Imagining a crypto world

Exploring the way the value of crypto is imagined by YouTube “fin”-fluencer

Team Members

Gian Oerlemans - gian.oerlemans@student.uva.nl
Martin Trans - martin.trans@student.uva.nl
Luca van Leer- luca.van.leer@student.uva.nl
Wian Slendebroek - wwslendebroek@gmail.com
Tristan Bannerman - tristanwbannerman@gmail.com
Amber Wiznitzer - amberwiz@hotmail.com
Kobe De Keere - k.dekeere@uva.nl
Stefania Milan - s.milan@uva.nl

Contents

Team Members 1

Contents 1

Summary of Key Findings 2

1. Introduction 2

2. Initial Data Sets 3

3. Research Questions 3

4. Methodology 3

5. Findings 6

6. Discussion 11

7. Conclusion 12

8. References 13

Summary of Key Findings

Through an abductive combination of qualitative and quantitative analysis of YouTube videos we able to discern the dominance of four main imaginaries that culturally shape the motivations of crypto owners: cryptopia, institutionalizaion, bulls and bears and crypto-lifestyle. Next to that were able to see how the rise and fall of certain imaginaries (between 2017 and 2023) responded to important events within the field of crypto (e.g. market capitalization value). Finally, we managed to distinguish two important dimensions that determine the variations of field positions: inside vs. outside and currency vs. asset.

1. Introduction

According to the Cambridge Centre for Alternative Finance, there were 101 million cryptocurrency users worldwide in 2020, making 300,000 transactions a day (Schär & Berentsen, 2020). All cryptocurrency combined have, by now, a market capitalization value that easily surpasses a trillion USD (see coinmarketcap.com). However, little is known to date about what makes people engage in this new type of currency or financial asset, and how they imagine the new digital economy and the effects it can have on our social structure. We are hard pressed to find answers in the existing literature, which largely focuses on cryptocurrency's technological architecture and effects on finance. But when it comes down to the crypto users or owners, little systematic empirical research has been done that offers insight in their motivations, aspirations, and imaginaries.

Originally set up as a trust-free system to digitally bypass the political manipulation of money, many doubt whether cryptocurrency will ever fulfil its role as money due to its "promissory gap" (Dodd 2017, 2018). Instead of "neutral" money, cryptocurrency rests on a vibrant ideological community and is marked by strong centralizing power dynamics (e.g., unequal distribution of coins, mining resources and access) (Bjerg, 2016). So rather than an apolitical project, the field of cryptocurrency is marked by strong ideological and moral engagement. Although a small number of studies are based on empirical data (Caliskan, 2020; Shaw, 2021), most research on crypto ownership involves exploring theoretical premises while relying only on merely anecdotal evidence. This project now wants to take an important first step in systematically collecting in-depth empirical data on the cultures—in particular imaginaries and moral world views—of crypto users. For this purpose we have decided to analyze popular YouTube channels to uncover which types of imaginaries they use and diffuse that can shape the motivation of crypto owners. Entering the field of crypto demands a quite steep learning process, hence field novices rely strongly on YouTube video tutorial to get a

first understanding. This make the study of YouTube content as excellent case to get a first understanding of the most dominant crypto imaginaries and how the have changed over time.

2. Initial Data Sets

The project took off without a preliminary dataset. The project organizers brought embedded knowledge of the field and had lists of relevant accounts on various social media platforms. This served as a first step in the data gathering process. This project is still in the initial phase and therefore the winter school project is set up as proof-of-concept. The goal is to find out which data is suitable for collection and analysis.

Videochannel corpus*:* functioned as seed list for our research https://docs.google.com/document/d/1JVwlXNCantbBY66lySSFXhxPQdBaFofpX99-kPP2io8/edit?usp=share_link

Video list: All videos on these channels and their transcripts respectively https://docs.google.com/spreadsheets/d/1iJ1cza4FfTvsOGrDEweBwUdNWCPSPpSF9SztFWijJAI/edit?usp=share_link

Coindata*:* Historical market data for pricing, market cap, total number of coins and performance: https://docs.google.com/spreadsheets/d/1Ph4bAethl01TCx80ShZo2WShwlFF8IIxyc4GPQTxB_k/edit?usp=share_link

3. Research Questions

How is crypto imagined as valuable?

