our classification scheme.

A **transaction** is the most granular level of blockchain data. It records a transfer of value between **addresses**. In Bitcoin and other cryptocurrencies, a transaction record contains pseudonymous input and output address(es) with the value to transfer or received associated with each address.

Transaction records are stored in a data structure called a **block**. Blocks hold and group a certain number of transactions. Multiple blocks are connected in a linked list called a ledger. **Nodes** are electronic devices that maintain and distribute a copy of the ledger in the blockchain network so that the data remains synchronized. **Miners** are special nodes in the peer-to-peer network that participate in verifying transactions and adding new blocks to the ledger, with the possibility to receive a reward.

An **entity** represents a real blockchain user or organization behind a transaction. When input and output addresses are pseudonymous, entities cannot be directly inferred from the blockchain. In addition, an address is meant to be used only once as a conventional practice in the cryptocurrency blockchain community for privacy and security purposes.

Yet, we can trace the activities of addresses on the blockchain without knowing the real-world identity of entities behind transactions. Research has shown that simple heuristics can be used to group pseudonymous input addresses into entities [1] [A13] [A15] [A16]. By simply grouping addresses, entities remain pseudonymous. Yet, external data sources exist that provide a list of addresses that belong to well-known entities such as WalletExplorer.com [22], Bitcoin Forum [32], and Blockchain.info [O10].

2.3 Types of blockchain

Blockchains can be categorized into three types: public blockchains, consortium blockchains, and private blockchains [10] [50].

- **Public blockchains** are open blockchains in which any participant can read, write, and submit transactions to the ledger. Any participant can join the consensus process to determine whether to add blocks and transactions to the ledger. Public blockchains are suitable for applications that are open for everyone and need fully decentralized systems. Bitcoin is a well-known example of a public blockchain.
- **Consortium blockchains** are semi-private blockchains that restrict the consensus process to the selected group of participants that are trusted by the system. This reduces the time to verify transactions and blocks but also makes the systems partially centralized to selected nodes. Permission to operate a node on a consortium blockchain is granted by the overseeing group of organizations.
- **Private blockchains** are fully controlled by an organization that determines the consensus of the blockchain ledger. The private blockchain owner has an authority to allow or restrict the read permission to participants. Private blockchains are centralized systems, similar to database systems, and usually suitable for applications which require high trust and privacy.

Some of the criteria described at the beginning of Sect. 2 differ for consortium and private blockchains, where, for example, consensus is determined by selected nodes that can be trusted, and therefore past records could theoretically be tampered with. In the remainder of the article, we focus on public blockchains as our systematic review did not uncover analyses or tools dedicated to private or consortium blockchains. In this article, we also do not consider solutions like sidechains that allow interoperability between blockchains or that speed up transaction validation (e.g., *Lightning Network*) because they introduce specificities that we consider out of the general scope of our survey.

3 RELATED WORK

A large number of research disciplines are interested in the blockchain, including algorithms, software formal verification, database systems, computer security, system architecture, data security, and economics. Here, we summarize

prior work on the state-of-the-art in blockchain research that we reviewed to inform our own classification scheme of blockchain VA tools.

Reviews on blockchain technology: Most of the existing literature reviews on blockchain research focus on a technical perspective. For example, Zheng et al. [50] presented a comprehensive review of the current advancement of blockchain technology. The authors describe common blockchain characteristics: *decentralization, persistency, anonymity, and auditability*; and then compare the differences among consensus algorithms. The article also lists real-world applications that can benefit from blockchain architectures. Bonneau et al. [8] provide an analysis of algorithms and protocols used specifically in Bitcoin and cryptocurrencies. The article highlights stability and security limitations in the current cryptocurrency blockchains and proposes future challenges. Yli-Huomo et al. [49] collected 48 research articles on blockchain technology which the authors summarized into 7 categories based on technical challenges and limitations. One of the technical challenges discussed in the article is *usability from a user's perspective*. The authors emphasized the necessity of analytical tools to improve the ability of users to analyze and detect patterns in the blockchain network. None of these past reviews refer to visualization, which is the focus of our review.

Reviews on blockchain analytics: Balaskas and Franqueira [3] examined analytic tools for the Bitcoin blockchain that are available on the internet and proposed a taxonomy based on analysis themes: *analysis of entity relationships, metadata, money flows, user behavior, transaction fee, and market / wallets*. The found tools are mainly able to track and monitor cryptocurrency values, and therefore are useful for detecting fraudulent transactions. Another article by Bartoletti et al. [4] surveyed Bitcoin and cryptocurrency analysis tools found in academic articles and websites. The tools in their survey were classified based on analysis goals: *anonymity, market analytics, cyber-crime, metadata and transaction fees*. For each analysis goal, the authors further specified the kind of blockchain-related data used in the tools, such as *transaction graphs, address tags, IP addresses, mining pools, exchange rates, and lists of DDoS attacks*; and listed all sources that they retrieved. Based on the survey, the authors developed a general framework for blockchain analytics and showed use cases of analyzing transaction fees and Bitcoin metadata. Both articles collected tools dedicated only to the Bitcoin blockchain and classified them based on analytics tasks rather than visualization of blockchain data; which is the goal of our present work.

Reviews on blockchain visualization: We found only one literature review on blockchain visualization. Sundara et al. [41] reviewed 8 Bitcoin tools available on the internet and provided a short description on *visual representations and implementations*. Most tools in their survey performed real-time monitoring for Bitcoin transactions. Nonetheless, the authors neither performed an exhaustive search nor proposed a method to classify the tools they found. In our previous work, we collected 46 online Bitcoin visualization tools using a systematic review approach and classified them based on *analysis tasks and visual representations* [46]. In this article, we extend our data collection to include other kinds of blockchains (e.g., Ethereum) from research articles and online

sources, and provide a more complete classification scheme that additionally considers blockchain data visualized, target audiences, and task domains.

4 DATA COLLECTION

We identified visualization tools for blockchain data from both academic articles and online sources. In this section, we describe the data collection procedure and the criteria we used to include literature related to blockchain visualization.

4.1 Identifying search idioms

The first step in our analysis was to determine the right search terms for identifying the most relevant articles. We chose four starting search terms: “*blockchain*”, “*bitcoin*”, “*cryptocurrency*”, and “*ethereum*”; as we expected them to result in good coverage of blockchain-related articles. To narrow down the literature to tools related to visualization techniques for blockchains, we used the character sequence “*visual*” to cover keywords such as “*visualization*”, “*visual analytics*”, “*visualizing*”, etc. In addition, we used the character sequences “*data analy*” and “*graph*” to return articles that did not specifically use any “*visual*”-related key terms. We decided to be relatively broad in our search terms in order to optimize for recall rather than precision of the search result.

4.2 Searching academic articles

We selected 6 scientific databases to retrieve articles: (1) IEEE Xplore, (2) ACM Digital Library, (3) ScienceDirect, (4) DBLP, (5) Springer Link, and (6) Google Scholar. We used search engines available in those databases and applied the above search idioms to retrieve relevant articles. The initial search was performed in April 2019, but we included newly published relevant articles that appeared later up to the time of submission.

After these individual searches, we combined resulting articles from these six databases and removed duplicates. We next screened the returned results by reading the title of returned articles one by one and selected articles that seemed to potentially include blockchain visualizations beyond simple charts. If the title did not clearly describe the relevance of an article, we additionally read the abstract before deciding on inclusion in our survey. Inclusion criteria were: (a) the article is related to VA on any blockchain, and (b) the article includes a data analysis on any blockchain technology and uses visualization to communicate results.

4.3 Searching online web-based visualization

To collect blockchain visualization tools that are available on the internet, we typed every search idioms from the combination of (“*blockchain*” OR “*bitcoin*” OR “*cryptocurrency*” OR “*ethereum*”) AND (“*analysis*” OR “*analytics*” OR “*visualization*” OR “*visual analytics*” OR “*graph*” OR “*chart*”) on Google Search and retrieved the first 100 results. We followed the link to each web page one by one, looked at it, and checked whether the web page contained blockchain visualizations. In the case of web pages that contained links to other visualization tools, we followed each link in the web page and added the link to our list. To be selected, the

web page had to contain interactive graphics showing raw or aggregated data that is stored on a blockchain. We excluded web pages that showed only market data on cryptocurrency exchanges (e.g., the current \$ value of a Bitcoin).

4.4 Data collection result

Using our systematic review approach, we collected the following number of sources:

Literature Survey Sources			
Visualization Articles	Analysis Articles	Online	Total
14	17	45	76

Most of the tools we found came from online sources (59%) while visualization articles represented only 19% of all sources. The remaining 22% of sources are data analysis articles that provide empirical analyses of the blockchain networks and communicate the findings using static images. We decided to include data analysis articles to understand possible questions that researchers are interested in and common visualization types they used to convey their results.

Throughout this article, we include references to our sources using the following naming scheme: visualization articles ([V#]), data analysis articles ([A#]), and online sources ([O#]). Full references to these sources are available in Table 1.

5 CLASSIFICATION SCHEME AND METHODOLOGY

In defining our classification scheme, we considered many visualization-related categories such as data, task, types of visualizations, or end-users. After several rounds of open coding with an evolving code-set, we converged on five main aspects for delineating blockchain visualization sources: (1) target blockchains, (2) blockchain data, (3) target audiences, (4) task domains, and (5) visualization types. Fig. 1 gives an overview of our classification scheme. In this section, we present the classification scheme we applied to each source as well as summary statistics that show how many sources included visualizations within the given category. As a result, the total counts and percentages we report do not necessarily correspond to 76 / 100%—the number of total sources we collected—as sources may have included multiple types of visualizations in the classification scheme.

5.1 Target Blockchains

Blockchain visualization sources in our survey were targeted at the following blockchains:

Number of sources for different target blockchains		
Bitcoin	Ethereum	Others
60	19	10

With 79%, data from the **Bitcoin** blockchain was the most common to be visually represented. This is not surprising as the Bitcoin blockchain is the oldest running cryptocurrency blockchain and still widely used nowadays. The **Ethereum** blockchain was the second-most common visually represented blockchain (25%). Only 13% of our sources were dedicated to **other** kinds of blockchains—all cryptocurrency blockchains—such as those of Namecoin [34], Litecoin [28], Dogecoin [44], and Dash [12].

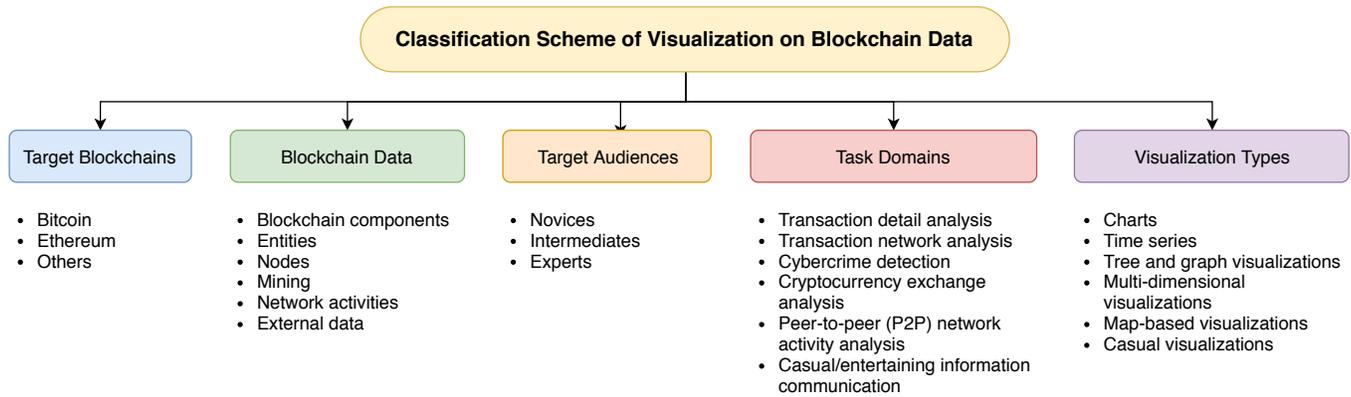


Fig. 1. The diagram shows the classification scheme of visualizations on blockchain data we developed after several rounds of assessing and characterizing sources we collected.

