Exploring blockchain data by central banks: outcomes from a Decentralised Finance (DeFi) hackathon

Anita Deflorio (Banca d'Italia); Alexander Hodbod (European Central Bank); Mikaël Kalfa (European Central Bank); Urszula Kochanska (European Central Bank); Eleni Koutrouli (Bank of Greece); Jose Manuel Carbó (Banco de España); Pedro Silva (European Central Bank); Thomas Teulery (European Central Bank – formerly)

Disclaimer: The content of this case study reflects the work of hackathon participants and the views of the authors. The views expressed do not necessarily reflect those of the authors' affiliated institutions.

Acknowledgments: The authors would like to thank Isabel Kerner and Alice Filippeta for their comments. Any errors and omissions are the sole responsibility of the authors.

Abstract

Decentralised Finance (DeFi) seeks to emulate various financial services provided by the traditional financial system but without centralised intermediaries, relying instead on automated protocols. DeFi protocols have gained a significant user-base since 2020 though they remain relatively unknown, difficult to assess as regards code quality, unsecure, untested, and posing significant risks to consumers and investors. In principle, DeFi is subject to the same risks and vulnerabilities observed in traditional finance such as excessive leverage and risk taking, liquidity mismatches or interconnectedness. However, the way DeFi is designed may generate certain novel vulnerabilities, for example due to the high interconnectedness with crypto-asset markets as a result of DeFi's unique governance setup, or because of specific types of software-related operational risk.

Readily available public data on DeFi generally lack granularity and are prone to gaps, while data retrieved directly from blockchains remain opaque and cumbersome to analyse also due to the lack of generally accepted standards and heterogeneous approaches across protocols. In this context, analysing more detailed data and developing methods for risk analysis are measures considered beneficial by central banks, as they seek ways to better understand and monitor this new segment of the crypto-asset markets, to capture it in official statistics and analyse how it connects with traditional financial markets.

Against this backdrop, this paper offers a case study elaborating on the findings of the March 2023 DeFi hackathon organised by the ECB. The objective of the hackathon was to deepen understanding of specific DeFi segments by offering participants hands-on exposure to detailed data on DeFi protocols and enabling analyses of relevance to the interests and mandates of central banks and banking

supervisors. Obtaining insights directly from blockchain data, rather than relying on data-providers, can be seen as pushing the boundaries in crypto-asset monitoring. The case study is organised in four sections.

The first section covers challenges related to blockchain data preparation for the hackathon. Specifically, this part elaborates on the systems that decode and expose DeFi protocol data via web endpoints. This method of retrieving blockchain data was used for the hackathon and allowed participants to explore all the available information concerning two distinct DeFi segments.

The second section elaborates on the features of DeFi credit/lending and the outcome of the analysis of Aave protocol data. This analysis sheds light on the potential risks that counterparts are exposed to when entering into DeFi agreements. The indicators developed by the hackathon participants focus on gaining insights into borrowers, loan features (including flash loans), and deposit pools across various blockchains. Flash loans have become one of the most creative tools in the DeFi industry and are either atomically executed and repaid within one individual transaction or reverted, thus they circumvent the borrower default and platform liquidity risks.

The third section presents the DeFi payment segment, with an examination of Sablier protocol data. The insights explored encompass the usage of DeFi payments, which includes streaming, across various blockchains. Additionally, this section delves into challenging aspects such as the feasibility to determine a transaction purpose or its geographical attribution, as well as identification of vendors or other key players.

In the fourth section the focus shifts to blockchain oracles, which are third-party services that enable smart contracts within DeFi applications to receive external data from outside of their ecosystem. A number of incidents relating to malfunctioning oracles or to the exploitation of the specific characteristics of oracles and DeFi applications are presented and analysed.

The concluding section of the case study summarises key findings related to DeFi and blockchain data for central banking. Additionally, it evaluates the efficacy of hackathons in acquiring knowledge and honing skills.

Keywords: Decentralised Finance (DeFi), DeFi lending, DeFi payment, blockchain oracles, blockchain data, experimental statistics

Contents

1 Introduction and motivation

Decentralised Finance (DeFi) seeks to emulate various financial services provided by the traditional financial system but without centralised intermediaries, relying instead on automated protocols. DeFi has been arguably around since the launch of Bitcoin in 2009 but it truly took off with the introduction of smart contracts on the Ethereum blockchain[1], which expanded the functionalities beyond the simple transfer of value among users. DeFi protocols have gained a significant user-base since 2020, though they remain relatively unknown, complex, difficult to assess as regards code quality, unsecure[2], and untested posing significant risks to consumers and investors.

In principle, DeFi is subject to the same risks and vulnerabilities observed in traditional finance such as excessive leverage and risk taking, liquidity mismatches or interconnectedness. However, the way DeFi is designed may generate certain novel vulnerabilities, for example due to the high interconnectedness with crypto-asset markets as a result of DeFi's unique governance setup, or because of specific types of software-related operational risk.

Readily available public data on DeFi generally lack granularity and are prone to gaps, while data retrieved directly from blockchains remain opaque and cumbersome to analyse also due to the lack of generally accepted standards[3] and heterogeneous approaches across protocols[4]. In this context, analysing more detailed data and developing methods for risk analysis will be beneficial for central banks and banking supervisors as they seek ways to understand this new segment of the crypto-asset markets and how it connects with traditional financial markets.

Against this backdrop, this paper offers a case study elaborating on the findings of the March 2023 DeFi Hackathon[5] organised by the ECB. The objective of the DeFi Hackathon was to deepen the understanding of specific DeFi segments by offering participants hands-on exposure to detailed data on DeFi protocols and enabling analyses of relevance to the interests and mandates of central banks and banking supervisors. Obtaining insights directly from blockchain data, rather than relying on data-providers, can be seen as pushing the boundaries in crypto-asset monitoring. The case study is organised in four sections.

The first section covers challenges related to blockchain data preparation for the hackathon. Specifically, this part elaborates on the systems that decode and expose

---

[1] Ethereum blockchain went live in July 2015

[2] Despite the decrease in 2023, DeFi hacking continues to represent the largest source of funds stolen.

[3] There exist token standards for Ethereum Virtual Machine(EVM) compatible blockchains (ERC 20, 721 and 777). They do not cover other blockchains and refer only to tokens, not to other crypto-assets.

[4] On the perceived problem of heterogeneous approaches across protocols, see F. Boissay, G. Cornelli et al, "Blockchain scalability and the fragmentation of crypto", BIS bulletin no. 56, pag. 3-4

[5] A hackathon is an event that brings interested people together typically to accomplish typically a programming or analytical objective in a short period of time. The word hackathon is a portmanteau of the words hacker and marathon. The DeFi Hackathon was open to participants from the European System of Central Banks (ESCB) and the Single Supervisory Mechanism (SSM). It lasted 48 hours.

DeFi protocol data via web endpoints. This method of retrieving blockchain data was used in the hackathon and allowed participants to explore all the available information concerning two selected DeFi segments.

The second section elaborates on the features of DeFi credit/lending and the outcome of the analysis of Aave data. This analysis sheds light on the potential risks that counterparts are exposed to when entering into DeFi agreements. The indicators developed by the hackathon participants focus on gaining insights into borrowers, loans features (including flash loans), and deposit pools across various blockchains. Flash loans have become one of the most creative tools in the DeFi industry and are either atomically executed and repaid within one individual transaction or reverted, thus they circumvent the borrower default and platform liquidity risks (see Box 1).