1) Which are the imaginaries popular within the field of crypto?
2) How does the popularity of certain imaginaries fluctuate over time?
3) Does the prevalence of certain imaginaries relate to specific historical events?
4) Which boundaries between field actors are drawn based on these imaginaries?

4. Methodology

Abductive principles of analysis (Tavory and Timmermans, 2014) informed our analysis. This means that we started with some preconceptions about the world of crypto and the possible imaginaries but wanted to, reiteratively, shape and reshape our knowledge

about it. More concretely it means that analysis of the data happened in several rounds. Seeing that an abductive analysis implies a double-checking of our inferences. Central to this approach is that we had both qualitative as well as quantitative rounds of coding and interpreting. This proof-of-concept project was unfolded in four stages:

Stage 1

Based on a seed list consisting of 16 YouTube channels (most popular channels that have crypto/bitcoin/ digital coins in their description with a minimum of 100k followers), we selected the latest posted video for a collective (4 independent coders) inductive coding. The goal of this round was to determine the most important imaginaries. This was done using Atlas.ti.

Concurrently, a spreadsheet was compiled of the publicly available videos posted by aforementioned channels by leveraging the Video List Module in YouTube Data Tools (Rieder, 2015). This list comprised entries for 19,607 videos, and included attributes such as unique video-IDs, titles, descriptions, date of release, duration, and public metrics of views, like-counts and comments. These were used to calculate a total engagement metric used for subsequent subsampling alongside a selection heuristic per channel.

Stage 2

Based on the qualitative analyses and group deliberation we determined the four most important imaginaries (see Findings sections) and the most important positions within the field of crypto (based on two axes, see Findings section). Finally, we also decided on the most important keywords, per imaginary, that will be employed during the quantitative stage 3.

Explorative Natural Language Processing (NLP) techniques were applied with the use of 4CAT (Peeters & Hagen, 2022) to derivative datasets consisting of the video titles and their respective timestamps. According to the abductive principles of analysis we decided to focus on representative subsampling and full video transcripts rather than video titles and descriptions as these were deemed too susceptible to practices of 'click-baiting' and Search Engine Optimization (SEO). Descriptions were generally repeating the same phrases and keywords which introduced avoidable bias to our corpus.

Stage 3

The objective of this stage is to check the prevalence of the imaginaries and see whether their salience fluctuates over time, in relation to certain events. The most engaged-with video for every other month per channel were selected for our subsample in working with captions. We further trimmed away videos in our subsample released before 2017 and after 2023, as not all channels were active outside those points in time, avoiding temporal bias. The subsample consisted on average of 60 videos per channel with a total of 434 videos. Transcripts were available for 402 videos, with the remaining 32 videos having the feature explicitly disabled by the channel creator.

We created a dataset of transcripts of the subsample by compiling a comma-separated text-file from the 'videoID'-column, which was fed into youtube-DL (ytdl-org, n.d.) Command Line Interface software which allowed us to fetch transcripts only, meaning that we did not have to store the videos, saving bandwidth and local storage. YoutubeDL was executed with the compiled txt-list as its input with the following flags: automatically write subtitles, English subtitles, subtitle format as vtt, skipping download of the video itself, title of each file as the corresponding video ID and language:

youtube-dl --write-auto-sub --sub-lang en --skip-download --sub-format vtt --id --batch-file batch-list-test.txt

This execution needed to be re-run every 40 seconds or so to avoid rate-limiting, and the compiled list needed parallel curation to avoid re-fetching the same video transcripts over and over. This workflow can most likely be optimized for further work. This resulted in a series of VTT files, which is not immediately compatible with our intended use of the data, as it contains information on time and context, where we ideally want clean text on which we can apply NLP techniques. To convert these files, retaining only the actual transcript words uttered in the videos, we adapted an existing python-script[1] to our needs, utilizing regular expressions to produce text-only txt-files of each video's transcript with only the video ID as its new title. The content of these text-files were then populated into a new column in their corresponding rows, matched through the video IDs, in an ad-hoc python notebook.