We did not find visualization tools on consortium and private blockchains, such as Hyperledger [45] and Dragonchain [14], likely because those kinds of blockchains are developed inside private organizations and data is not publicly available to analyze and visualize. However, many of the visualization techniques we surveyed can apply with modification to private and consortium blockchains as the underlying technological concepts are often similar.

5.2 Blockchain Data

We categorized seven different types of blockchain data that were represented:

Number of sources with different blockchain data					
Blockchain Components	Entities	Nodes	Mining	Network Activities	External Data
56	16	11	15	27	18

The most common category of data visualized was **blockchain components** (74%). Blockchain components are fundamental data types stored in public ledgers, including transactions, addresses, and blocks. **Entities** data require pre-computation to identify blockchain users from anonymous addresses and was used in 21% of all sources.

Nodes play an important role in the blockchain network. However, node information was only presented in 14% of our sources. Nodes ensure consensus through mining, to verify transactions and store them in the public ledger. Statistics of **mining** activity was presented in 20% of our sources. The data can be directly calculated from the blockchain such as, for Bitcoin, the average miner's speed to solve the proof-of-work problem (*hash rate*), mining difficulties over time, and the amount of reward to the successful miners.

Network activities information were the second most common data source (36%)—usually displayed using aggregated statistics describing the whole network. Network activity data usually included time-based data on the number of unique addresses used, the total number of transactions recorded in a given time period, the number of transactions waiting to be confirmed (*mempool*), and the number of unspent transaction outputs (*UTXOs*).

External data appeared in 24% of all sources to convey meaningful contexts such as cryptocurrency exchange rates,

online news, socio-economic data (e.g., percentage of internet users, gross domestic product (GDP) per capita, or the human development index (HDI)), social media information, or Google Trends data. The most common external data source was data on cryptocurrency exchanges, in particular, to describe conversion rates of a cryptocurrency value to a government-backed currency, such as the exchange rate of Bitcoin to US Dollar.

5.3 Target Audience's Levels of Analysis

We categorized three types of target audiences for blockchain visualization tools and visualizations communicated in data analysis articles based on levels of analysis that audiences demand to the tool:

Number of sources for different target audiences		
Novices	Intermediates	Experts
16	30	30

The majority of existing blockchain visualizations were targeted at users required for analyses at intermediate or expert levels. Sources for **intermediate users** (39%) aimed for active blockchain users such as miners, cryptocurrency traders, or enthusiasts who may be interested in monitoring blockchain activities or look at individual transactions or blocks. Among those sources, only two visualization articles targeted this audience.

Sources for blockchain **experts** (39%), such as economists or fraud investigators, targeted users who may perform in-depth analyses of activities on the blockchain. Sources for experts commonly allowed the flexible investigation of data at multiple levels of scale (from individual transactions to network activities) and using multiple (potentially pre-computed) dimensions of data. All data analysis articles (17 sources) presented analyses of blockchain networks using specific calculated measures, such as the growth of blockchain adoption, the degree of connectivity between entities, or the centrality of entities in the transaction network. On the other hand, most visualization articles (12 out of 14 sources) proposed interactive systems that allow data analysis experts to engage in exploratory analysis; while the other two targeted intermediate users. We found only one

online source for expert users that allowed to track Bitcoin value movement.

Of the sources we found 21% targeted **novice** audiences—all were online sources aimed at casually informing curious visitors about the Bitcoin blockchain.

5.4 Task Domains

We categorized our blockchain visualization tools and data analysis articles into six focus task domains: (1) transaction detail analysis, (2) transaction network analysis, (3) cybercrime detection, (4) cryptocurrency exchange analysis, (5) peer-to-peer (P2P) network activity analysis, and (6) casual/entertaining information communication. These task domains are not mutually exclusive but helped us to detect goals for the development, analysis, and exposure of existing tools.

Number of sources with different task domains					
Detail	Network	Cybercrime	Exchanges	P2P	Casual
24	24	12	10	27	12

P2P network analysis (36%) was the most common task domain we found in blockchain visualization tools, followed by transaction detail analysis (32%), and transaction network analysis (32%). We observed that cybercrime detection (16%) is a task domain focused on investigating fraudulent activities and cyberattack events in the blockchain, but still largely missed in the existing sources. Moreover, we found visualizations for cryptocurrency exchange analysis (16%), and for casual/entertaining information communication (13%).

Transaction Detail Analysis: Transaction detail analysis tools often expose basic statistics on the level of individual transactions, of blocks, and sometimes related to individual blockchain users (entities) such as individual people, exchange platforms, dark marketplaces, gambling services or companies.

Since the most common application context for blockchains currently is cryptocurrencies, it is not surprising that all 24 sources in this task domain had to do with the communication of basic information on financial transactions—17 online sources and 7 visualization articles.

Transaction Network Analysis: A blockchain transaction network is a bipartite graph connecting addresses through transactions. Half of the sources in this task domain were from data analysis articles (12 out of 24 sources). These articles analyzed transaction networks and described the structures and dynamics of blockchain transaction networks. These articles included visualizations that represented measures calculated from the transaction network, such as the number of addresses (node), and the distribution of transactions received (in-degree) and sent (out-degree) over time.

Moreover, there were 7 visualization articles and 5 online sources focused on visualizing large blockchain transaction networks. These also allowed for interactive exploration of transaction networks based on specific events or group of entities. These tools generally did not report statistical network measures but focused on representing the variety or temporal dynamics of address connections.

Cybercrime Detection: Cybercrime is a serious threat to the use of blockchains. This task domain is particularly common for the cryptocurrency community because of the historic frequency of fraudulent activities (e.g., money laundering and illegal trading) as well as cyberattacks (e.g., denial-of-service and Sybil attacks) on most cryptocurrency blockchains.

We found a total of 12 sources that discussed work related to fraudulent activities in the network—4 analysis articles focused on specific fraudulent events, such as laundry services, online drug market places, denial-of-service attacks, and anonymity of users. Besides these, there were 7 visualization articles that proposed fraud detection tools while only 1 online source allowed experts to investigate criminal activities in blockchains. All of the existing cybercrime detection tools were designed to investigate cybercrime and fraudulent activities after they occur. Therefore, we still lack tools to monitor and automatically detect potential cybercrime activities in real-time.

Cryptocurrency Exchanges Analysis: Cryptocurrency exchanges are an important target domain particularly to cryptocurrency blockchains such as Bitcoin. Tools in this target domain present exchange market statistics together with financial data related to different cryptocurrencies, such as the exchange rate between a cryptocurrency value to the US Dollar. In contrast to the transaction detail analysis task domain, tools that targeted cryptocurrency exchanges focused on external data (mostly currency conversion rates and trading volume) rather than looking at fine-grained information on any specific transactions or their aggregation.

In this review, we did not systematically collect all tools that focused on market-related data without also including some data stored on a blockchain. Instead, we included 10 sources that visualize cryptocurrency exchange statistics together with blockchain data—9 online sources and 1 visualization article. A comprehensive review of online cryptocurrency exchange sources can be found in our previous work [46].

P2P Network Activity Analysis: Several sources targeted peer-to-peer (P2P) network activity analysis. This target domain concerns the presentation of aggregated statistics that gives an overview of activities in the P2P network, such as mining, transaction rates, transaction volume, mempool statistics, sometimes coupled with inferred geographic locations. We found a total of 22 interactive visualization tools—21 online sources and 1 visualization article—for analyzing P2P network activities. On the other hand, 5 analysis articles presented longitudinal analyses of blockchain network characteristics. In contrast to transaction network analysis domain, P2P network analysis focuses the entire P2P blockchain network, i.e. what is its state and how well does it work as a whole system.

Casual/Entertaining Information Communication: In addition to the more serious analysis target domains outlined above, we also found 12 sources exclusively from the web that were built to attract the attention of novice audiences to blockchain technologies and engage them through casual information visualization.

5.5 Visualization Types

We analyzed visual encodings in blockchain visualization tools and found six common visualization types:

Number of sources with different visualization types					
Charts	Time Series	Graphs	MD Vis	Maps	Casual
31	40	29	5	10	12

Time series (53%) were the main visualization type for blockchain data, followed by basic charts (41%), and tree and graph visualizations (38%). This is not surprising because blockchain components contain time-stamped information and addresses are connected via transactions forming transaction networks. Other visualization types often showed blockchain data in a specific context. For example, map-based visualizations (13%) displayed global blockchain node distributions. We found casual information visualizations (16%) only in casual/entertaining sources. Multi-dimensional visualizations (7%) were used to encode blockchain component data that contain multiple attributes.

Charts: Charts showed predominantly two or three (never four) data dimensions using basic representation types such as bar charts, pie charts, histograms, scatterplots, and heatmaps. We counted basic charts with a time dimension as “time series” but found 31 sources that included basic charts without a time dimension—17 online sources, 12 data analysis articles, and 2 visualization articles.

Time Series: Time series were the most common visualization type because timestamps are an essential attribute in blockchain data. We found 40 sources that had at least one time series element—20 online sources, 13 data analysis articles, and 7 visualization articles. Most commonly time series showed the activity of a blockchain address or entity summarized across different time granularities. Time series were often presented as *line plots* and *bar graphs* with a temporal x-axis. We also found other visualization techniques for time-oriented data including *tile maps* [30], a heat map with calendar divisions to encode activity statistics with one or two temporal dimensions.

Tree and Graph Visualizations: Tree and graph visualizations were a common choice to represent money flows and transaction networks in the sources we surveyed. These representations typically showed the connection of transactions from input addresses to output addresses. We found 29 sources with trees or graphs—10 visualization articles, 10 data analysis articles, and 9 online sources. *Node-link diagrams* were the most common technique to show the connectivity of blockchain components. A few sources used different graph visualization techniques, such as an *adjacency matrix*, a *Circos diagram* [27], or *customized visualizations*.

Multi-dimensional Visualizations: We refer to multi-dimensional visualizations as visualizations designed for showing data of higher dimensions than basic charts, including *small-multiple glyphs*, *self-organizing maps*, a *classification tree*, a *3D scatterplot*, *spider charts*, and a *parallel coordinates plot*. We found only 5 sources in our survey that included multi-dimensional visualizations—3 visualization articles and 2 online sources.

Map-based Visualizations: Map-based visualization was a common technique to display geographical information associated with the blockchain. We found 10 different



Fig. 2. BitInfoChart [O9] shows the historical Bitcoin balance of an address in both Bitcoin currency and exchange rate to the US Dollar.

sources—8 online sources and 2 data analysis articles. All of these sources in our survey visualized *thematic maps*, showing statistical information about a blockchain related to geographic area. We found 3 point maps, 2 density maps, 4 choropleth maps, and 2 virtual globes in 3D. For example, Lischke and Fabian [A10] included 2 map-based visualizations: an area map and a choropleth map.

Casual Visualizations: We grouped a set of non-standard, custom-made graphical representations of blockchain data as *casual information visualizations* [36]. All casual visualizations came exclusively from 12 online sources. These sources did not use common charts or plots as described above. Instead, they depicted basic blockchain components in unique ways to attract attention. Visualizations used included 3D animated balls, 3D toy models of buildings and roads, or a flying balloon in a 360-degree view. We even found one data physicalization project for blockchain data: [O25], [O38].

6 DETAILED ANALYSIS PER TASK DOMAIN

Since task domains are an important distinguishing factor in our classification scheme, we describe patterns of sources for each task domain in greater detail. We highlight some representative examples and discuss blockchain data and visualization types that are commonly used in those sources.