The third section presents the DeFi payment segment, with an examination of Sablier[6] data. The insights explored encompass the usage of DeFi payments, which includes streaming, across various blockchains. Additionally, this section delves into challenging aspects such as the feasibility to determine a transaction purpose or its geographical attribution, as well as identification of vendors or other key players.

In the fourth section the focus shifts to blockchain oracles[7] which are third-party services that enable smart contracts within DeFi applications to receive external data from outside of their ecosystem. A number of incidents relating to malfunctioning oracles or to the exploitation of the specific characteristics of oracles and DeFi applications are presented and analysed. Given their relevance, blockchain oracles were used for the hackathon's challenge that was not related to data.

The concluding section of the case study summarises key findings related to DeFi and blockchain data for central banking. Additionally, it evaluates the efficacy of hackathons in acquiring knowledge and honing skills. The hackathon played a crucial role in fostering and strengthening collaboration between central banks on crypto-assets and DeFi monitoring, also motivating the creation of an Eurosystem group to better monitor crypto-assets.

---

6 At the time of the hackathon, Sablier was one of the biggest DeFi payment protocols in terms of TVL. It offers e.g. money streaming which encompasses open-ended continuous payments. Such payments can be used to establish a real time direct link between the value transfer and the service provision.

7 Oracles are decentralised middleware entities (intermediaries) that connect smart contracts to validated resources outside their native blockchains. Oracles are widely used in DeFi to provide e.g. price feeds. Oracles serve as bridges connecting any blockchain with data from both other blockchains as well as off-chain systems.

2 2. Data preparation

2.1 Data preparation for the hackathon

While blockchain data are openly accessible, they remain opaque and cumbersome to analyse directly. As a result, the developers of the DeFi protocols deployed systems that decode and expose protocol data via API endpoints[8]. For preparing the hackathon, the ECB used a blockchain indexer to extract and reformat the complete data from two protocols, distributed across several blockchains. This gave hackathon participants access to detailed data to explore and understand DeFi protocols for two distinct DeFi areas. The two protocols that were examined are the Aave lending protocol and the Sablier payments protocol. They were chosen as illustrative examples of the respective DeFi areas, as Aave has a significant market share in terms of total value locked (TVL) whereas Sablier, at the time of the hackathon, offered *streaming* - a novel payments mechanism which was considered worthy of exploration.

Main blockchains structure their data as Merkle Trees[9]. Each block is represented by one tree, containing all the transactions that occurred since the last block addition. At the root of the tree the block hash is stored. This hash is the output of a cryptographic function that takes as inputs the main identifying attributes of a block; thereby purposefully replacing transactions data with a hash[10]. Each block also stores the hash from the preceding block thereby creating a chain between all the blocks. Given the growing amount of data stored in blockchains, this technology provides for a secure and efficient method for data integrity and verification. Regarding the storage of and access to data, hashes serve as a useful lookup handle[11] item, to easily index and retrieve the granular data[12].

By definition, a transaction solely represents the transfer of value/ownership from one address to another. However, for blockchains supporting smart contracts, the same concept is used to interact with the protocols, storing the operation specifications in the body of the transaction alongside metadata, the transaction being a facilitating vessel only. For example, when users participate in asset management activities via a designated DeFi protocol or engage in a protocol's governance, they instigate actions through transactions; however, these transactions may not exclusively aim to facilitate fund transfers and alter token balances. Consequently, the significance of incorporating supplementary transaction metadata becomes apparent. While this data is freely accessible in its purest form from

---

8 API: Application Programming Interface, that enables cross-system interactions

9 Bitcoin, Ethereum, Binance, Avalanche are all blockchains present in the hackathon and use Merkle Trees. Abitrum, Optimism, Polygon are all layer-2 protocols connected to Ethereum therefore also based on Merkle Trees.

10 Vitalik Buterin "*Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform"*, 2014

11 Element to identify the data item, serving as an identifier, a useful search id/identifier.

12 Stefan Büttcher, Charles L. A. Clarke, Gordon V. Cormack "*Information Retrieval: Implementing and Evaluating Search Engines*", 2010, MIT Press

common blockchain explorers, some third parties offer in-between extract, transform, and load (ETL) features and make the processed data available as in a datamart (reshaped, transformed, and enriched in meaning). The Graph[13] offers the service of hosting of such product[14], with the datamarts called *subgraphs*. Implementable by any individual, some protocols, such as Aave or Sablier, purposely develop their own subgraphs to ease and smoothen the experience of data consumers, to boost attractiveness and promote transparency. The ECB extracted in batches the data for Aave and Sablier parsing indexers' structures and deriving queries. For transparency, the indexer source code is made public. It enables validation of both the process and the derived data, by comparing them to the smart contracts' source code stored on chain.

The extracted Blockchain data points were saved individually in a *key-value* datastore for robustness in case of extraction disruption (each data observation was physically saved individually). They underwent data quality checks including completeness and consistency. The final steps covered structuring and offering the data in the cloud, utilizing Amazon Web Services (AWS) to allow DeFi participants to query the data.

2.2 Challenges and assumptions on the data used in the analysis

The hackathon presented two data-focused challenges and one challenge that did not require any specific data. The data challenges involved analysing Aave and Sablier data in 48 hours. Hackathon participants were given access to 130 Amazon Simple Storage Service (S3) tables through the AWS console. To analyse the blockchain data, the participants were instructed to consult public online sources for detailed information about the content of the tables, which reflected a typical challenge that anybody looking at these data would face. It was difficult to find any comprehensive information about the content of the tables. Although the table names provided some indication of their content, this was not enough and not completely reliable.

Given the time constraints, the hackathon participants prioritised looking for certain variables and established a workflow using Amazon Athena[15] to query data. They subsequently transferred the data to a SageMaker Notebook[16] to analyse them with Python. In the case of Aave, their initial step was to examine the first rows and descriptions of each table to locate key variables like number of users, lending rates, lending amounts, repayment, liquidation amounts, etc. Furthermore, the participants

[13] https://thegraph.com

[14] The Graph offered free hosting service of data indexers in the past and now offers a *pay-per-query* service. Payments for the queries can be settled in The Graph's own crypto-asset - GRT.

[15] Amazon Athena is a serverless, interactive analytics service providing a simplified way to analyse big data using SQL or Python. Athena is built on open-source Trino and Presto engines and Apache Spark frameworks, with no provisioning or configuration effort required.

[16] An Amazon SageMaker notebook instance is a machine learning (ML) compute instance running the Jupyter Notebook App.

noticed that the tables spanned across three different blockchains: Ethereum, Matic, and Avalanche.

Once a desired variable and table were identified, the units were derived and standardised. It was challenging as the detailed additional information and guidelines were not provided by the hackathon organisers. Taking the example of Aave loans on Ethereum, the participants had to create indicators concerning the loan outstanding amounts and borrowing rates which depend on the underlying assets. All blockchain amounts are stored as integers rather than decimal numbers and each crypto-asset has its own decimal system (maximum smallest division). Where one euro can be divided up to one cent, one USDT (Tether, a stablecoin pegged to the USD), for instance, can be divided up to 1 millionth of a USDT (or 6 decimals). Ether has 18 decimals and a percentage point of a rate is stored on 25 decimals. Therefore, the raw information in the blockchain data, for example, an outstanding amount of 10, could represent different values in different crypto-assets. It could refer to e.g. 10 millionth USDT or 10 Wei (the smallest denomination of Ether, 1 ETH = 10 to the power of 18 Wei), or for rates 10 to the power of 25 percent. To standardize these values, additional tables with the mapping of underlying assets and the necessary conversion rates (e.g. dividing by 10^18 for Wei) were looked for and used. Additionally, to allow for meaningful comparisons, the outstanding loan amounts could be converted into USD where a crypto-asset to USD exchange rate was available. All this extra information needed was spread across the various tables. Having investigated all the tables with mappings and auxiliary information, several assumptions still needed to be made in order to transform and link the data for meaningful comparison (see Table 1.2.1).