Important note: relying (mostly, as we can only assume that each creator has not uploaded their own custom subtitles) on Google's internal transcript engines resulting in certain words not being picked up properly if they're not acknowledged as such individual words. For example words like defi (short decentralized finance) was often transcribed as d5 or defy, dfy. This had an effect on the word search algorithms we used and asks for additional data cleaning.

[1] https://gist.github.com/glasslion/b2fcad16bc8a9630dbd7a945ab5ebf5e

Stage 4

In this final stage we focused on visualizing the data and the findings. For this we opted for a timeline approach that diachronically represents the sequence of events between 2017 and now. By plotting market cap value and popularity of imaginaries (based word counts and N-grams).

The newly populated CSV-files containing each video's transcript were imported as a custom data source in 4CAT, and a series of NLP techniques were applied to the dataset. The quantitative NLP-generated visualizations informed our qualitative analysis and vice versa in a dialectic manner, as we experimented with different combinations of stopwords and subsamples, across various applications of tokenization[2], lemmatization (i.e. converting words to their root-versions, e.g. 'running' becomes 'run') and time-delimiting (i.e. producing documents per year vs per month) that brought different aspects of our dataset to the foreground. We finally produced a rankflow (Rieder, n.d.) from a series of word-counts divided per year as described above.

5. Findings

Four imaginaries

There are four overarching imaginaries that are very dominant within the corpus. These imaginaries, that can be found in most of the analyzed videos, are as follows:

1. Cryptopia (label: cryptopia)

Although there are different types to be distinguished (as will be discussed in the next paragraph), crypto users do form a community of their own, as most of them share certain ideological views on the market and on society. They have a strong belief that decentralized finance, or DeFi - a financial technology that promotes the use of peer-to-peer transactions - has the future (Bjer, 2016; Dodd, 2017, 2018, Maurer et al 2013)). DeFi is meant to challenge the current centralized financial system by taking out the middleman and eliminating the fees that financial companies (such as banks) charge for

[2] https://www.nltk.org/api/nltk.tokenize.html

using their services. This way, crypto currency should offer a certain independence. This is essential for the crypto community for two reasons. First of all, crypto users see the independence that crypto seemingly offers as a way to fight global inequality and 'bank the unbanked', enabling those who (for whatever reason) do not own a bank account to transfer money anyway. And secondly, many crypto users share a common distrust of central authorities. Crypto is considered as a guarantee that when the debt society, that these authorities supposedly created, finally comes crumbling down, crypto users will not be impacted by this societal collapse: after all, they are not dependent on the services of financial companies (Golumbia, 2016). The crypto community considers this eventual collapse caused by the 'old' economy to be inevitable, which strengthens their belief in crypto as the future.

2. Between bulls and bears (llabel: bulls and bears)

Lots of analysis about the crypto market are about 'bulls' and 'bears'. A bull market is when market conditions are favorable for investments, for example when users expect the prices to go up soon. A bear market is the complete opposite, when market conditions are not favorable and a crash of the cryptocurrency is possibly about to happen. The four types of crypto-investors we are about to introduce are looking at this quite differently. The first group are the day-traders. Day-trading is a short-term strategy for investors. They buy and sell their cryptocurrencies on one day with the aim of making money in a quick manner. So their goal is to profit from changing prices throughout the day, and thus this group makes use of the bull markets a lot. The second group contains long-term investors. They are planning on using crypto for a longer period of time, typically for many years. It's a form of passive income for them, and that's why the fluctuations in the cryptomarkets don't do much to this group. The third group are the HODLers (Hold On for Dear Life). They will never sell their cryptocurrencies, whatever happens to the markets. This fits very well into the ideological view, based on decentralized finance and an alternative economy, as explained earlier. It's only possible to say whether a person is a long-term investor or a HODLer after they stop investing, or when they speak out about it. The fourth group are the 'whales'. These people own much of a certain cryptocurrency, which causes them to have great influence on the market price when they come or leave, and therefore for them to be able to influence a bull or bear market.