6.1 Transaction Detail Analysis

The main goal of transaction detail analysis is to analyze transaction patterns for individual blockchain components (i.e. transactions, addresses, and blocks) or derived entities in blockchain networks. We distinguish three types of transaction detail analyses based on the blockchain data visualized: (1) visualization of financial transactions, (2) visualization of blocks, and (3) visualization of multiple entities.

Visualization of financial transactions: Visualizations of this category allow intermediate users to search and explore the details of cryptocurrency value transactions, addresses, and blocks, for example for the Bitcoin ([O7], [O9], [O10], [O11]), Ethereum ([O1], [O9], [O10], [O11], [O21], [O24], [O26]), Litecoin ([O9], [O11]), and Dash ([O9], [O18]) blockchains.

Most tools in this task domain focus on representing financial transaction activities in a specific address or entity such as the total received, sent, or the balance amount over time in the form of *time series*. BitInfoCharts [O9] is a representative example in this category that uses *line plot* time series to show the balance amount of individual addresses for several cryptocurrencies, including a conversion rate to US Dollar (Fig. 2). Other time series visualizations have also

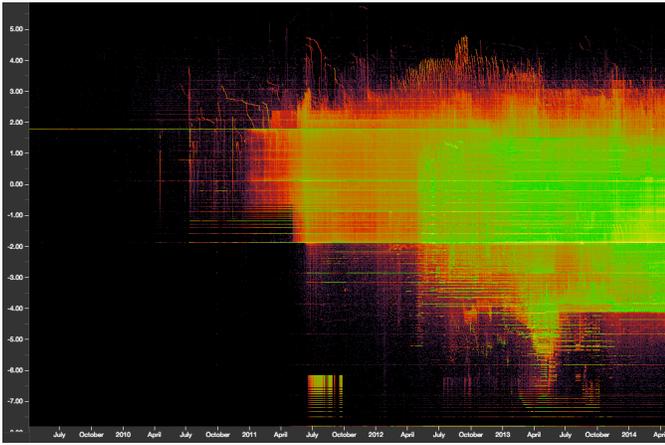


Fig. 3. Schretlen et al. [V12] proposed a tile map encoding which the authors later applied to the distribution of Bitcoin amount exchanged over time. The x-axis represents the time of transactions and the y-axis represents the \log_{10} Bitcoin amount. The frequency of transactions in each value is encoded as the color scale. It is designed to help detect outlier activities or interesting patterns for further investigation. (Image from a public presentation, used with permission of Uncharted Software Inc.)

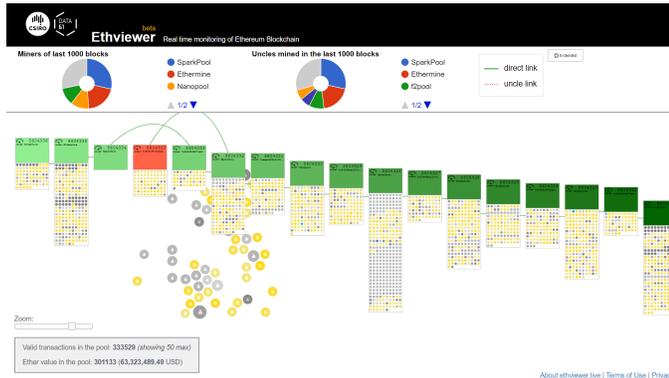


Fig. 4. Ethviewer [O23] visualizes the Ethereum blockchain in real-time. Each circle represents a transaction in the mempool waiting to be collected in the block. The color encodes different types of transactions and the size encodes the value associated with the transaction. When the transactions are included in the block, the screen shows the animated circles moving to the box. The color of the box header changes from red to green as the block is confirmed. At the top of the webpage, two pie charts show information about gas associated with transactions.

been used. Blockchain Explorer [V8], for example, visualizes weekly or monthly transaction volumes as a *tile map* [30].

In contrast to most sources in this task domain that show aggregated statistics on transactions and addresses, Schretlen et al. [V12] proposed an interactive visualization for exploratory analysis of transaction data stored in the Bitcoin blockchain. It used a large scale *tile map* (Fig. 3) to display the distribution of Bitcoin transaction values.

Visualization of blocks: We found 3 sources that visualized the content of blocks: [V4], [V14], [O31]. Chawathe [V4] applied a *self-organizing map* to create a low-dimensional representation of transactions in a block. The self-organizing map is visualized as a hexagonal grid of wind rose plots to show the main characteristics of transaction groups in a block. Another tool, Ethviewer [O23], shows the real-time transaction pool in Ethereum. The tool shows a chain of linked

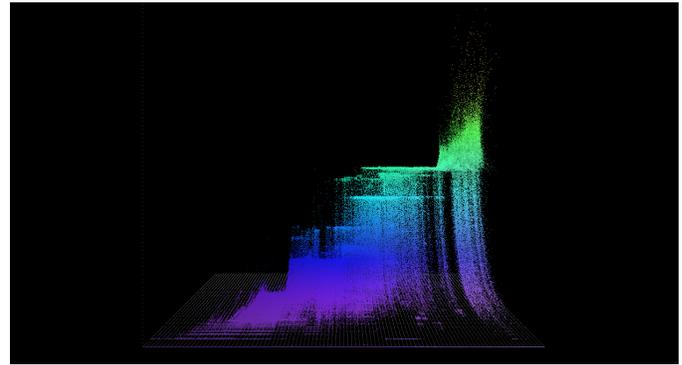


Fig. 5. OXT Landscapes [O31] provides an interactive 3D scatterplot to explore attributes of all blocks in the Bitcoin blockchain. Each dot represents a block in the Bitcoin blockchain in 3-dimensional space. Each dimension encodes an attribute describing the block, including block size, number of transactions, number of addresses in the block, or total transaction fee.

blocks as a *node-link diagram* (Fig. 4). OXT Landscapes [O31] is the only source that uses 3D visualization to represent attributes of blocks as a *3D scatterplot* (Fig. 5).

Visualization of multiple entities: We found 2 sources that presented financial information of entities allowing experts to explore single or a group of entities and drill down to see transaction behavior: [V7], [V14]. Attributes that characterize entities were usually represented using *multi-attribute visualizations*.

BitConduite [V7] is a visual analytics tool for exploring entities in Bitcoin using multiple views (Fig. 6). It allows analysts to filter groups of entities visually from a *classification tree*. The tool also clusters groups of entities that have similar activity patterns, and encodes them as *radar charts* to represent quantitative attributes of entities, such as the number of transactions, time active, and the average number of input addresses per transaction. BitExTract [V14] is another visual analytics tool that also falls in the category of entity visualizations, focusing on the analysis of activities among Bitcoin exchanges, including transactional volume, market share, and connectivity between exchanges. (Fig. 7 A, B, C).

6.2 Transaction Network Analysis

Transaction network analysis sources showed generally three kinds of information: (1) transaction networks, (2) the network of entities, and (3) value flows tracing the transfer of cryptocurrency values through transactions over time. These sources were always represented as *tree and network visualizations*. In particular, *node-link diagrams* were most often used to show the connectivity among blockchain components. A common technique to arrange nodes was the *force-directed graph layout*.

A **transaction network** is a directed bipartite graph connecting addresses via a transaction. There are two kinds of nodes: one type for addresses and one for transactions. Two kinds of directed edges exist in such a graph. Input edges connecting input address(es) to a transaction and output edges connecting a transaction to output address(es).

We found real-time transaction networks in three online visualization tools: [O8], [O17], [O18]. For example, Daily-Blockchain [O17] shows a live Bitcoin transaction network

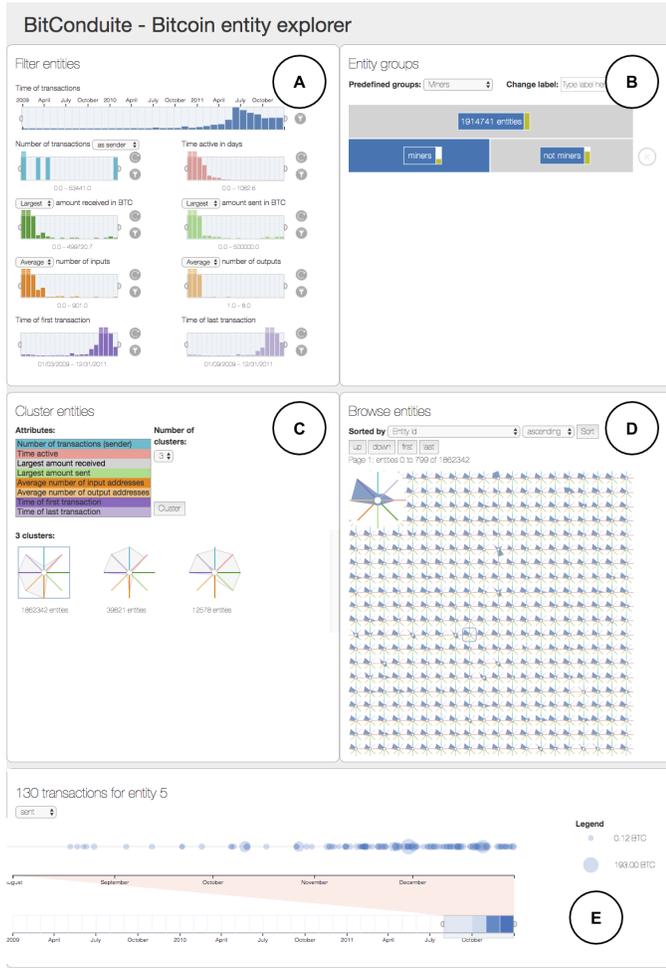


Fig. 6. BitConduite [V7] is a tool to analyze entity groups for the Bitcoin blockchain. (A) The filter view provides a *time series* and multiple *histograms* to filter Bitcoin entities. (B) The tree view is a *classification tree* used to show the result of applied filters. (C) After that, entities are clustered into groups. Averages of each metric visualized as *star glyphs*. (D) Entities in a selected group are presented as *glyphs* encoding their attributes. (E) Data analysis experts can select an entity and explore its activity timeline.

where the nodes evolve over time. Users can zoom into and hover over nodes to see additional information. However, the transaction network is growing over time which decreases the performance of the graph rendering. Bitforce5 [O8], in contrast, only shows a limited number of the most recent transactions. Therefore the performance remains the same over time.

Several visualization articles proposed tools to explore transaction networks based on specific events: [V2], [V6], [V9], [O36]. The BlockchainVis [V2] tool displays a fully connected transaction network of a transaction or an address entered by the user (Fig. 8). McGinn et al. [V9] proposed a system to display a transaction network on a large screen on which users can pan, zoom, and hover over to get a better overview or more detail. Blockchain 3D Explorer [O36] is the only tool in this domain that visualizes a transaction network as a *3D graph*. It also supports virtual reality systems for Google Cardboard to explore the blockchain network in an immersive way. Instead of showing a static transaction network as a node-link diagram, Bitcoin Entity Explorer [V6]



Fig. 7. BitExTRACT [V14] is a multiple-view visual analytics tool to compare the transaction activities and relationships of Bitcoin exchange entities. (A) The comparison view is an interactive *parallel coordinates plot* to compare financial activities of Bitcoin exchanges ordered by user-defined ranks (rows in each bar) over months (axes). (B) The exchanges list view displays transactional volumes of Bitcoin exchanges over months as a *time series bar chart*. (C) The massive sequence view uses a *heat map*-like representation to show the transaction surplus of exchanges through their entire history of those entities. (D) The connection view is a node-link diagram encoding the volume of transactional exchanges. Each node is sorted and colored based on the continent of origin of the exchange. The edge thickness shows the intensity of activities between two exchange nodes. The view becomes an ego-network graph when analysts drag the exchange node of their interest into the center of the network.