Table 1.2.1 Standardized loan data from Aave

(The table encompasses variables from tables ave_eth_borrow and ave_eth_reserves)

Column Name	Description	Example Value	Assumption / Note
Loan_ID	Unique identifier for the loan transaction	123456	Auto-incremented or based on transaction hash
User_Address	Borrower's blockchain address	0xUSER...	One user can have multiple addresses
Timestamp	Date and time of the loan, in human-readable format	2023-11-23 12:00:00	Converted from Unix epoch time
Asset_Symbol	Symbol of the asset borrowed	USDC	Derived from the 'underlying asset' address
Amount_Borrowed	The standardized amount of the asset borrowed	1000	Standardized based on the 'decimals' column
Amount_Borrowed_USD	The USD equivalent of the borrowed amount	1000	Assumes real-time conversion or historical price data

In standardising units, the hackathon participants brought the loan amounts to the level of major units of crypto-assets while the original data had often 10^6 to 10^18 decimal places. The raw borrowing rates were also represented by large numbers, and we assumed that a 25 decimal place precision for one percentage point should be applied, and we standardised accordingly. The borrowing rate was also adjusted for the perpetual nature of loans, assuming a daily rate for standardisation.

Another challenge was to disentangle the information on the underlying asset address, as it was often concatenated with a liquidity pool address and sometimes

additionally also with a user address and even a transaction hash. The rationale for such storage is unclear as in many tables there were overlapping information between variables (i.e. several variables concatenated a combination of assets, pools or user addresses or transactions hashes).

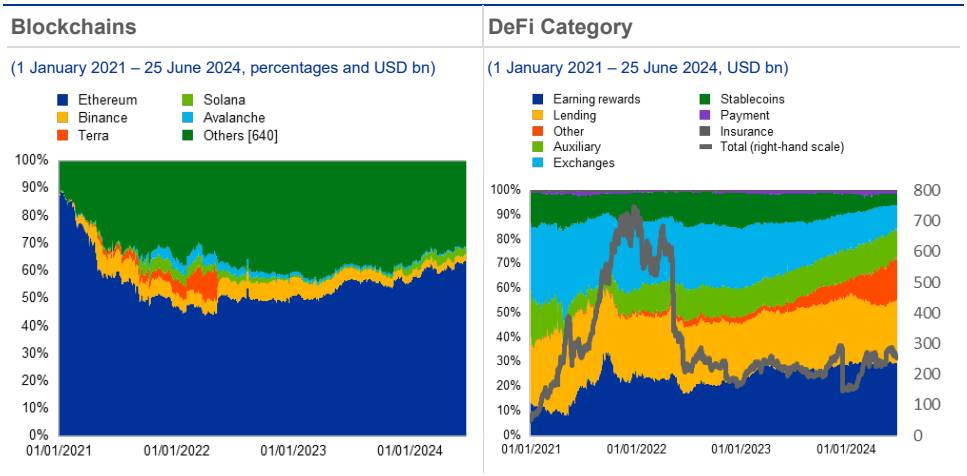
When it comes to Sablier, the data challenges were relatively smaller compared to those of Aave. Firstly, there were only 30 tables containing information spread across 6 blockchains (Ethereum, Arbitrum, Avalanche, Binance Smart Chain, Polygon, Optimism), compared to Aave's 135 tables spread over 3 blockchains (Ethereum, Avalanche, Polygon). Given that almost every table existed with a similar structure in all blockchains (plus a few more tables specific to the Ethereum blockchain), the hackathon participants had on average around 33 distinct tables for Aave per chain versus 5 distinct tables for Sablier per chain (the Ethereum blockchain featured 38 distinct tables for Aave versus 6 for Sablier). The number of variables per table was also lower on average in Sablier compared to Aave. Moreover, most of the Sablier table names clearly described their contents, unlike Aave's tables, which were highly interconnected and made gathering information complex. Almost all Sablier tables in this dataset contained everything needed for the analysis or depended on one additional table only. Still, some assumptions had to be made when mixing information from different tables, as with Aave.

3 Exploring DeFi lending

3.1 Features of DeFi lending protocols

A DeFi platform (protocol) constitutes a decentralised application (dApp)[17] that offers financial services such as lending, trading, or exchanging crypto-assets without banks or brokers. DeFi platforms exploit blockchain technology and smart contracts to allow users to interact directly with each other in a peer-to-peer manner and with financial markets.

Chart 3.1
Total value locked (TVL) by blockchain and by DeFi category of protocols



Source: DeFi and authors' calculations.
Notes: The Total Value Locked (TVL) in all of decentralised finance (DeFi) protocols. TVL represents the sum of all assets deposited in DeFi protocols earning rewards, interest, new coins and tokens, fixed income, etc. TVL might be overestimated due to token re-usage. The DeFi categories cover:
1) Exchanges, protocols that allow users to swap and trade crypto-assets including derivatives and indices
2) Lending, protocols that allow users to borrow and lend assets including against Non-Fungible Tokens (NFTs) or tokenised Real World Assets (RWA) as collateral, as well as without collateral
3) Insurance, protocols that offer coverage against losses caused by events typically in the DeFi ecosystem, such as hacking, malfunctioning of exchanges or smart contracts
4) Payments, protocols that allow users to pay/send/receive crypto-assets
5) Earning rewards, protocols that reward for staked assets including borrowed ones with crypto-assets, offer yield aggregation from various protocols
6) Stablecoins, protocols that mint their own stablecoins including using collateralised lending, provide algorithmic coins to stablecoins, launch new projects and coins
7) Auxiliary, protocols that bridge tokens from one network to another, support DeFi services, and connect data from the outside world (off-chain) with the blockchain world (on-chain) "Oracles"
8) Others: protocols that allow users to bet on future results, have gaming components, offer marketplaces for buying/selling/renting NFTs, obscure the information about transactions, involve Real World Assets (RWA) and their tokenisation, and integrate into social media activities.

Currently, the Ethereum blockchain plays a fundamental role in the DeFi ecosystem, although other emergent blockchains[18] with similar financial services have been

[17] A decentralized application (DApp) is a software application distributed across multiple nodes in a peer-to-peer (P2P) network rather than on a centralised server or authority. DApps are similar to other software applications supported by a website or mobile device, but they leverage on blockchain technology for data storage and processing.

[18] Solana, Cardano, Polkadot, Polygon or Avalanche among others, see e.g. www.coinmarketcap.com.

-Cardano: Proof of Stake blockchain founded in 2017. In 2021 the Alonzo hard fork was launched, bringing the smart contract functionality to blockchain, opening it to dApps development.

onboarded (see Chart 3.1 left). DeFi protocols can be grouped in various categories, e.g. based on the financial functions they emulate (see Chart 3.1 right). The DeFi ecosystem saw a remarkable growth in early 2021 as measured e.g. by Total Value Locked (TVL), followed by a decrease in 2022 (a year that coincides with the first interest rate increase by the FED since long, and where a string of large failures occurred in the crypto markets, e.g. failures of the stablecoin TerraUSD, the crypto hedge fund Three Arrows Capital, crypto lenders BlockFi and Celsius or the exchange FTX). Currently, the DeFi ecosystem seems to start growing again. It is hard to predict how it will further develop in the future, but there are possibilities that some of the innovative aspects of DeFi could be integrated into the future financial world.