3. Institutionalization: only when it's in our interest (label: institutionalization)

Crypto-investors and content makers have an ambiguous relationship with institutionalization (Hates, 2019). On the one hand, they don't like any form of regulation and institutionalization which can harm their interests and the interests of

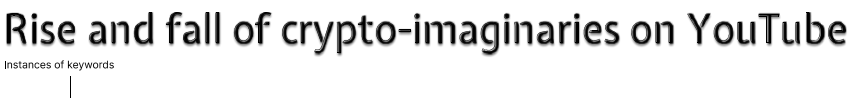
cryptocurrencies. In practice this means they are critical towards central banks (like the ECB and the FED), the debt system, the Consumer Price Index (CPI; inflation), regulation and the enforcement of it, taxes and the law. On the other hand, when there is form of institutionalizatidon which can protect them and their interests, they don't hesitate to use it. For example, when a bank wants to use a certain cryptocurrency, it creates a social safety net in case there's a crash. An example of such a bank is the American bank Silvergate, with FTX as their main (former) client (it's still not totally clear what the FTX crash means for Silvergate, but the collapse definitely did have its impact).

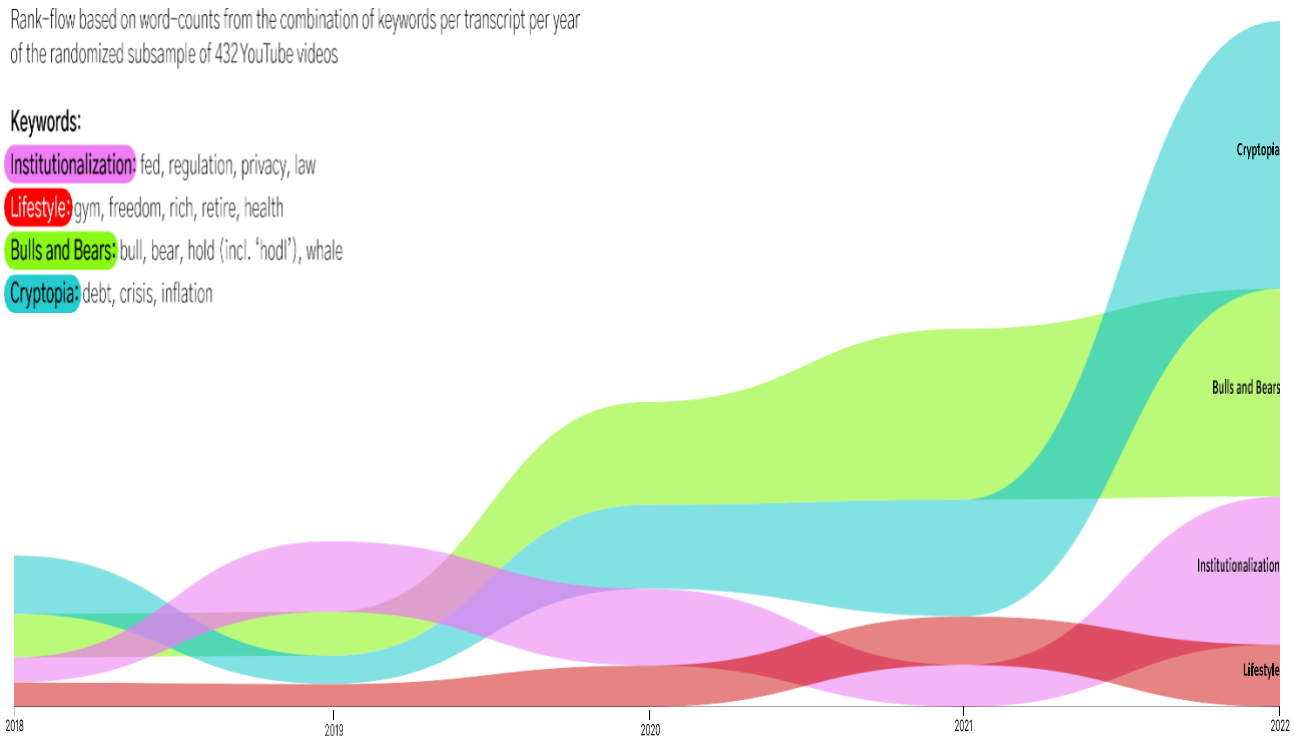
4. The grind never stops (label: lifestyle)

An interesting aspect of the crypto community is visible on YouTube. The community of subscribers and viewers see the content makers as life coaches. The content makers encourage their followers to be fit and live a healthy life. Their viewers are often gamers that want to get rich quick as digital nomads (= someone who earns money online, independent from their physical location). In this way they try to be financially independent, which they view as financial freedom. As a consequence, they can retire early.