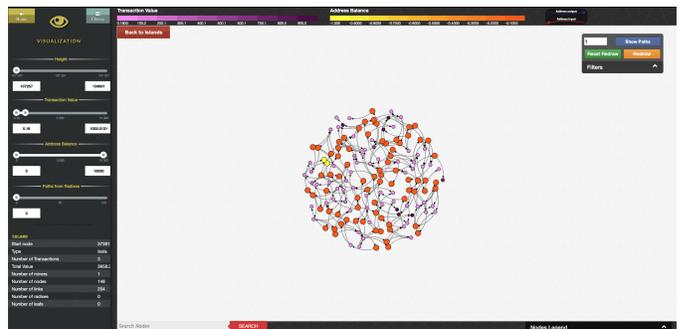


Fig. 8. BlockchainVis [V2] allows for creating transaction networks and analyzing the connectivity of transaction exchanges in Bitcoin. The tool provides a filter panel that allows analysts to filter out unimportant nodes from the graph based on block height, transaction value, and address balance.

is an exception in that it presents a transaction activity timeline of a chosen entity with a timeline-based squarified graph layout connecting input and output addresses over time (Fig. 9).

A **network of entities** shows the connectivity between entities in the blockchain network: [V14], [A15], [O19]. Nodes represent entities and edges represent connectivity through transactions. For example in Bitcoin, an edge represents the total amount of exchanged values between two entities and is absent if no value was exchanged.

The Bitcoin Big Bang [O19] is an online visualization presenting a network of entities as a *node-link diagram* connecting well-known wallets and highlighting the transaction volume between them. The color of nodes represents the type of nodes, such as payment processors, dark marketplaces, and gambling services. It adds a temporal dimension to the node-link diagram by arranging the node distance from

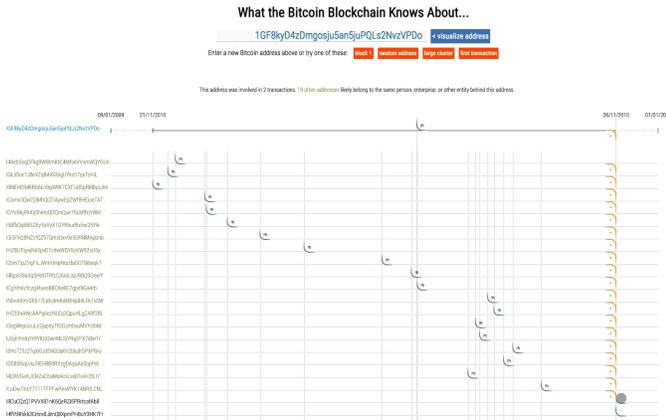


Fig. 9. The Bitcoin Blockchain Entity Explorer [V6] shows a transaction activity timeline of addresses that likely belong to the same entity. The timeline shows an entity's list of sending and receiving addresses that are involved in transactions and highlights the sending and receiving activities with glyphs that encode transfer amount and direction. This figure shows the activity of a Bitcoin miner who combines 20 mining transactions and sends the combined value to two output addresses. Green addresses on the left likely all belong to the same miner.

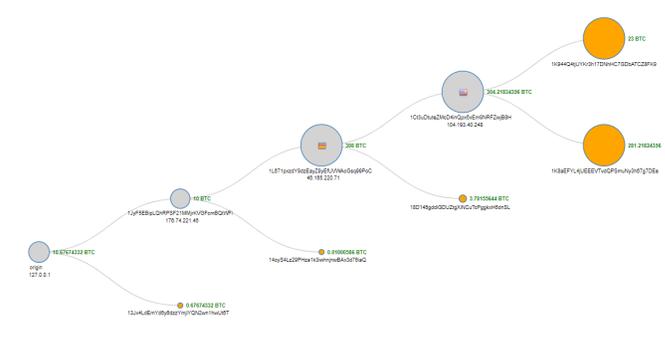


Fig. 11. Blockchain.info [O10] visualizes the flow of Bitcoin money as a tree graph. The size of the nodes is proportional to the total value sent or received from an input or output address.

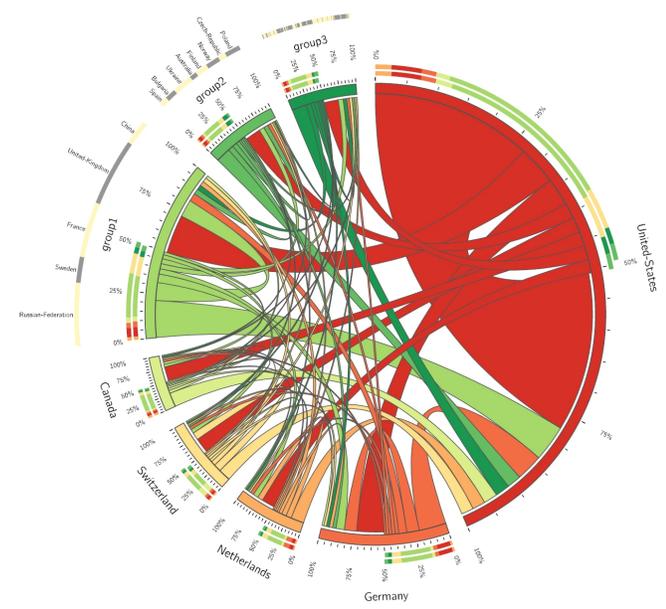


Fig. 10. Parino et al. [A15] present the movement of Bitcoin value among countries in 2013 as a *Circos* diagram. The flow of transaction exchanges is a ribbon in which the size is proportional to the amount transferred between the two countries. The external circle shows the total sending and receiving money for each country. (CC BY 4.0)

the center based on the time of their first appearance. Users can select a node to highlight the transaction flow on that node. BitExtract [V14] has a connection view that shows the relationship of exchange entities using a *circular network layout* to investigate the interaction of entities in the blockchain network (Fig. 7 D). Parino et al. [A15] describe a flow network of Bitcoin transactions aggregated at the country-level. The authors use a *Circos* diagram [27], also known as *dependency wheel*, to visualize the total transaction exchanged between major countries (Fig. 10).

A *value flow* presents traces of cryptocurrency value given a particular transaction or address of interest. We most

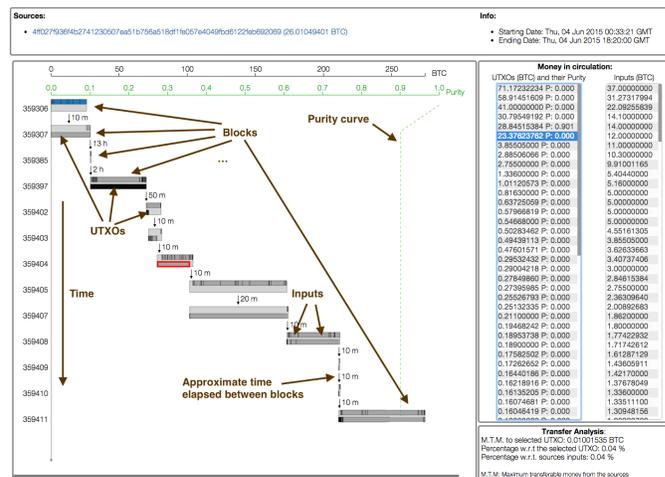


Fig. 12. BitConeView [V5] is a graphical tool to analyze the pattern of Bitcoin value flows over time by means of transactions. Each block contains spending transactions in the upper bar and unspent transaction output (UTXOs) in the lower bar. The width of each bar indicates the amount of a transaction. (© 2015, IEEE)

commonly saw it represented as a *tree diagram* connecting the flow of values in chronological order. In this layout, a node represents a transaction or address and an edge represents the amount of value exchanged. We found 6 sources that visualized value flows: [V1], [V5], [V13], [A17], [O10], [O12]. For example, Blockchain.info [O10] provides a *tree diagram* in which users can click through tree levels to follow value flow from connected input and output addresses (Fig. 11). Instead of presenting the value flow as a tree structure, BitConeView [V5] provides a unique diagram showing the value flow of a seed transaction as it appears in blocks from top to bottom (Fig. 12).

All of the examples above present static graphs that do not consider the timestamp of transactions. We found a unique value flow graph in BitInfoCharts [O9] that visualizes the flow of transactions over the entire history of a cryptocurrency blockchain (Fig. 13) as a kind of *node-link diagram* arranged using a linear layout. The same kind of graph also appeared in McGinn et al. [A12] as an *adjacency matrix* representation.

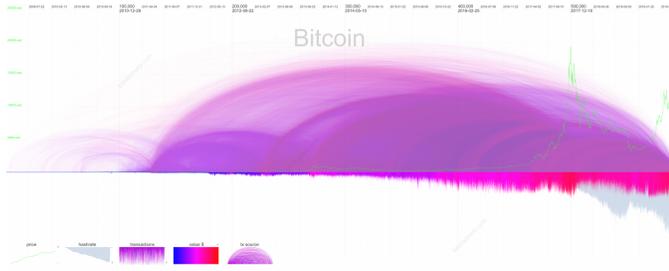


Fig. 13. The Bitcoin visualization in BitInfoCharts [O9] shows the transaction exchanges network between blocks for the entire Bitcoin network. The graph uses a linear graph layout to display the evolution of block height on the x-axis. The edges encode the total value of transaction outputs in a block that has been spent in the next blocks. Additional line and bar charts are added to show descriptive information on the blockchain including price, hash rate, and the number of transactions.

6.3 Cybercrime Detection

The cybercrime detection task domain includes tools that are able to detect suspicious transactions and entities or investigate cyber-attack events. Sources in this task domain also include shared characteristics with sources in the transaction detail analysis and transaction network analysis domains, but additionally, have very specific user tasks and subsequently focused features for cybercrime detection. Current blockchain visualization tools for cybercrime detection focus on two questions: (1) value flow analysis to see how cryptocurrency value is propagated and (2) transaction network analysis to see how the blockchain network reacted in light of cybercrime events.

One way to detect fraudulent financial activities in cryptocurrency blockchains is to analyze value flows. Exemplary tools dedicated to **in-depth value flow analysis**: [V1], [V5], [V13], [O12]. BlockSeer [O12] is an online money flow analytical tool that allows blockchain experts to construct a deep transaction flow graph to trace money laundering and stolen Bitcoins. Di Battista et al. [V5] and Ahmed et al. [V1] proposed transaction graph construction tools dedicated to analyzing the mix of Bitcoin stolen money in the transaction flow (i.e. taint analysis). To analyze the degree of money mixing from the original transaction, Di Battista et al. [V5] introduced a *purity* measurement, the degree that a seed transaction is mixed with other transactions. Ahmed et al. [V1] use a *First-In-First-Out (FIFO)* algorithm to track the diffusion of tainted transactions in both forward (i.e. starting from a stolen coin to the following transactions) or backward (i.e. tracing the previous transactions until the origin of a tainted coin is found) directions. The authors developed an interactive visualization tool to display taint propagation as a *node-link tree visualization*.

We also found visualization sources for **transaction network analysis on specific events or group of entities**: [V2], [V9], [V11], [V13], [A4], [A6], [A12], [A14], [A16]. For example, BitVis [V13] uses multiple graph visualizations with a filtering panel to display transaction networks for detecting abnormal and suspicious Bitcoin entities. Two articles from McGinn et al. [V9], [A12] show how their tool can be used to visualize a transaction network during cybercrime attacks, including denial-of-service attacks where an attacker tries to fill up a block with spam transactions (Fig. 14). The

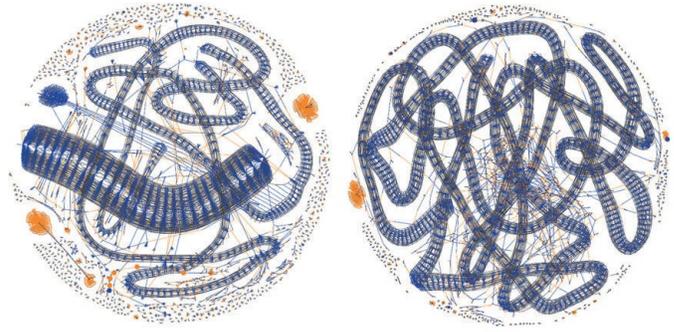


Fig. 14. McGinn et al. [V9] displayed transaction networks during transaction rate attack incidents in the Bitcoin network at block heights 364133 (left) and 364618 (right). (CC BY 4.0)



Fig. 15. Blockchain.info [O10] provides a candlestick chart displaying the historical price of cryptocurrency exchanges.