DeFi lending is one of the largest DeFi categories in terms of TVL. In this segment, borrowers interact with smart contracts that pool liquidity supplied by lenders (liquidity providers). Differently from traditional financial institutions, in Defi lending platforms the interest rate on borrowed amounts is set automatically, depending on e.g. market conditions (loan demand, pool size) and/or parameters defined in the governance process.

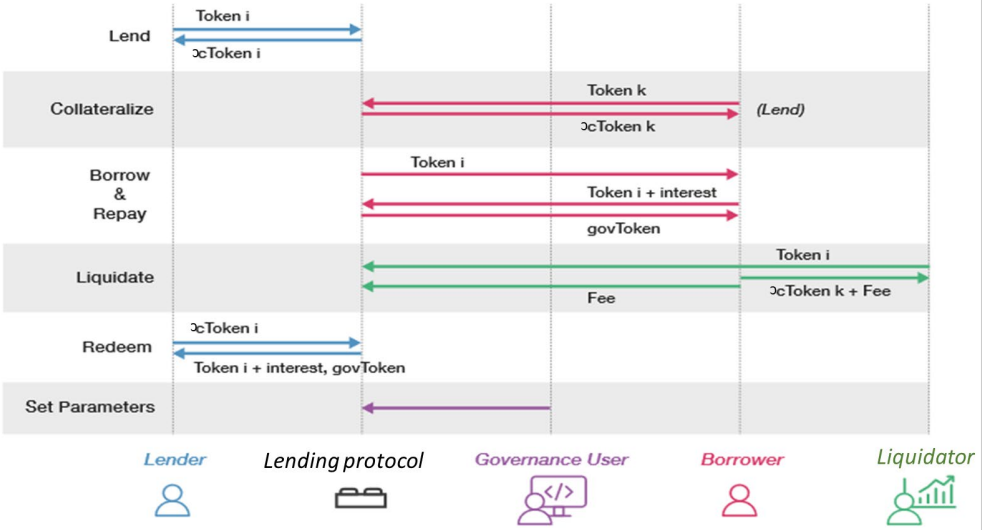
In terms of features of DeFi lending (see Figure 3.1 below)[19], lenders deposit crypto-assets, and in return receive other tokenized assets that allow them to redeem deposits later in time, plus a reward or fee. Lenders bear the credit risk associated with the loan (see section 3.3.1). Borrowers pay interest on borrowed funds, where interest rates can follow different models, such as threshold-based (Kinked) rates, linear rates and non-linear rates[20]. The accumulated interests create the reserves from which lenders' rewards are paid.

- Solana: officially launched in 2020, it is based on innovative hybrid consensus model (proof of history consensus mechanism combined with underlying Proof of stake consensus protocol) aiming to improve scalability;
- Polkadot: known as a layer-0 metaprotocol, it is a multchain network, which can process several transactions on different chains in parallel, behaving as a "parachain". Due to this feature it is able to improve scalability;
- Polygon: Launched in October 2017, at the beginning known as "Matic Network", it is a structured platform for Ethereum scaling. Specifically, it is a Layer 2 scaling solution able to transact up to 65.000 transactions per second on a single side chain;
- Avalanche: launched in 2020, it aims to improve scalability through its unique architecture which consists of three individual blockchain (X-Chain, C-Chain, P-Chain) based on distinct purpose and different consensus mechanisms. The platform interoperates with Ethereum, through bridges development.

[19] To this point, refer to BIS Working Papers No 1066, The Technology of Decentralized Finance (DeFi) R. Auer, B. Haslhofer, S. Kitzler, P. Saggese, F. Victor. Monetary and Economic Department, January 2023

[20] L. Gudgeon, D. Perez., S. Werner, W. J. Knottenbelt "DeFi Protocols for Loanable Funds: Interest Rates, Liquidity and Market Efficiency"arXiv:2006.13922v3 (q-fin.GN)15 Oct, 2020, pag.3.

Figure 3.1 Features of lending protocols



Source: The Technology of Decentralized Finance (DeFi) BIS Working paper No 1066 (adopted).
Notes: xTokens refer to other tokenized assets that allow them to redeem deposits later in time,

Borrowers are usually expected to provide collateral to manage the counterparty risk of default. Lending protocols in DeFi platforms require overcollateralization due to crypto-asset volatility, which means that borrowers deposit as collateral a greater amount than the loan's value itself (with the loan usually equal to 75% of the collateral value, depending on the quality of provided assets).

Borrowing can also be uncollateralised as e.g. in flash loans. Flash loans are either atomically executed and repaid within one individual transaction or reverted, thus they circumvent the borrower default and platform liquidity risks. Flash loans have become one of the most creative tools in the DeFi industry. The main applications of these loans are arbitrage, self-liquidation, collateral swapping, and refinancing. Arbitrage refers to making profits by taking advantage of the price difference of a certain asset on different platforms. Self-liquidation implies using a flash loan to reduce losses, for instance when users want to avoid paying liquidation fees when the liquidation is performed by the platform itself. Collateral swapping comes in play if users borrow from a platform with a certain token as collateral, and then want to change the collateral into another one. Finally, refinancing allows users to successfully obtain a cheaper loan without external funding[21]. Flash loans are also used to exploit the vulnerabilities in smart contracts to carry out attacks and steal a large amount of wealth[22].

[21] Xie, Y., Kang, X., Li, T., Chu, CK., Wang, H. (2022). Towards Secure and Trustworthy Flash Loans: A Blockchain-Based Trust Management Approach. In: Yuan, X., Bai, G., Alcaraz, C., Majumdar, S. (eds) Network and System Security. NSS 2022. Lecture Notes in Computer Science, vol 13787. Springer, Cham. https://doi.org/10.1007/978-3-031-23020-2_28

[22] Qin, K., Zhou, L., Livshits, B., Gervais, A. (2021). Attacking the DeFi Ecosystem with Flash Loans for Fun and Profit. In: Borisov, N., Diaz, C. (eds) Financial Cryptography and Data Security. FC 2021. Lecture Notes in Computer Science, vol 12674. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-64322-8_1

Box 1
Insight on Flash loans

Flash loans are a type of uncollateralized loan available in decentralized finance (DeFi). Unlike traditional loans, these loans do not require borrowers to provide collateral. Instead, they enable borrowers to borrow funds instantly and without principal if the borrowed amount is repaid within the same transaction block (single atomic transaction).

Flash loans are facilitated by smart contracts on blockchains as Ethereum. These smart contracts ensure loan transactions are either fully executed or completely reversed within the same single atomic transaction.

Flash loans are used for various purposes such as arbitrage trading, liquidation of undercollateralized positions, collateral swapping or to profit from temporary market inefficiencies. They exhibit some key features like instant availability, lack of capital requirements or flexibility in terms of the amounts borrowed.

Key risks

Instant availability and **uncollateralized positions** may carry significant risks. If the borrowed funds are not repaid within the same transaction block, the entire transaction is reversed, potentially resulting in losses for the borrower and any counterparties involved as there is no recourse for lenders in case borrowers default on their loans. Additionally, they are vulnerable to **price manipulation** and support forms of **market manipulation**, as they enable large amounts of capital to be deployed and withdrawn within a short period of time.

The reliance on smart contracts to execute transactions exposes them to **smart contract risk.** Smart contract vulnerabilities or bugs can be exploited and leading to loss of funds or deficient performance. When used for arbitrage trading, flash loans are exposed to **price volatility risk** if sudden price movements or slippage lead to losses and arbitrage opportunity disappears before the loan is repaid.