4 Main Imaginaries and their keywords:

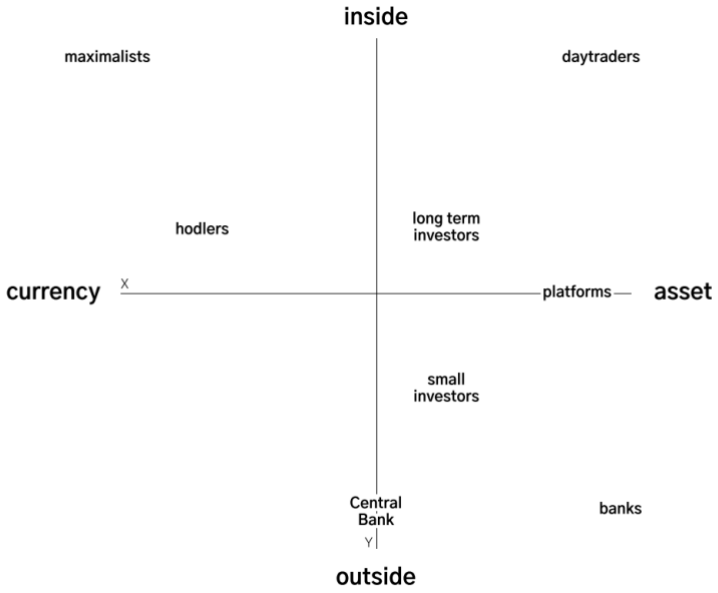
1. **Institutionalization**: fed, regulation, privacy, law
2. **Lifestyle:** gym, freedom, rich, retire, health
3. **Bulls and Bears:** bull, bear, hold (incl. 'hodl'), whale
4. **Cryptopia:** debt, crisis, inflation





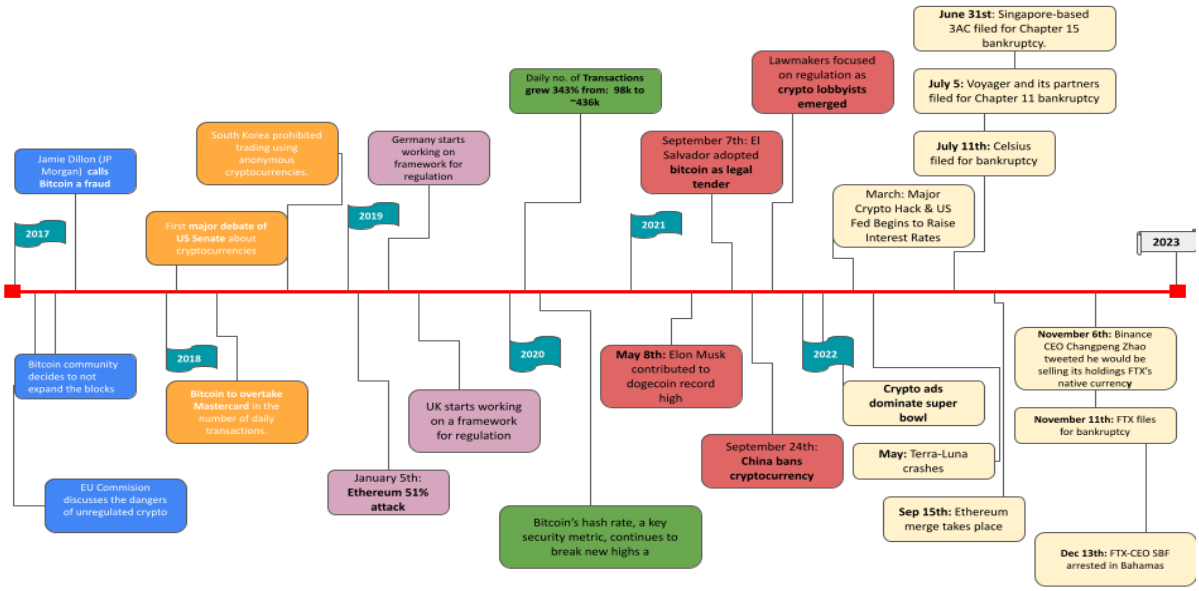
Field positions

Through in-depth analysis of the 16 most recent published videos and a collective deliberation about the content and codes we eventually came up with classification of different crypto actors based on two field dimensions. The first dimensions refers to how central the actors are to field (inside or outside the field). The second dimensions entails the distinction between those who see crypto as potential new type of money or currency. For them, cryptocurrency is more than merely an interesting investment. On the other side of the spectrum we find those who see crypto coins primarily as digital asset, a new form of investment.