BlockChainVis tool [V2] is another transaction network tool that allows to filter specific parts of the transaction network during an event of interest (Fig. 8). It has been used to analyze the *WannaCry ransomware incident* on May 12th, 2017 [A6]. Other data analysis articles performed ad-hoc analyses of transaction networks during attacks on the Bitcoin network, including money laundry services [A4], online drug marketplaces [A14], and the network of Bitcoin thefts [A16].

6.4 Cryptocurrency Exchange Analysis

Cryptocurrency exchanges convert cryptocurrency values into real-world currencies, such as US Dollars. As such, much of the data related to these exchanges is not stored on the blockchain itself. Our analysis of cryptocurrency exchange data is limited to those sources that relate any externally captured data to data stored on the blockchain. Our sources in this task domain either (1) cover the conversion of cryptocurrency value to US Dollar for blockchain components such as individual addresses or blocks (see Sect. 6.1) or (2) provide an additional view that relates information on blockchain components to market statistics, such as historical price, trading volume, and market capitalization. A market statistics view informs intermediate and expert users about blockchain value in the external environment, for example, as converted to US Dollar or Euro. Market statistics can help to understand how the mechanisms inside the blockchain network are affected by the cryptocurrency economy at large.

As already described in the “transaction detail analysis” task domain section (Sect. 6.1), a first type of source visualizes **conversion rates for cryptocurrency values**: [O6], [O9], [O10], [O21]. The second type of sources visualizes

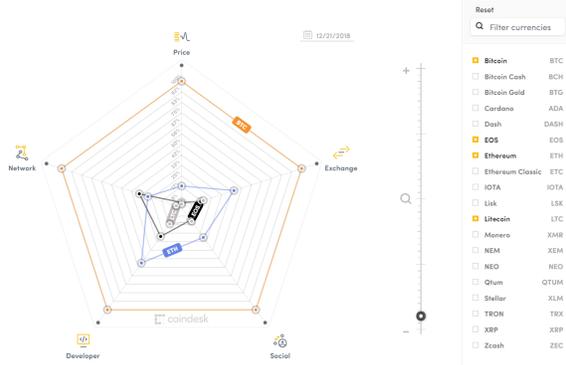


Fig. 16. The Crypto-Economics Explorer in CoinDesk [O16] provides a comparison view of multiple market measures for different blockchains.

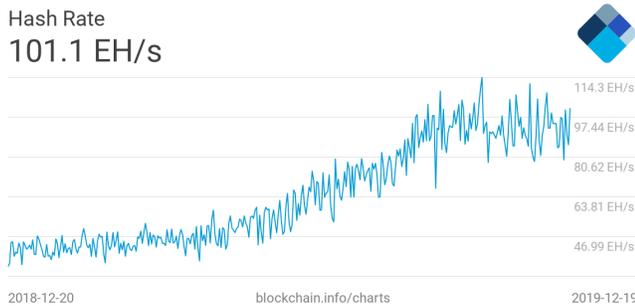


Fig. 17. Blockchain.info [O10] provides time series charts to show aggregated statistics of the Bitcoin network. This example shows the hash rate of Bitcoin over time. The number of nodes in each clustered location is encoded as the bubble size.

market statistics: [O10], [O16], [O21]. These sources mostly used *time series* to display the historical exchange rate and market volume for different time scales (i.e. hours, days, weeks, months) in addition to more detailed information on individual transactions, addresses or blocks. For example, Blockchain.info [O10] provides a market view for various cryptocurrencies (Fig. 15). CoinDesk [O16] is a unique online tool in this category that shows summarized measures of the size and investment opportunities of several cryptocurrencies. It presents a *spider chart* to compare multiple measures related to price, exchanges, social media, developers, and the overall network size (Fig. 16).

6.5 P2P Network Activity Analysis

Blockchains are decentralized systems running with client nodes in a peer-to-peer (P2P) network architecture. Understanding the activities within the P2P network helps intermediate and expert users to track the current status of a block as a result of overall activities among participants in the network. Sources in this task domain use two kinds of visualizations: (1) *time series* to show the aggregated statistics of the P2P network, and (2) *map-based visualizations* to show the geographical distribution of blockchain usage around the world.

Most sources in this task domain present **P2P network statistics**, calculated from aggregated node activities in the blockchain network over time. All sources use time series visualizations to represent changes of the blockchain network



Fig. 18. EthStats.net [O35] provides a real-time dashboard to monitor Ethereum network status.

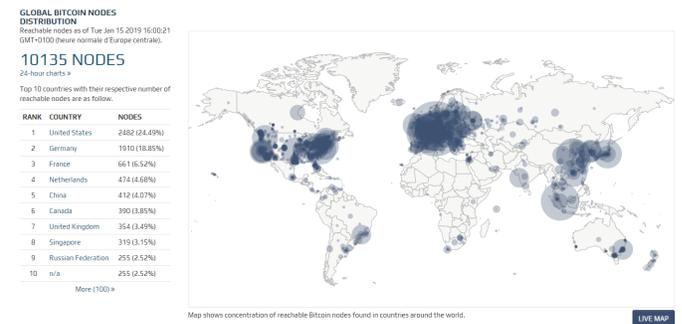


Fig. 19. BitNodes [O45] shows the geographical distribution of the number of active Bitcoin nodes in a point map. The number of nodes in each clustered location is encoded as the bubble size.

over time. For example, Blockchain.info [O10] provides a long list of time series charts to display a wide range of Bitcoin network statistics, for example, the total hash rate (Fig. 17), average block size, total transaction fee, and mining difficulty. A dashboard proposed by Bogner [V3] presents time series and basic charts on Ethereum statistics and highlights outlier data using anomaly detection techniques. EthStats.net [O35] is a rare example that provides a real-time dashboard for monitoring network status and active nodes in the Ethereum blockchain (Fig. 18).

Analyzing the **global distribution of a blockchain network** involves observing the density of blockchain nodes and transactions around the world. Public blockchain data does not inherently include geographic information about senders, receivers, or blockchain nodes. However, when nodes in the blockchain network have associated IP addresses, these can be used to infer the geographic location of a node with a degree of uncertainty [7], [23]. The geographic origin of a transaction can then be inferred from the IP address of the first node that relayed it [25], [39].

We found 9 sources that display the number of nodes active in the blockchain P2P network ([A10], [A15], [O18], [O20], [O21], [O35], [O45]), and transaction origins ([A10], [O4], [O44]). All of them display geography information in *map-based visualizations*—the only task domain that used this kind of visualization type. For example, BitNodes [O45] implemented a node crawler to gather reachable node locations to estimate the global distribution of Bitcoin nodes (Fig. 19). Different types of map-based visualizations we saw

are discussed in Sect. 5.5.

6.6 Casual/Entertaining Information Communication

Sources in this task domain generally provided original and experimental visualization of blockchain components that are distinct from any of the ones used for the above task domains. Our sources encoded, for example, attributes of transactions and blocks as custom objects—often in 3D—with animation and real-time updates.

To show the wide variety of visual encodings in this category we briefly discuss a few examples: BitBonkers [O3] shows live Bitcoin transactions as 3D balls falling on a plate each time a new transaction is broadcasted to the network. BitcoinCity [O34] represents Bitcoin transactions as 3D toy models of buildings along the road that are moving as new transactions are created. BitListen [O33] presents transactions as animated bubbles floating on the screen, producing notes that combine into improvised music. Symphony of Blockchains [O30] includes a combination of interactive visual representation of Bitcoin data that allows web visitors to browse blocks as a 3D visual representation and navigate through a flight-simulator mode, along with background audio representing the network hash rate and using a unique tone for each of the transactions in the block. Bitcoin VR [O43] is an open-source project that visualizes Bitcoin transactions as balloons flying over a 360-degree view.

7 DISCUSSIONS AND OPEN CHALLENGES

This section reports general observations that we gathered while compiling this survey about blockchain data visualization. In particular, we discuss the state-of-the-art of existing visualization practice in regards to blockchains as well as opportunities for future research.

7.1 Blockchain Visualization: Research vs. Practice

Blockchain enthusiasts and startups have quickly established the need to better communicate what happens in and around blockchains. In our survey, online sources that use visualizations outnumber by far dedicated visualization articles that describe how to visualize blockchain data for various types of tasks. Yet, most online sources only use and require simple charts or time series visualizations, where research input might not necessarily be required. Where research can contribute the most is by offering systems for in-depth analysis and this is, indeed, what most current visualization articles focus on. Yet, we also see opportunities for research to offer more fluid interaction and exploration capabilities, in particular for those online sources that offer simple graph visualizations and explorations across time. In particular, we saw a need for interactive tools that allow users to explore how their activities manifest on the blockchain and to show what data can or could be inferred about them through their blockchain use. This could in particular help novices assess and adjust their transaction patterns.

7.2 Blockchain Analysis vs. Visual Analytics

In our survey, we covered visualizations published as part of blockchain analysis articles—for which conveying the result of the analysis and not the design of visualizations were the focus. As such, visualizations in the analysis articles were meant for communication of scientific results rather than exploration. Most analysis articles focused on transaction network analysis—indicating a current research focus for which, according to our survey, few tools exist that experts could make use of. As a result, most analysis articles only showed time series plots or basic charts of network measures. In addition, we observed that blockchain analysis articles often did not conduct an in-depth analysis of the entire blockchain, probably because of the large size of the blockchain and the lack of simple ways to explore and statistically analyze the data in its entirety. Blockchain analysis articles often focused on a higher-level data analysis of the blockchain network (i.e. the global blockchain network or longitudinal study of P2P network analysis). Considering the demand of data analysis experts and decision-makers to better understand blockchain activities, it would be an opportunity for the VA community to come forward with more advanced tools that support both higher-level and more in-depth analyses of blockchain data. The growing size and dynamic nature of blockchains require techniques that work on multiple levels of data aggregation and show data updates well. Tools also need to provide complete overviews of the network and also allow experts to interactively drill-down and see close details of transactions or individual actors on the network. In particular, tools that allow experts to take on specific viewpoints such as individual entities in the network (e.g., people, enterprises, miners), historic events (e.g., cyber-attacks), or network-related events (e.g., halving days or forks) are still missing. Finally, there is a lack of tools tailored to the specific needs of particular experts, in particular economists and blockchain managers. Economists want to understand the activities on the blockchains and compare them with related economic activities in the real world. Consortium blockchain managers need to understand e.g., how their blockchain evolves, according to their plans and how it compares to other blockchains.

7.3 The Dominance of Cryptocurrency Blockchains

Cryptocurrencies are the most widely used applications of blockchains nowadays. All visualization sources we found addressed cryptocurrency blockchains. Most sources visualized Bitcoin data since it is well-known, adopted, and has a high number of users. The second most frequently visualized blockchain is Ethereum. All Ethereum-related visualizations focus on the cryptocurrency aspect of value exchanges among entities, not the smart contract functionality that makes Ethereum different from Bitcoin. We did not find any source dedicated to other blockchain types, including consortium and private blockchains. Even though the concept of cryptocurrency blockchain sources should be able to apply to other blockchains, there are some differences in the detailed mechanism (e.g., the transaction data structure and mining protocol) that need to be considered in the design of visualization systems. The visualization of private and consortium blockchains, or blockchains for non-cryptocurrency

use cases such as in healthcare [26], mobility [18], supply chain management [38], or for government services [47] are fruitful areas for future works.