Borrowers may not thoroughly assess the risks associated with the intended use of flash loans. This lack of due diligence can lead to unexpected losses or unintended consequences. One is related with **counterparty risk**, as flash loans involve counterparties that provide liquidity or participate in transactions. In case those counterparties default or engage in fraudulent activities, borrowers may face losses. The other concerns **regulatory risk** as DeFi and flash loans operate in an evolving and uncoordinated regulatory landscape. Regulatory changes or crackdowns could impact the legality or availability of flash loans in certain jurisdictions.

Do flash loans endanger traditional finance?

Flash loans in DeFi pose **minimal direct risk** to finance systems due to their relatively small size compared to the overall financial market. However, they could indirectly impact the financial system. First, as part of the DeFi ecosystem, they are **interconnected with financial markets** and spillover effects can occur in case of significant event or vulnerability in DeFi. Second, the **regulatory scrutiny** in the DeFi space can indirectly impact financial institutions that interact and provide services to DeFi platforms and users. Finally, DeFi and flash loans are forms of financial **innovation**

that could lead to **increased competition** and pressure on financial institutions to adapt and innovate.

However, the extent of this impact depends on various factors, including the growth and evolution of the DeFi ecosystem and regulatory responses to DeFi activities.

How flash loans compare with other loans?

Flash loans differ from other loans in several key aspects such as collateral requirement, instant availability, transaction reversibility or cost and fees (Table 3.1).

Table 3.1 Flash loans and traditional loans

Flash loans	Traditional Loans
Collateral is not required, and borrowers can access funds without any principal too.	Borrowers to provide collateral such as real estate or securities.
Instant availability of funds, allowing borrowers to access liquidity within seconds.	Require application and approval process, which can take days/weeks to complete.
Repaid within the same transaction block on the blockchain if not, transaction is reverted, and loan is cancelled.	Fixed repayment schedules, penalties for late payments or in case of default.
Smaller transaction sizes for short-term, high-frequency trading strategies in DeFi.	Larger transaction size for real economy purposes such as business operations or private consumption.
Lower fees as they eliminate the need for financial intermediaries and collateral.	Larger transaction size for real economy purposes such as business operations or private consumption.

Overall, while flash loans offer unique opportunities for DeFi participants to access liquidity and execute complex trading strategies, they come with significant risks. They require careful consideration and risk management due to their inherent risks.

Price fluctuations can lead to insufficiently collateralised loan positions and if a borrower does not provide additional collateral, the loan might be liquidated. Liquidators trigger the process via a smart contract. Any network participant equipped with some prerequisite information about insufficiently collateralised loans can be a liquidator[23]. A liquidation mechanism typically involves a nonatomic English auction process or an atomic fixed spread strategy.[24] Sometimes liquidators can repay only part of the borrowed position and receive in return a fraction of the borrower's collateral at a discount with respect to the market price. At the new market prices, either the remaining borrower's collateral is sufficient to back the fraction of loaned crypto-assets that were not liquidated, or it will be subject to subsequent liquidations. Part of the liquidation fees can be retained within the protocol.

In case of a race between the borrower (trying to post more collateral and keep the loan active) and the liquidator (trying to liquidate to get hold of some of the collateral) the issue of mempools[25] should be highlighted. Mempools are pre-chains where

[23] See e.g. https://docs.aave.com/developers/guides/liquidations

[24] Kaihua Qin, Liyi Zhou, Pablo Gamito, Philipp Jovanovic, Arthur Gervais An empirical study of DeFi liquidations: incentives, risks, and instabilities; IMC '21: Proceedings of the 21st ACM Internet Measurement Conference November 2021Pages 336–350 https://doi.org/10.1145/3487552.3487811

[25] Mem is a short for memory.

network nodes add a transaction before miners/validators pull it from there and add it to the blockchain (validate it). Such features might impair market integrity, as in some cases a savvy liquidator could pay miners an extra fee to ensure that their transaction gets listed ahead of the borrower within the same block[26].

The last layer of the lending protocols involves governance. Holders of the tokens accepted for the governance of the protocol can take part in deciding on procedures, vote and execute smart contracts. Typically, proposals concerning the collateralisation thresholds or interest rates are voted.

3.2 Hackathon challenge 1: Aave

The Aave Protocol is a decentralised non-custodial lending and borrowing protocol where users can participate as liquidity suppliers (i.e. lenders), borrowers, and liquidators. Lenders provide liquidity to the platform by depositing eligible crypto-assets they own into *liquidity pools* which are governed by open-source smart contracts. Deposited funds are stored in a smart contract, in exchange for tokens (*aTokens*) representing the lender's position including an interest in return of the crypto assets provided. Borrowers are able to borrow from these liquidity pools after posting eligible[27] crypto-assets collateral in an overcollateralised fashion. Borrowers mint the debt token (*Aave tokens)*. The interest rate of an Aave pool is decided algorithmically by the smart contract and depends on the available funds within the lending pool. The more users borrow an asset and as a result deplete a lending pool's liquidity, the higher its interest rate rises. A lending pool can consist of several cryptocurrency assets, for instance ETH, DAI, and USDC.

In Aave, when the loan Health Factor[28] drops below 1, any liquidator can call the public pool function *liquidationCall*, by repaying parts or all of the outstanding debt, while profiting from the liquidation spread. Aave specifies that only a maximum of 50% of the debt can be liquidated within one liquidationCall execution (referred to as a close factor) if the Health Factor is above the corresponding close factor threshold, or 100% if the Health factor falls below. Borrowers in Aave can also engage in one-block borrow transactions or "flash loans" (see par. 3.1), which do not require collateralization.

Aave supports token ERC technology (Ethereum network) and over time extended to other blockchains (e.g. Avalanche, Polygon, Optimism. Arbitrum, Fantom, Harmony). The development of Aave has been carried out via its Aave Improvement Proposal

[26] See e.g. Inside the 'mempool,' where crypto risks hide

[27] Tokens accepted as collateral in the Aave protocols undergo a selection process based on community governance decision. The selection criteria include asset reliability (time-tested smart contracts), and stability (non-volatile assets, e.g stablecoins) liquidity in view of ensuring collateral realisation the event of loan liquidation.

[28] The Health factor is defined as the weighted average of the Liquidation Threshold (percentage at which a position is defined as undercollateralized) of the collateral assets and the values of borrows.

processes (AIPs) which are voted by holders of the governance tokens (AAVE, aToken, stkAave) or delegated participants[29].

Chart 3.2

Aave instances

Source:DefiLlama. Aave V3 Techncal Paper; January 27, 2022 https://github.com/aave/aave-v3

Since its creation, the Aave protocol has seen numerous updates, from the ETHLend to the current version 3 of Aave (V3) (see chart 3.2). Starting in 2017 as a peer-to-peer lending platform (under the name *ETHLend), the protocol initially* matched individual lenders and borrowers. Then it evolved to liquidity pools-based lending and was rebranded in 2020 to Aave. Aave's V1 version brought innovations such as Flash Loans and aTokens, but used a pooled risk model (all assets being at risk of liquidation). Aave's version 2 (V2), released in December 2020, supported more assets and introduced some new features such as debt tokenization, credit delegation[30], or the deposit pool centric architecture. Subsequently, V3 was released in March 2022[31]. Among other improvements, it introduced *Portal*, a bridging[32] tool for the Aave covered networks, isolated liquidity pairs[33] (supply/collateral pairs in which borrowers can only borrow one asset at a time with a specific collateral) and isolation mode (assets are borrowable/lendable without affecting other assets in the wallet), improving risk management tools. Specifically, V3 offers capital efficiency - enhancements of the borrowing power when collateral and borrowed assets are correlated in price. Finally, Aave Arc is a permissioned market instance restricted to

[29] Information available on the site https://docs.aave.com

[30] Credit delegation which allows a depositor to deposit funds to earn interest, and delegate borrowing power to other users.