Time line

The timeline was constructed by gathering all major crypto headlines from that year. The timeline starts with the first large-scale bull run of crypto in 2017. Due to these bull runs crypto started breaking into the mainstream more often. Nevertheless, most information was retrieved from specific crypto reporting websites that solely focus on crypto news. The final cut of events in the timeline is reduced due to the prevention of becoming too granular. Furthermore, a large number of headlines for certain years became less interesting as they are not as impactful as events that transpired on a later date in a more mature market. Nevertheless, we have kept a decent selection of events that have directly impacted the trajectory of development for the crypto industry, such events include; regulation debate, large-scale hacks, country-wide bans, and major infrastructure upgrades.



6. Discussion

The research has allowed us to take a deep dive into the crypto community and understand what drives the different groups within the “fin-fluencers” on youtube. The main findings of this research have been the four main imaginaries *cryptopia, bulls and bears, institutionalization, and lifestyle*. These imaginaries demonstrate to what groups within the market attribute value to. People with an interest in investing in crypto will be met with these imaginaries if they were to do their primary research on youtube. Quickly they will have to decide which imaginaries speak to them the most and where on the matrix they would position themselves on the basis of what they seek in their purpose for crypto investing. By understanding the main dialogue within the market and which is most prevalent, we can roughly estimate where the market will head next. This has implications for future users of crypto and how their market will move. Moreover, by understanding the development of this dialogue over time in comparison to timeline events, market cap, and YoY growth of coins minted we roughly start to recognize patterns as to what are the biggest drivers in this market.

In general, we found that there was a shift in the way that the cryptocurrency YouTubers talked about cryptocurrency, from using language that we found to be based on short term trading to using language that showed a more ideological and long term understanding of the technology. This shift we saw correlated to the cryptocurrency market capitalization. As the cryptocurrency market capitalization grew so too did the number of cryptocurrencies this economic reality drove the language and the imaginaries of the crypto YouTubers. A shift of this nature highlights the difficulty that

has been affecting the cryptocurrency world and showcases how a group continues to proselytize and save face by changing the imaginaries that they are engaging with. Moving from the realm of trading and short term gain to an imaginary of commitment to the technology is a large shift and makes it clear that the declining value of cryptocurrency is being felt heavily by the community in ways that go beyond their wallet.
These findings are interesting for a number of reasons. First, we are able to understand more about this very large and disparate crypto-community that has not been adequately researched despite its outsized influence economically, technologically, and culturally. Second, these findings make it clear to us that there is much more work to be done in this realm. This project makes it clear that there is much more for us to uncover using our methods and approach and that scaling up will yield better results. Third, this research is significant because we can view this project as a proof of concept of the research as a whole and of our methods that we used. With that in mind this gives other researchers tools to conduct similar research on communities on YouTube.

7. Conclusion

In summation our research makes it clear that there is much more work to be done to properly assess the imaginaries that color the cryptocurrency community. As stated before this project proves that there is a viable path to do this type of research on YouTube videos and that for us to continue this work we should scale up. With regards to our findings we know that there has been a shift in the realm of cryptocurrency imaginaries since the year 2017. This shift is to be expected given the growth and turmoil that has affected the world of cryptocurrency (and the world at large) but it would be interesting to look at the trends starting from 2009. Although it has been limited to just one platform, this could act to be a proof of concept for understanding crypto market dynamics based on its community if performed on a variety of platforms. The community remains a key part of crypto-market drivers

The matrix that we created to help understand the crypto community imaginaries is also a useful tool that could be included in future studies of cryptocurrency imaginaries. The research has led to our discovery of four main imaginaries. These Imaginaries give an insight into how the fin-fluencers community attribute value to crypto. We have found a distinct difference in what appeals to certain groups of people to this market.