7.4 The Missing Context of Blockchain Data

The majority of visualization sources presented detail about blockchain components and overviews of network activities in the transaction detail analysis and P2P network analysis task domains. What is missing in most sources is information that provides context for monitoring and analysis of activities in the blockchain, including the identification of entities, geographic information, social network activity, or historic events.

In the big cryptocurrency blockchains, users remain anonymous by using multiple non-identifiable addresses to send and receive cryptocurrency value. Identity is only revealed if people or enterprises post their addresses openly and connect them to other pieces of identification, such as forum user names, their websites, etc. If one wants to understand how cryptocurrencies are used in regards to what is known about fiat currencies, information about which addresses belong to the same entities (such as individual users, businesses receiving Bitcoin for payment, or exchanges) is required. Heuristics exist that help to cluster addresses and identify entities with a degree of uncertainty [37] but entity-based visualizations are, nevertheless, rare (e.g., [V7], [V14], [O10]).

The blockchain network involves many anonymous participants interacting with each other through committing and validating transactions in the P2P network. It would be interesting to study the collective activities of participants in the blockchain network in light of historical events, such as from volatility of market prices, cyber-attacks (e.g., Mt. Gox hack, Bitcoin theft, and denial-of-service attacks), government regulation, or changing in mining rewards. The existing visualization sources have been used mostly for dedicated task domain (i.e. transaction detail analysis, transaction network analysis, or P2P network analysis) without providing the capability to investigate historical events in a holistic view.

7.5 Open Blockchain Visualization Challenges

Blockchain technology produces a large transactional dataset, rich in details, including sophisticated maintenance mechanisms which are interesting for analysis. The complexity of working with blockchain data comes both from both technical aspects and its social component related to many social networks in general. Blockchain data is more complex than most social networks due to its pseudonymous use of addresses and the nature of the data is carries, usually monetary value. Therefore, it is unlikely that simple views will ever be able to convey the richness of information it carries.

Most visualization sources we surveyed focused on using common chart types (i.e. time series and basic charts) with basic interaction techniques (i.e. querying and zooming), that are not sufficient for advanced analysis tasks. Here we list several opportunities to improve existing visualizations used and opportunities to develop new dedicated representations.

Multiple views visualization of blockchain data: Sources from data analysis articles and online sources usually

provide many single view charts showing a particular blockchain measure over time. Single disconnected views make it difficult to relate multiple blockchain characteristics to each other. BitExTract [O23] is one example that broke the trend and proposes a dashboard with multiple chart elements for analyzing transaction activities among Bitcoin exchange entities. Yet, additional sophisticated interaction techniques for visual comparison [17] would help to connect views and generate more comprehensive insights.

New visual representations for transaction network analysis: Existing network visualization sources present transaction networks and value flows as static graphs at specific points of interest (i.e. a point in time, a specific block, or for a group of entities). However, those tools mostly do not consider the temporal evolution of the network and, in particular, changes of blockchain connectivity over time; we saw no dynamic network visualizations [6]. Besides, blockchain networks have specific network properties which could benefit from dedicated network layouts. For example, they are directed and time-oriented. For Bitcoin, the addresses can be clustered and so the raw graph can be simplified using well-known heuristics [20], leading to simpler visualizations.

Uncertainty visualization: Much of the contextual information related to Bitcoin comes with a degree of uncertainty. For example, heuristics to cluster Bitcoin entities are not sure to capture Bitcoin entities with 100% accuracy and IP-addresses of nodes in the P2P network are not necessarily reliable indicators of the geographic location of a node. In addition, analysis tools that may label certain transaction patterns as fraudulent or belonging to certain services (e.g., exchanges, mixing services, etc.) may make false predictions. Any uncertainty in the data should be made evident in the visualization [29] [35] and expose where viewers should be cautious about inferring insights and making decisions on the data.

Progressive visual analytics: Exploring Blockchain data involves navigating over large amounts of data for computing aggregated values on selections of the transactions or over time windows. These operations are usually simple to compute but take a long time. Work shown in research articles usually does computation offline to allow visualization tools to remain interactive. However, doing the operations offline means that the data exploration is limited to pre-computed values, and all the interactive tools we reviewed were limited in that respect. For continuously computing derived data when the Blockchain evolves, for e.g., maintaining the clustered Bitcoin entities up-to-date, techniques inspired by Boukhelifa et al. [9] could be applied. To allow more open-ended explorations on Bitcoin data, novel tools could rely on progressive data analysis and visualization [16], [40]. For example, Kinkeldey et al. report that BitConduite [V7] provides dynamic queries on time and attribute values to visualize aggregated information about Bitcoin transactions, but each filter operations make take a minute or so to complete depending on the amount of loaded data. Performing these operations progressively using methods reported by Moritz et al. [31] would drastically reduce the interactive latency and greatly improve the efficiency of exploring Bitcoin data, and could perhaps even cope with the complete Bitcoin blockchain data.

Evaluation: In our survey, we found that 9 out of 14

visualization articles evaluated the usability of their tools by either demonstrating case studies ([V2], [V8], [V9], [V11], [V13], [V14]) or performing user evaluation ([V5], [V7], [V9]). The fact that the majority of Bitcoin visualizations are published without formal task analyses or evaluation (at least as evident for research articles), is a clear sign that more visualization research is needed in this domain. This could be achieved by providing easier research access to updated blockchain data or developing easy to deploy analysis infrastructures. Some work in this direction has been started with e.g., BlockchainDB [21] and BlockSci [24] but will need further development to become usable on visual analytics infrastructures. This would allow researchers on Bitcoin analysis tools to focus on designing analysis tools rather than the data backend needed to extract blockchain data, compute usage metrics, and make them accessible for quick visual analysis.

8 CONCLUSION

Blockchain is a promising technology that will change the way we make electronic exchanges and maintain the integrity of data in untrusted and decentralized systems. In this work, we systematically reviewed 76 blockchain data visualization sources—14 visualization articles, 17 data analysis articles, and 45 online web-based tools. We classified those sources based on blockchain data, task domains, and visualization types, and describe the different kinds of tools in each task domain.

Among blockchain visualization sources in our systematic review, Bitcoin data is the most visualized, while the number of Ethereum sources has been increasing in recent years. Apart from financial applications for cryptocurrency blockchains, we did not see sources dedicated to other application domains, such as smart contracts, consensus, or private blockchains.

Most of the sources communicate basic aggregated statistics of blockchain elements using time series charts. Tree and network visualizations have been used to analyze the connected activities in the blockchain. However, blockchains do not include information that help to cluster addresses for pseudonymous users, which is necessary for certain types of analysis. The entity information needs to be derived from entity labeling datasets or computed by applying heuristics to group related addresses.

Many online blockchain visualization tools, target the P2P network analysis and transaction detail analysis task domains. Those tools basically allow intermediate users to explore the detail of blockchain components and look at the overview of network activity over time. On the other hand, expert users are interested in transaction network analysis and cybercrime detection (according to the number of data analysis articles we saw), but these types of analyses are not common in existing visualization tools.

In summary, we provide the first survey on blockchain visualization in a landscape that is still rapidly changing, with new applications appearing regularly and existing blockchains changing protocols and structures. As such, our survey provides a first overview of the blockchain visual analytics space and can help to further survey, observe, and classify emerging tools. Due to the increased adoption of

blockchains in recent years, the need for more VA tools will grow and we outline several fruitful opportunities for research on blockchain visual analytics. Application areas include the exploration and monitoring of activities in the blockchain network and more advanced tools for understanding different uses for blockchains.

ACKNOWLEDGMENTS

This research work has been carried out under the leadership of the *Institute for Technological Research SystemX*, and therefore granted with public funds within the scope of the French Program *Investissements d’Avenir*.

REFERENCES

- [1] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, “Evaluating user privacy in bitcoin,” in *Proc. Conference on Financial Cryptography and Data Security*, Springer, 2013, pp. 34–51. DOI: 10.1007/978-3-642-39884-1_4.
- [2] A. M. Antonopoulos, *Mastering Bitcoin: Programming the Open Blockchain*, 2nd. O’Reilly Media, Inc., 2017, ISBN: 978-1491954386.
- [3] A. Balaskas and V. N. Franqueira, “Analytical tools for blockchain: Review, taxonomy and open challenges,” in *Proc. Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, IEEE, 2018, pp. 1–8. DOI: 10.1109/CyberSecPODS.2018.8560672.
- [4] M. Bartoletti, S. Lande, L. Pompianu, and A. Bracciali, “A general framework for blockchain analytics,” in *Proc. Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers*, ACM, 2017, pp. 1–6. DOI: 10.1145/3152824.3152831.
- [5] D. Bayer, S. Haber, and W. S. Stornetta, “Improving the efficiency and reliability of digital time-stamping,” in *Sequences II: Methods in Communication, Security, and Computer Science*, Springer, 1993, pp. 329–334. DOI: 10.1007/978-1-4613-9323-8_24.
- [6] F. Beck, M. Burch, S. Diehl, and D. Weiskopf, “The state of the art in visualizing dynamic graphs,” in *EuroVis - State of the Art Reports*, R. Borgo, R. Maciejewski, and I. Viola, Eds., The Eurographics Association, 2014. DOI: 10.2312/eurovisstar.20141174.
- [7] A. Biryukov, D. Khovratovich, and I. Pustogarov, “Deanonymisation of clients in bitcoin p2p network,” in *Proc. Conference on Computer and Communications Security (CCS)*, ACM, 2014, pp. 15–29. DOI: 10.1145/2660267.2660379.
- [8] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, “Sok: Research perspectives and challenges for bitcoin and cryptocurrencies,” in *Poc. Symposium on Security and Privacy (SP)*, IEEE, 2015, pp. 104–121. DOI: 10.1109/SP.2015.14.
- [9] N. Boukhelifa, F. Chevalier, and J.-D. Fekete, “Real-time Aggregation of Wikipedia Data for Visual Analytics,” in *Proceedings of Visual Analytics Science and Technology (VAST 2010)*, Los Alamitos, CA, USA, United States: IEEE Computer Society, Nov. 2010, pp. 147–154. DOI: 10.1109/VAST.2010.5652896. [Online]. Available: <https://hal.inria.fr/hal-00690084>.
- [10] V. Buterin, *On public and private blockchains*, 2015. [Online]. Available: <https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/>.
- [11] CoinMarketCap OpCo, LLC, *Cryptocurrency Market Capitalizations*, Accessed: June, 2019. [Online]. Available: <https://coinmarketcap.com>.