[31] AAVE Version 3 Whitepaper: https://github.com/aave/aave-v3core/blob/master/techpaper/Aave_V3_Technical_Paper.pdf

[32] Protocol V3 allows to burn aTokens on the source network and at the same time minting them on the destination network. In this way the assets can be moved on different networks though a bridge.

[33] Isolation Mode was inspired by MakerDAO approach for exposure management: borrowers suppling an isolated asset as collateral cannot supply other assets as collateral, though they can still supply for yield gaining. More sophisticated risk parameters have been introduced, as supply and borrow caps and granular borrowing power control, which enable the possibility to lower borrowing power of an asset without impacting existing borrowers.

institutions which undergo a KYC procedure to comply with AML standards introduced in January 2022[34].

Considering Aave's significant share among the DeFi lending protocols, interesting developments over time, as well as a perceived high availability of corresponding public information, Aave blockchain data were selected as one of the challenges of the DeFi Hackathon. The hackathon participants' output was assessed against a set of analytical questions and indicators (see Table 3.2).

Table 3.2 Challenge 1. Analytical questions/indicators to cover in the analysis

What are the indicators on DeFi lending (e.g. number of users/new users, lending rates, lending amounts, maturities of loans, repayment of loans, liquidation factor, liquidation amounts, collateral used, debt tokenisation, credit delegation and insight into deposit pools, etc.)?
How indicators on DeFi lending compare across various blockchains?
What are the risk characteristics of the loans and potential mitigants (e.g. collateral)? Can metrics to identify loans with different credit risk assessment be defined ("credit ratings"), including to identify substandard or non-performing loans? How do the risk characteristics of DeFi loans compare to credit risk factors that are analysed for bank loans? What factors influence the risk characteristics?
What information is available on borrowers?
Can KYC-enabled permissioned parts of the protocol be identified (segregated permissioned pools of 'whitelisted' users that have passed Know Your Customer)? Do the indicators for KYC-enabled permissioned part of the protocol differ from those for the non-KYC-enabled part and if yes how?
Is there any information on flash loans? What indicators help to analyse the risks from flash loans?
Is any information on "mempool" available which could be used by sophisticated actors to force liquidations by creating artificial price volatility?

3.3 Findings of the hackathon

This section provides insights into the outcome of the hackathon challenges and directions of prospective further exploration. The hackathon participants had 48 hours at our disposal to analyse Aave historical data spanning until February 2023. The participants did not know each other and largely did not have any prior exposure to Aave protocol and its blockchain data. Constrained by the hackathon form, the analysis did not strive for comprehensiveness, but for demonstrating the potential for granular analysis performed directly on raw data rather than relying on third party data aggregation.

The hackathon participants started the analysis by establishing key performance indicators (KPIs) by blockchain. KPIs included the number of users, lending rates, lending amounts and liquidity factors. Out of more than 1.25 million unique addresses, an overwhelming majority where on Matic: 87%, followed by Ethereum: 10%, and Avalanche: 3%.

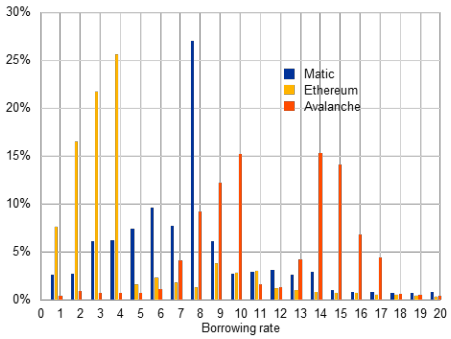
[34] Fireblocks handles verification: https://www.fireblocks.com/blog/permissioned-defi-goes-live-with-aave-arcfireblocks/

Chart 3.3

Loans features by blockchain

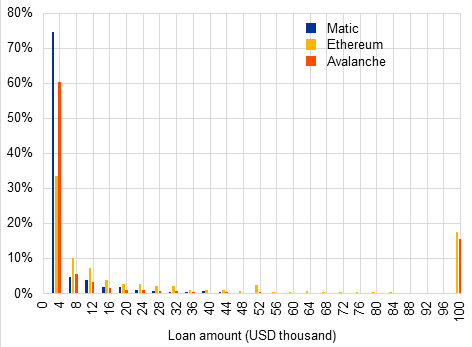
Borrowing rates

(histograms)



Loan amounts

(histograms, loans amount are clipped at USD 100,000)



Source: Aave and authors' calculations.

A close look at lending rates indicated that loans on Avalanche were more expensive than on other blockchains suggesting higher perceived risks (possibly from relative lower liquidity), with almost half of the loans having an interest rate of more than 5%. On the other hand, three quarters of all loans on Ethereum, and Matic had an interest rate below 4% (see Chart 3.3 left).

Regarding loan values in USD, larger loans were primarily linked to Ethereum (see Chart 3.3 right), featuring a median value of USD10,000, with a quarter of all loans exceeding USD50,000. In contrast, the median loan value for Matic stood at USD45, while Avalanche occupied a middle position, with a median loan value of USD10,000, yet with 15% of its loans surpassing USD100,000.

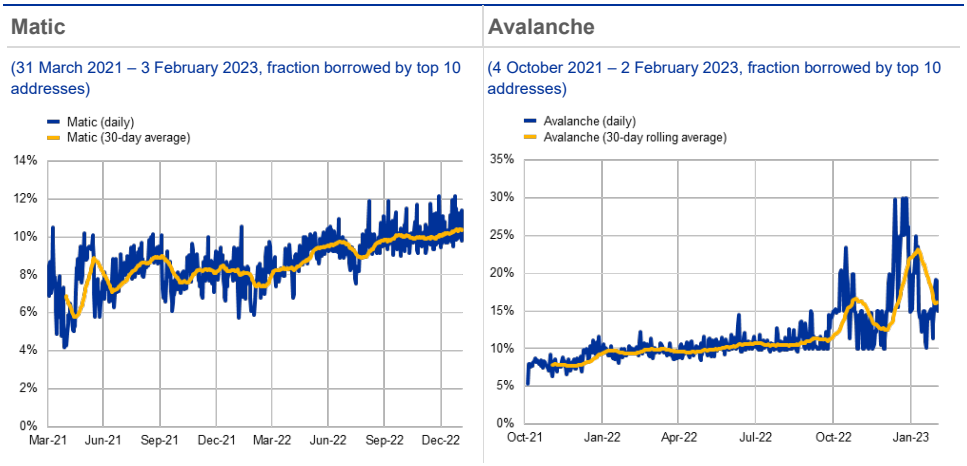
This allowed the hackathon participants to identify different types of users per blockchain which are aligned with the different characteristics of each blockchain. For example, Matic was known for its cost-efficiency performance explaining that we saw a large number of users, and smaller loans. On the other hand, Ethereum is the most established of the three analysed blockchains, and at the time of analysis it followed a costly proof-of-work (PoW) validation mechanism instead of proof-of-stake (PoS) used on Matic and Avalanche. PoW was considered more robust and secure, so larger loan amounts were expected. Finally, Avalanche's riskier loans could be attributed to the fact that this blockchain was the newest, therefore its liquidity was lower and potentially could attract users more open to novel and speculative instruments.