- *The maximalits* who believe in a cryptopia that would replace our debt-markets that we currently use. This group is a fundamental believer of DeFi replacing our current traditional markets.
- *The bulls and bears* with the majority of bulls being formed by day traders. The bears are formed mostly by HODLers who lean more toward a cryptopian belief

of the cryptoindustry. Key aspect of a HODLers is that they will never sell until their position unless leaving the market for good.

As mentioned before this research does have its limitations with regard to maintaining a feasible timeframe for conducting it during the Winter School. This has led to results being limited to only one platform only. However, as mentioned before this could act as a proof of concept to better understand the community behind crypto. For future practices new digital methods tools for data collection on other platforms such as Discord can be useful. Crypto-communities like to migrate over time to always stay ahead and therefore the tools required to conduct research should adapt accordingly.

One benefit crypto-researchers will have compared to traditional financial markets is that everything is open-source due to the public nature of distributed ledgers. Therefore not relying on exchange data reserved for institutional traders only. Therefore I am convinced that future research could extend on what has been established in this report.

Finally, we are proud of the work we have done and look forward to carrying it on.

8. References

Bjerg, O. (2016). How is Bitcoin Money*? Theory, Culture & Society* 33(1): 53–72.

Çalışkan, K. (2020a). Platform Works as Stack Economization: Cryptocurrency Markets and Exchanges in Perspective. *Sociologica* 14(3): 115–142.

Çalışkan, K. (2020b) Data Money: The Socio-technical Infrastructure of Cryptocurrency Blockchains. *Economy and Society* 49(4): 540–561.

de Nooy, W. (2003). Fields and Networks: Correspondence Analysis and Social Network Analysis in the Framework of Field Theory. *Poetics* 31(5): 305–327.

Dodd, N. (2017) Utopian Monies: Complementary Currencies, Bitcoin, and the Social Life of Money. In: *Utopian Monies: Complementary Currencies, Bitcoin, and the Social Life of Mone*y. Princeton: Princeton University Press.

Dodd, N. (2018). The Social Life of Bitcoin. *Theory, Culture & Society* 35(3): 35–56.

Golumbia, D. (2016). *The Politics of Bitcoin: Software as Right-Wing Extremism.* Minneapolis: University of Minnesota Press.

Hayes, A. (2019) The Socio-Technological Lives of Bitcoin. Theory, Culture & Society 36(4): 49–72.

Maurer, B., Nelms, T.C. & Swartz, L. (2013). "When Perhaps the Real Problem is Money Itself!": The Practical Materiality of Bitcoin. *Social Semiotics* 23(2): 261–277.

Peeters, S., & Hagen, S. (2022). The 4CAT Capture and Analysis Toolkit: A Modular Tool for Transparent and Traceable Social Media Research. *Computational Communication Research*, *4*(2), 571–589. https://doi.org/10.5117/CCR2022.2.007.HAGE

Schär, F. & Berentsen, A. (2020) *Bitcoin, Blockchain, and Cryptoassets: A Comprehensive Introduction*. Cambridge: MIT Press.

Schneider, N. (2019). Decentralization: An Incomplete Ambition. *Journal of Cultural Economy* 12(4): 265–285.

Shaw, L. (2021). The Inevitable Sociality of Money: The Primacy of Practical Affirmation over Conceptual Consensus in the Construction of Bitcoin's Economic Value. *Socio-Economic Review*, Online first: DOI: 10.1093/ser/mwab009

Rieder, B. (2015). *YouTube Data Tools* (1.3). https://tools.digitalmethods.net/netvizz/youtube/

Rieder, B. (n.d.). *RankFlow* [JavaScript]. Retrieved 30 January 2023, from https://github.com/bernorieder/RankFlow (Original work published 2015).

Tavory, I., & Timmermans, S. (2014). *Abductive analysis: Theorizing qualitative research*. University of Chicago Press.

ytdl-org. (n.d.). *Youtube-dl* [Python]. youtube-dl. Retrieved 30 January 2023, from https://github.com/ytdl-org/youtube-dl (Original work published 2010).