- [12] Dash, *Dash - Dash Is Digital Cash You Can Spend Anywhere*, Accessed: June, 2019. [Online]. Available: <https://www.dash.org/>.
- [13] O. Dib, K.-L. Brousmiche, A. Durand, E. Thea, and E. B. Hamida, "Consortium blockchains: Overview, applications and challenges," *International Journal On Advances in Telecommunications*, vol. 11, no. 1&2, 2018. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-02271063>.
- [14] Dragonchain, Inc., *Dragonchain - Blockchain as a Service for Enterprises and Developers*, Accessed: June, 2019. [Online]. Available: <https://dragonchain.com/>.
- [15] Enterprise Ethereum Alliance, *Enterprise Ethereum Alliance*, Accessed: April, 2019. [Online]. Available: <https://entethalliance.org/>.
- [16] J.-D. Fekete, D. Fisher, A. Nandi, and M. Sedlmair, "Progressive data analysis and visualization (dagstuhl seminar 18411)," 2019. DOI: 10.4230/dagrep.8.10.1.
- [17] M. Gleicher, "Considerations for visualizing comparison," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 24, no. 1, pp. 413–423, 2017. DOI: 10.1109/TVCG.2017.2744199.
- [18] M. Gösele and P. Sandner, "Analysis of blockchain technology in the mobility sector," *Forschung im Ingenieurwesen*, pp. 1–8, 2018. DOI: 10.1007/s10010-019-00315-y.
- [19] S. Haber and W. S. Stornetta, "How to time-stamp a digital document," in *Proc. Conference on the Theory and Application of Cryptography*, Springer, 1990, pp. 437–455. DOI: 10.1007/3-540-38424-3_32.
- [20] M. Harrigan and C. Fretter, "The unreasonable effectiveness of address clustering," in *Proc. Conference on Advanced and Trusted Computing*, IEEE, 2016, pp. 368–373. DOI: 10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0071.
- [21] M. El-Hindi, M. Heyden, C. Binnig, R. Ramamurthy, A. Arasu, and D. Kossmann, "Blockchaindb-towards a shared database on blockchains," in *Proc. Conference on Management of Data*, ACM, 2019, pp. 1905–1908. DOI: 10.1145/3299869.3320237.
- [22] A. Janda, *WalletExplorer.com: smart Bitcoin block explorer*, Accessed: August, 2019. [Online]. Available: <https://www.walletexplorer.com/>.
- [23] P. L. Juhász, J. Stéger, D. Kondor, and G. Vattay, "A bayesian approach to identify bitcoin users," *PloS one*, vol. 13, no. 12, e0207000, 2018. DOI: 10.1371/journal.pone.0207000.
- [24] H. Kalodner, S. Goldfeder, A. Chator, M. Möser, and A. Narayanan, "Blocksci: Design and applications of a blockchain analysis platform," 2017. arXiv: 1709.02489 [cs.CR].
- [25] D. Kaminsky, *Black Ops of TCP/IP 2011 (Black Hat USA 2011)*, 2011. [Online]. Available: <https://www.slideshare.net/dakami/black-ops-of-tcpip-2011-black-hat-usa-2011>.
- [26] S. Khezr, M. Moniruzzaman, A. Yassine, and R. Benlamri, "Blockchain technology in healthcare: A comprehensive review and directions for future research," *Applied Sciences*, vol. 9, no. 9, p. 1736, 2019. DOI: 10.3390/app9091736.
- [27] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra, "Circos: An information aesthetic for comparative genomics," *Genome Research*, vol. 19, no. 9, pp. 1639–1645, 2009. DOI: 10.1101/gr.092759.109.
- [28] Litecoin Project, *Litecoin - Open source P2P digital currency*, Accessed: June, 2019. [Online]. Available: <https://litecoin.org/>.
- [29] A. M. MacEachren, R. E. Roth, J. O'Brien, B. Li, D. Swingley, and M. Gahegan, "Visual semiotics & uncertainty visualization: An empirical study," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 18, no. 12, pp. 2496–2505, 2012. DOI: 10.1109/TVCG.2012.279.
- [30] D. Mintz, T. Fitz-Simons, and M. Wayland, "Tracking air quality trends with sas/graph," in *Proc. SAS User Group International Conference (SUGI)*, 1997, pp. 807–812. [Online]. Available: <https://support.sas.com/resources/papers/proceedings/proceedings/sugi22/INFOVIS/PAPER173.PDF>.
- [31] D. Moritz, B. Howe, and J. Heer, "Falcon: Balancing Interactive Latency and Resolution Sensitivity for Scalable Linked Visualizations," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, ACM Press, 2019. DOI: 10.1145/3290605.
- [32] S. Nakamoto, *Bitcoin Forum*, Accessed: August, 2019. [Online]. Available: <https://bitcointalk.org/>.
- [33] —, *Bitcoin: A peer-to-peer electronic cash system*, 2008. [Online]. Available: <http://bitcoin.org/bitcoin.pdf>.
- [34] Namecoin, *Namecoin*, Accessed: June, 2019. [Online]. Available: <https://namecoin.org/>.
- [35] A. T. Pang, C. M. Wittenbrink, and S. K. Lodha, "Approaches to uncertainty visualization," *The Visual Computer*, vol. 13, no. 8, pp. 370–390, 1997. DOI: 10.1007/s003710050111.
- [36] Z. Pousman, J. Stasko, and M. Mateas, "Casual information visualization: Depictions of data in everyday life," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 13, no. 6, pp. 1145–1152, 2007. DOI: 10.1109/TVCG.2007.70541.
- [37] F. Reid and M. Harrigan, "An analysis of anonymity in the bitcoin system," in *Security and Privacy in Social Networks*. New York, NY: Springer New York, 2013, pp. 197–223. DOI: 10.1007/978-1-4614-4139-7_10.
- [38] S. Saberi, M. Kouhizadeh, J. Sarkis, and L. Shen, "Blockchain technology and its relationships to sustainable supply chain management," *International Journal of Production Research*, vol. 57, no. 7, pp. 2117–2135, 2019. DOI: 10.1080/00207543.2018.1533261.
- [39] M. Santamaria Ortega, "The bitcoin transaction graph anonymity," 2013. [Online]. Available: <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/23562/9/msantamariaOTFM0613memoria.pdf>.
- [40] C. D. Stolper, A. Perer, and D. Gotz, "Progressive visual analytics: User-driven visual exploration of in-progress analytics," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 20, no. 12, pp. 1653–1662, 2014. DOI: 10.1109/TVCG.2014.2346574.
- [41] T. Sundara, "Study on blockchain visualization," *International Journal on Informatics Visualization (JOIV)*, vol. 1, no. 3, pp. 76–82, 2017. DOI: 10.30630/joiv.1.3.23.
- [42] M. Swan, *Blockchain: Blueprint for a new economy*. O'Reilly Media, Inc., 2015, ISBN: 978-1491920497.
- [43] D. Tapscott and A. Tapscott, *Blockchain revolution: how the technology behind bitcoin and other cryptocurrencies is changing the world*. Portfolio, 2018, ISBN: 978-1101980149.
- [44] The Dogecoin Project, *Dogecoin*, Accessed: June, 2019. [Online]. Available: <https://dogecoin.com/>.
- [45] The Linux Foundation, *Hyperledger - Open Source Blockchain Technologies*, Accessed: June, 2019. [Online]. Available: <https://www.hyperledger.org/>.
- [46] N. Tovanich, N. Heulot, J.-D. Fekete, and P. Isenberg, "A systematic review of online bitcoin visualizations," in *Poster Proc. of the European Conference on Visualization (EuroVis)*, J. Madeiras Pereira and R. G. Raidou, Eds., The Eurographics Association, 2019. DOI: 10.2312/eurp.20191148.
- [47] S. Underwood, "Blockchain beyond bitcoin," *Communications of the ACM*, vol. 59, no. 11, pp. 15–17, Oct. 2016, ISSN: 0001-0782. DOI: 10.1145/2994581.

- [48] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, pp. 1–32, 2014. [Online]. Available: <https://gavwood.com/paper.pdf>.
- [49] J. Yli-Huomo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where is current research on blockchain technology?—a systematic review," *PLoS one*, vol. 11, no. 10, e0163477, 2016. DOI: 10.1371/journal.pone.0163477.
- [50] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: A survey," *International Journal of Web and Grid Services*, vol. 14, no. 4, pp. 352–375, 2018. DOI: 10.1504/IJWGS.2018.095647.

VISUALIZATION ARTICLES

- [V1] M. Ahmed, I. Shumailov, and R. Anderson, "Tendrils of crime: Visualizing the diffusion of stolen bitcoins," 2019. arXiv: 1901.01769 [cs.CY].
- [V2] S. Bistarelli and F. Santini, "Go with the-bitcoin-flow, with visual analytics," in *Proc. Conference on Availability, Reliability and Security*, ACM, 2017, p. 38. DOI: 10.1145/3098954.3098972.
- [V3] A. Bogner, "Seeing is understanding: Anomaly detection in blockchains with visualized features," in *Proc. Joint Conference on Pervasive and Ubiquitous Computing and Symposium on Wearable Computers (UbiComp/ISWC)*, ACM, 2017, pp. 5–8. DOI: 10.1145/3123024.3123157.
- [V4] S. Chawathe, "Monitoring blockchains with self-organizing maps," in *Proc. Conference on Trust, Security and Privacy In Computing and Communications and Conference on Big Data Science And Engineering (TrustCom/BigDataSE)*, IEEE, 2018, pp. 1870–1875. DOI: 10.1109/TrustCom/BigDataSE.2018.00283.
- [V5] G. Di Battista, V. Di Donato, M. Patrignani, M. Pizzonia, V. Roselli, and R. Tamassia, "Bitcoveview: Visualization of flows in the bitcoin transaction graph," in *Proc. Symposium on Visualization for Cyber Security (VizSec)*, IEEE, 2015, pp. 1–8. DOI: 10.1109/VIZSEC.2015.7312773.
- [V6] P. Isenberg, C. Kinkeldey, and J.-D. Fekete, "Exploring entity behavior on the bitcoin blockchain," in *Poster Proc. Conference on Information Visualization (InfoVis)*, 2017. [Online]. Available: <https://hal.inria.fr/hal-01658500>.
- [V7] C. Kinkeldey, J.-D. Fekete, T. Blascheck, and P. Isenberg, "Visualizing and analyzing entity activity on the bitcoin network," 2019. arXiv: 1912.08101 [cs.HC].
- [V8] H. Kuzuno and C. Karam, "Blockchain explorer: An analytical process and investigation environment for bitcoin," in *Proc. Symposium on Electronic Crime Research (eCrime)*, IEEE, 2017, pp. 9–16. DOI: 10.1109/ECRIME.2017.7945049.
- [V9] D. McGinn, D. Birch, D. Akroyd, M. Molina-Solana, Y. Guo, and W. J. Knottenbelt, "Visualizing dynamic bitcoin transaction patterns," *Big Data*, vol. 4, no. 2, pp. 109–119, 2016. DOI: 10.1089/big.2015.0056.
- [V10] R. Norvill, B. B. F. Pontiveros, R. State, and A. Cullen, "Visual emulation for ethereum's virtual machine," in *Proc. Network Operations and Management Symposium*, IEEE, 2018, pp. 1–4. DOI: 10.1109/NOMS.2018.8406332.
- [V11] F. Oggier, S. Phetsouvanh, and A. Datta, "Biva: Bitcoin network visualization & analysis," in *Proc. Conference on Data Mining Workshops (ICDMW)*, IEEE, 2018, pp. 1469–1474. DOI: 10.1109/ICDMW.2018.00210.
- [V12] P. Schretlen, N. Kronenfeld, D. Gray, J. McGeachie, E. Hall, D. Cheng, N. Covello, and W. Wright, "Interactive data exploration with 'big data tukey plots'," in *Poster Proc. Conference on Visualization (VIS)*, 2013. [Online]. Available: <https://uncharted.software/assets/interactive-data-exploration.pdf>.

- [V13] Y. Sun, H. Xiong, S. M. Yiu, and K. Y. Lam, "Bitvis: An interactive visualization system for bitcoin accounts analysis," in *Proc. Crypto Valley Conference on Blockchain Technology (CVCBT)*, IEEE, 2019, pp. 21–25. DOI: 10.1109/CVCBT.2019.000-3.
- [V14] X. Yue, X. Shu, X. Zhu, X. Du, Z. Yu, D. Papadopoulos, and S. Liu, "Bitextract: Interactive visualization for extracting bitcoin exchange intelligence," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 25, no. 1, pp. 162–171, 2018. DOI: 10.1109/TVCG.2018.2864814.