The concentration of borrowers in the Matic and Avalanche networks was also analysed (see chart 3.4). Looking at the fractions of the total debt of the 10 addresses who borrowed most each day, the concentration indicator amounted to around 10% on Matic and was higher on Avalanche. This level of concentration remained relatively stable for Matic, while it was volatile for Avalanche, indicating an increase in the activity of the main borrowing addresses, reaching levels of up to 30% of total debt. Such concentrated debt possibly indicates systemic vulnerabilities and risks and may contribute to the higher loan rates observed on Avalanche. More

research could be conducted to understand the liquidity and stability of these dynamics.

Chart 3.4

Borrowing addresses concentration



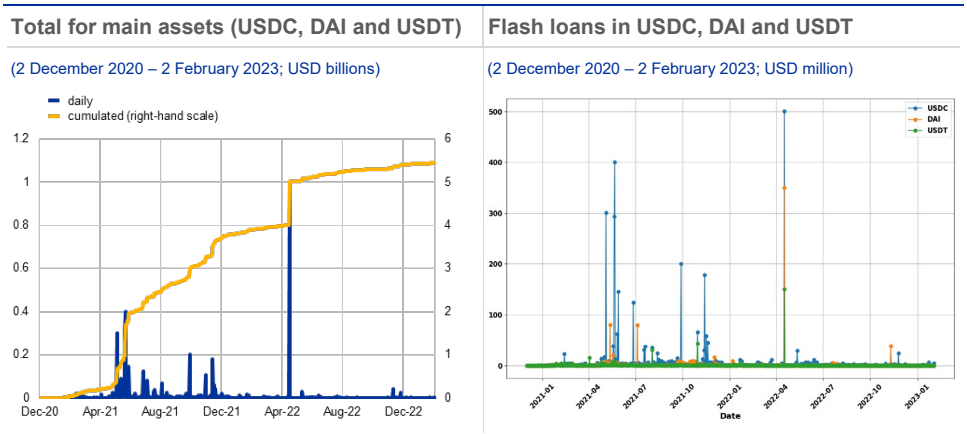
Source: Aave and authors' calculations.

Other indicators constructed and analysed covered loan volumes in selected crypto-assets and the respective concentration of lending addresses. Furthermore, concerning flash loans, instances of borrowing of several hundred million in USDC were observed. In particular, the value of flash loans in the analysed sample reached the threshold of 200 million dollars on four days, peaking at 1 billion dollars on 17 April 2022 (see Chart 3.5 left panel), on Ethereum on Aave. The primary crypto-assets used in these flash loans were USDC, DAI, and USDT. To compare these magnitudes with collateral debt on Aave in Ethereum, Chart 3.6 left hand panel shows that while flash loans represented only a small percentage of all operations on most days, they accounted for more than 200% on four days, and on 17 April 2022, they represented 1200% compared to collateral loans on Aave. By design, these large loans were taken without posting any collateral and repaid within the same transaction, after being used to generate profit, possibly through arbitrage or by exploiting vulnerabilities in smart contracts through so-called flash loan attacks. In fact, on 17 April 2022 there was an incident in which flash loans were used to attack the stablecoin protocol Beanstalk, stealing 182 million dollars[35].

[35] https://www.theverge.com/2022/4/18/23030754/beanstalk-cryptocurrency-hack-182-million-dao-voting

Chart 3.5

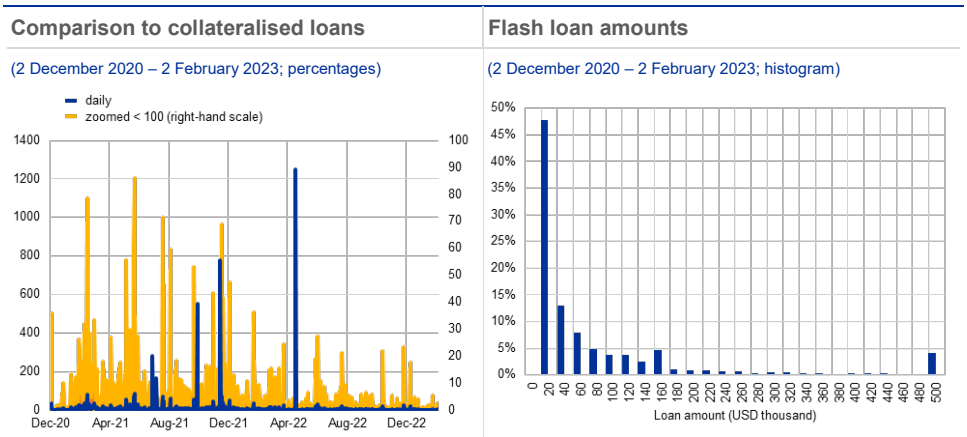
Evolution of flash loans on Ethereum



Source: Aave and authors' calculations.

Chart 3.6

Flash loans



Source: Aave and authors' calculations.

3.3.1 Commentary on risk

As DeFi mimics traditional finance concepts it is worthwhile to reflect on traditional risk management concepts in the context of DeFi lending, such as counterparty credit risk, liquidity risk or operational risk, having in mind however that in DeFi the risks are to a large extent borne in a direct manner by the DeFi users rather than by a central intermediary, the bank.

Counterparty credit risk on the Aave protocol is mostly handled through the overcollateralization of loans and the effective enforcement of liquidation (see sections 3.1 and 3.2) which are meant to prevent the issue of non-performing loans to arise. In that sense, DeFi lending on Aave should be compared to securities financing transactions (SFTs) rather than to retail lending. It enables retail participants to perform transactions of such nature (with crypto-assets) in contrast to the traditional banking system where SFTs are available mostly to wholesale or

professional clients. Along this line, an interesting idea to investigate could be the comparison of the Aave protocol risk management framework for rating the collateral assets and deriving their risk parameters (e.g. loan to value limit, liquidation threshold) with the Basel framework for the treatment of collateralised transactions which sets haircuts for different assets[36].

As for the **liquidity risk**, the loans taken by borrowers have no maturity (i.e. are perpetual) however they can be forced into early repayment or liquidation by for instance increasing funding costs in case of a variable rate, or market movements affecting the valuation of their collateral which the borrowers cannot compensate by posting more collateral. The funding costs are determined by the liquidity supply and demand of the borrowed asset. There is a risk for floating rate loans when liquidity providers withdraw their deposits in a given liquidity pool. This may trigger higher funding cost for the borrowers in the pool and potential liquidation which can then turn into fire sales of the collateral and trigger further liquidation.

Regarding the **operational risk**: several sources of novel operational risks can stem from DeFi lending compared to traditional finance. The correct functioning of the protocol relies on the liquidation of overcollateralized loans. But as collateral is composed by crypto-assets there is a risk that a devaluation of collateral used by the protocol triggers losses to market participants and to the protocol itself. An example of such risk is the quasi-overnight devaluation of circa 20% of the USDC stablecoin when part of its reserves were found to be in custody with failing Silicon Valley Bank. Another operational risk is the one associated with the use of smart contracts either from the protocol itself or from the tokens in circulation on Aave (e.g. the collateral tokens). The protocol being algorithmically run is exposed to the risk of IT security vulnerabilities in the smart contract code that if discovered can be exploited by hackers to steal users' funds[37]. In a similar manner, flaws in collateral asset smart contracts not controlled by Aave may result in collateral becoming worthless and triggering massive liquidation of all borrowers that have pledged that collateral.

Finally, the valuation of all tokens on the Aave protocols is based on market prices sourced from an oracle and thus bear the downstream risk in case of oracle failure or manipulation (see chapter 5).