DATA ANALYSIS ARTICLES

- [A1] I. Alqassem, I. Rahwan, and D. Svetinovic, "The anti-social system properties: Bitcoin network data analysis," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018. DOI: 10.1109/TSMC.2018.2883678.
- [A2] E. H. Aw, R. Gera, K. Hicks, N. Koeppen, and C. Teska, "Analyzing preferential attachment in peer-to-peer bitcoin networks," in *Proc. Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, IEEE, 2018, pp. 1242–1249. DOI: 10.1109/ASONAM.2018.8508273.
- [A3] A. I. Badev and M. Chen, "Bitcoin: Technical background and data analysis," FEDS Working Paper, 2014. DOI: 10.2139/ssrn.2544331.
- [A4] T. de Balthasar and J. Hernandez-Castro, "An analysis of bitcoin laundry services," in *Proc. Nordic Conference on Secure IT Systems*, Springer, 2017, pp. 297–312. DOI: 10.1007/978-3-319-70290-2.
- [A5] M. Bartoletti and L. Pompianu, "An analysis of bitcoin op_return metadata," in *Proc. Conference on Financial Cryptography and Data Security*, Springer, 2017, pp. 218–230. DOI: 10.1007/978-3-319-70278-0_14.
- [A6] S. Bistarelli, M. Parrocchini, and F. Santini, "Visualizing bitcoin flows of ransomware: Wannacry one week later," in *Proc. Italian Conference on Cybersecurity (ITASEC)*, 2018. DOI: 10.1080/13567888.2017.1335101.
- [A7] T.-H. Chang and D. Svetinovic, "Data analysis of digital currency networks: Namecoin case study," in *Proc. Conference on Engineering of Complex Computer Systems (ICECCS)*, IEEE, 2016, pp. 122–125. DOI: 10.1109/ICECCS.2016.023.
- [A8] T. Chen, Y. Zhu, Z. Li, J. Chen, X. Li, X. Luo, X. Lin, and X. Zhang, "Understanding ethereum via graph analysis," in *Proc. Conference on Computer Communications (INFOCOMM)*, IEEE, 2018, pp. 1484–1492. DOI: 10.1109/INFOCOM.2018.8486401.
- [A9] G. Di Battista, V. Di Donato, and M. Pizzonia, "Long transaction chains and the bitcoin heartbeat," in *EuroPar 2017: Parallel Processing Workshops*, Springer, 2018, pp. 507–516. DOI: 10.1007/978-3-319-75178-8_41.
- [A10] M. Lischke and B. Fabian, "Analyzing the bitcoin network: The first four years," *Future Internet*, vol. 8, no. 1, p. 7, 2016. DOI: 10.3390/FI8010007.
- [A11] D. D. F. Maesa, A. Marino, and L. Ricci, "Uncovering the bitcoin blockchain: An analysis of the full users graph," in *Proc. Conference on Data Science and Advanced Analytics (DSAA)*, IEEE, 2016, pp. 537–546. DOI: 10.1109/DSAA.2016.52.
- [A12] D. McGinn, D. McIlwraith, and Y. Guo, "Towards open data blockchain analytics: A bitcoin perspective," *Royal Society Open Science*, vol. 5, no. 8, p. 180298, 2018. DOI: 10.1098/rsos.180298.
- [A13] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage, "A fistful of bitcoins: Characterizing payments among men with no names," in *Proc. Internet Measurement Conference*, ACM, 2013, pp. 127–140. DOI: 10.1145/2504730.2504747.

- [A14] L. Norbutas, "Offline constraints in online drug marketplaces: An exploratory analysis of a cryptomarket trade network," *International Journal of Drug Policy*, vol. 56, pp. 92–100, 2018. DOI: 10.1016/j.drugpo.2018.03.016.
- [A15] F. Parino, M. G. Beiró, and L. Gauvin, "Analysis of the bitcoin blockchain: Socio-economic factors behind the adoption," *EPJ Data Science*, vol. 7, no. 1, p. 38, 2018. DOI: 10.1140/EPJDS/S13688-018-0170-8.
- [A16] F. Reid and M. Harrigan, "An analysis of anonymity in the bitcoin system," in *Security and Privacy in Social Networks*, Springer, 2013, pp. 197–223. DOI: 10.1109/PASSAT/SocialCom.2011.79.
- [A17] D. Ron and A. Shamir, "Quantitative analysis of the full bitcoin transaction graph," in *Proc. Conference on Financial Cryptography and Data Security*, Springer, 2013, pp. 6–24. DOI: 10.1007/978-3-642-39884-1_2.
- [O24] Everex, *Ethplorer*, Accessed: August, 2019. [Online]. Available: <https://ethplorer.io/>.
- [O25] P. D. Filippi, *Plantoids: Blockchain-based life forms*, Accessed: August, 2019. [Online]. Available: <http://plantoid.org/>.
- [O26] GasTracker.io, *Gastracker.io*, Accessed: August, 2019. [Online]. Available: <https://gastracker.io/>.
- [O27] L. Hendriks, *Bitcoin Transaction Visualization*, Accessed: August, 2019. [Online]. Available: <http://bitcoin.interaqt.nl/>.
- [O28] C. Hines and G. Akerman, *Federal Bitcoin*, Accessed: August, 2019. [Online]. Available: <https://federalbitcoin.herokuapp.com/>.
- [O29] J. Hoenicke, *Johoe's Bitcoin Mempool Statistics*, Accessed: August, 2019. [Online]. Available: <https://jochen-hoenicke.de/queue/>.
- [O30] Input Output HK Limited, *Symphony of Blockchains: 3D Blockchain Explorer*, Accessed: August, 2019. [Online]. Available: <https://symphony.iohk.io/>.

ONLINE SOURCES

- [O01] Aleth.io, *EthStats.io*, Accessed: August, 2019. [Online]. Available: <https://ethstats.io/>.
- [O02] —, *Alethio*, Accessed: August, 2019. [Online]. Available: <https://aleth.io/>.
- [O03] BitBonkers, *BitBonkers - A Bitcoin Blockchain Transaction Visualisation*, Accessed: August, 2019. [Online]. Available: <https://bitbonkers.com/>.
- [O04] Bitcoin Globe, *Bitcoin Globe*, Accessed: August, 2019. [Online]. Available: <http://bitcoinglobe.com/>.
- [O05] BitcoinWisdom.com, *BitcoinWisdom*, Accessed: August, 2019. [Online]. Available: <https://bitcoinwisdom.com/>.
- [O06] Bitfly gmbh, *Etherchain*, Accessed: August, 2019. [Online]. Available: <https://www.etherchain.org/>.
- [O07] bitFlyer, *chainFlyer*, Accessed: August, 2019. [Online]. Available: <https://chainflyer.bitflyer.jp/>.
- [O08] BitForce5, *BitForce5*, Accessed: August, 2019. [Online]. Available: <http://bitforce5.com/>.
- [O09] BitInfoCharts, *BitInfoCharts*, Accessed: August, 2019. [Online]. Available: <https://bitinfocharts.com/>.
- [O10] Blockchain Luxembourg S.A., *Blockchain Explorer*, Accessed: August, 2019. [Online]. Available: <https://www.blockchain.com/explorer>.
- [O11] Blockchair, *Blockchair*, Accessed: August, 2019. [Online]. Available: <https://blockchair.com/>.
- [O12] BlockSeer, *BlockSeer*, Accessed: August, 2019. [Online]. Available: <https://www.blockseer.com/>.
- [O13] BTC.com, *BTC.com*, Accessed: August, 2019. [Online]. Available: <https://btc.com/>.
- [O14] K. Cieřla, *Bitcoinity*, Accessed: August, 2019. [Online]. Available: <https://data.bitcoinity.org/>.
- [O15] Coin Dance, *Coin Dance*, Accessed: August, 2019. [Online]. Available: <https://coin.dance/>.
- [O16] CoinDesk, Inc., *CoinDesk*, Accessed: August, 2019. [Online]. Available: <https://www.coindesk.com/data>.
- [O17] DailyBlockchain, *DailyBlockchain*, Accessed: August, 2019. [Online]. Available: <https://dailyblockchain.github.io/>.
- [O18] DashRadar, *DashRadar*, Accessed: August, 2019. [Online]. Available: <https://dashradar.com/>.
- [O19] Elliptic Enterprises Limited, *The Bitcoin Big Bang*, Accessed: August, 2019. [Online]. Available: <https://info.elliptic.co/hubfs/big-bang/bigbang-v1.html>.
- [O20] ethernodes.org, *The Ethereum Nodes Explorer*, Accessed: August, 2019. [Online]. Available: <https://www.ethernodes.org/>.
- [O21] Etherscan, *Etherscan*, Accessed: August, 2019. [Online]. Available: <https://etherscan.io/>.
- [O22] EtherView, *EtherView*, Accessed: August, 2019. [Online]. Available: <https://etherview.now.sh/>.
- [O23] Ethviewer, *Ethviewer*, Accessed: August, 2019. [Online]. Available: <http://www.ethviewer.live/>.
- [O31] Katana Cryptographic Ltd., *OXT*, Accessed: August, 2019. [Online]. Available: <https://oxt.me/>.
- [O32] E. Kerstein, *Bitcoin Visuals*, Accessed: August, 2019. [Online]. Available: <https://bitcoinvisuals.com/>.
- [O33] M. Laumeister, *BitListen*, Accessed: August, 2019. [Online]. Available: <https://www.bitlisten.com/>.
- [O34] P. Marrucho, *BitcoinCity*, Accessed: August, 2019. [Online]. Available: <http://bitcoincity.info/>.
- [O35] M. Oancea, *Ethereum Network Status*, Accessed: August, 2019. [Online]. Available: <https://ethstats.net/>.
- [O36] K. Small, *Blockchain 3D Explorer*, Accessed: August, 2019. [Online]. Available: <https://blockchain3d.info/>.
- [O37] Satoshi.info, *Statoshi.info*, Accessed: August, 2019. [Online]. Available: <https://satoshi.info/>.
- [O38] D. Stupp, *ON BRINK: Live Physicalization of the Bitcoin Blockchain*, Accessed: August, 2019. [Online]. Available: <http://dataphys.org/list/on-brink-live-physicalization-of-the-bitcoin-blockchain/>.
- [O39] TradeBlock, *TradeBlock*, Accessed: August, 2019. [Online]. Available: <https://tradeblock.com/bitcoin/>.
- [O40] TX Highway, *TX Highway*, Accessed: August, 2019. [Online]. Available: <https://txhighway.com/>.
- [O41] J. Vornberger, *Bitcoin Monitor*, Accessed: August, 2019. [Online]. Available: <https://www.bitcoinmonitor.com/>.
- [O42] G. Walker, *Bitcoinrain*, Accessed: August, 2019. [Online]. Available: <http://bitcoincrain.io/>.
- [O43] R. K. Wilson, *Bitcoin-VR*, Accessed: August, 2019. [Online]. Available: <https://bitcoin-vr.github.io/>.
- [O44] Wizbit, *Wizbit Realtime Bitcoin Globe*, Accessed: August, 2019. [Online]. Available: <https://blocks.wizb.it/>.
- [O45] A. Yeow, *Bitnodes*, Accessed: August, 2019. [Online]. Available: <https://bitnodes.earn.com/>.



Natkamon Tovanich is a Ph.D. student at Institute for Technological Research SystemX and Université Paris-Saclay. He is also a member of the Aviz research team at Inria, France. He received the Erasmus Mundus joint master's degree in Information Technologies for Business Intelligence from Université Libre de Bruxelles, Belgium, Université de Tours, France, and CentraleSupélec, France in 2018. His main research areas are information visualization and visual analytics with a focus on blockchain data.



Nicolas Heulot is a research engineer at Institute for Technological Research SystemX, France. He received the Ph.D. in Computer Science from University of Paris-Sud, France in 2014. His main research areas are Visual Analytics and Blockchain.



Jean-Daniel Fekete is the scientific leader of the Inria project team Aviz that he founded in 2007. He received the Ph.D. in Computer Science in 1996 from University of Paris-Sud, France, joined Inria in 2002 as a confirmed researcher, and became senior research scientist in 2006. His main research areas are Visual Analytics, Information Visualization and Human-Computer Interaction. He is a Senior Member of IEEE.



Petra Isenberg is a research scientist at Inria, France in the Aviz team. She received her PhD in Computer Science from the University of Calgary in 2010. Her main research areas are information visualization and visual analytics with a focus on collaborative work scenarios, interaction, and evaluation.