[36] Basel Framework calculation of risk weighted assets for credit risk

[37] See e.g. a bug labelled as "critical" occurred in November 2023: "Aave pose assets on Avalanche, Polygon, Optimism,and Arbitrum" https://coinpaper.com/2538/aave-pauses-assets-on-avalanche-polygon-optimism-and-arbitrum

4 Exploring DeFi payments

4.1 Features of DeFi payment protocols

DeFi payment protocols offer the ability to pay, send and receive crypto-assets in a largely decentralised way utilising smart contracts. The Total Value Locked (TVL) in DeFi payment protocols reached a peak in November 2021 but amidst a general crypto-asset meltdown dropped down significantly. Payment protocols represented at the time of the Hackathon around 0.01% of Total Value Locked in all DeFi sectors, approximately $250M, according to Defillama.[38]

Although widely used, it is debatable whether TVL is the best metric to measure the popularity and adoption of a payment protocol since it is inflated by the price of the assets. There are other measures that can be considered more robust to evaluate payment protocols like the number of transactions, the average value of transactions or the total value of transactions in a certain period.

There are a few DeFi payment protocols, each with a unique focus on a specific service. There are protocols facilitating instant real-world payments by offering a point-of-sale solution for spending crypto-assets at merchants, e.g. Flexa[39]. Flexa achieves this by using a hybrid-approach utilising a centralised framework and a decentralised network. In order to facilitate more flexible transactions, the Lightning network has gained interest as a second layer for Bitcoin facilitating “instant” micropayments. In addition, some protocols aim to facilitate the on/off ramp concerning official currencies by offering cross-chain decentralized payment networks and collateral staking to issue stablecoins (e.g. Ramp DeFi). Furthermore, some protocols target specific areas like business-to-business (B2B) payments. They enable the creation and tracking of invoices with automated payments and integration with various payment gateways (e.g. Request Network) or web3[40] related payments (e.g. Sablier).

A number of Defi payment systems, e.g. Sablier, Llamapay and Superfluid, use a streaming mechanism for payments, enabling continuous or interval payments based on a specified frequency and timeline. These protocols can be used for high-frequency payrolls (instead of monthly based), subscriptions or instant, second based, payments. Money streaming is a noteworthy innovation that distinguishes itself from services offered by traditional payment solutions, capturing the interest of users and investors.

The hackathon focused on these DeFi protocols offering the money streaming feature. Within this domain Sablier was chosen for in-depth analysis due to its non-complexity and popularity at the time of the hackathon. Sablier represents a small

[38] https://defillama.com/protocols/Payments

[39] https://flexa.network/

[40] Web3 is the common term to refer to the next evolution of the World Wide Web. An open, transparent, decentralized and user-centric internet. The term is often associated to technology innovations such as blockchain or decentralized applications (DApps).

part of the DeFi payments ecosystem, as presented in section 4.2, however it is indicative of the possibilities offered by this ecosystem because of both the streaming functionality and the innovative characteristics of DeFi payments. Since the Hackathon, DeFi payments protocols doubled in terms of size benefitting from the comeback of crypto-asset markets, usage of stablecoins and the development of the Lightning Network built on top of the Bitcoin blockchain. The Lightning Network is gathering the support of numerous projects targeting different use cases[41].

4.2 Hackathon challenge 2: Sablier

Sablier is a protocol for real time payments that can be used in scenarios where payments are recurrent or/and need to occur over a specific duration such as salaries, subscriptions, donations or other cases that require regular or continuous payments (explained below). More recently it is also used for services specific to DeFi governance, such as vesting[42] or airdrops[43].

In the Sablier protocol, the payer can define the start and end time of the streaming and the amount that can be transferred per second. After depositing an amount for a specific recipient address, smart contracts start “streaming” the tokens towards the recipient (see Chart 4). The transfer of value on the protocol follows a process akin to an hourglass mechanism (hence the protocol name). A proportionate fraction of the entire contract value is transferred each second until the contract is mature.

Money streaming encompasses frictionless and closed-end or open-ended / continuous payments. They can be used to establish a real time direct link between the value transfer and the service provision, without delays, risk of missed payment and without intermediaries. Such features may facilitate the creation of streaming exchanges, saving accounts, cashflow advances (allowing people to take credit based on their advances recorded on a blockchain), continuous actions (e.g. auctioning of an advertising space)[44]. Sablier claims to offer to users, mainly active in web3, a range of new possibilities such as the easy creation of continuous payment streams to send and receive funds in real-time over a specified duration and transparent execution of payment streams by leveraging on smart contracts or a full control of funds without relying on a central authority namely a credit institution or a central bank.

The web interfaces for streaming money and for receiving the streamed money are developed and administrated by Sablier Labs, the developer of the Sablier Protocol. The protocol architecture is open to the community for further developments after a period of four years subject to a Business Source License. The license restricts the

[41] Lightning Network Landscape

[42] Vesting refers to the gradual or conditional release of tokens to stakeholders like employees, founders, investors or community members.

[43] Airdrop refers to the free distribution of tokens or coins to the holders of a specific existing crypto asset. It is usually part of marketing or promotional strategies in a context of an initial coin offering (ICO) or as a reward to participants of an existing crypto assets project. Airdrops entail risks such as user data collection.

[44] Streams, a New Defi Primitive by Francesco George Renzi (CEO & Co-founder at Superfluid) at TRTM 3.0

utilization of the code in a commercial production environment. The protocol covers main blockchains (e.g. Ethereum, Polygon, Optimism), and tokens (ERC-20 and ERC-721 non-fungible tokens).

Sablier started by providing an overlying layer to ease access to the service through an application. Since its inception in June 2019 the protocol had two main updates. Sablier V1.0[45], launched in December 2019, added important features as a mobile version, access to wallets besides Metamask, bug-fixes and more simplicity and usability of the interface when streaming money. Sablier V2.0[46], introduced in July 2023, extended functionalities to non-linear streaming payments or payment streaming from non-fungible tokens (NFTs), was deployed in further blockchains, and added security features and a technology licensing system.

The TVL of the Sablier protocol reached a peak in the beginning of November 2021 with 1.5 billion USD. Since then the TVL decreased sharply just to 4.5 million USD (December 2023)[47]. At the time of the hackathon, Sablier V1.0[48] had the largest share of the total Sablier TVL, however the TVL of Sablier V2.0 kept increasing.

The hackathon participants were called to dive into the Sablier data provided in the hackathon context and produce a set of indicators and responses to several analytical questions (see Table 4.1).

Table 4.1 Challenge 2. Analytical questions/indicators to cover in the analysis

1 What are the indicators on users of this protocol?
2. Can the purpose of transactions be distinguished?
3. Can geographical angle of transactions and users be grasped based on the data?
4. Is money streaming widely used?
5. Are there any differences in transactions and users on various blockchains?
6. Can vendors be identified?
7. Is it possible to obtain information by type of sectors/services of the economy (e.g. e-commerce, travel and tourism, entertainment, Luxury & Fashion, etc)?
8. Are there any insights available based on the data on the major players?
9. What are potential risks from this protocol?
10. Can one infer crypto-assets payments attitudes from the data presented?

Source: ECB

4.3 Findings of the hackathon

Having generated various indicators to understand the dynamics of the Sablier protocol revealed *inter alia* that the protocol was mainly used for payments of short duration (see Chart 4.1 left). When comparing the deposits and withdrawals of

[45] Sablier December 2019

[46] Sablier July 2023

[47] Please note that the calculation of TVL also was adapted and as of October 2022 does not include vesting tokens.

[48] Version of the protocol launched in December 2019 to replace the app. Sablier V1 added a mobile version, extended the acceptance to other wallets besides Metamask, fixed bugs detected and add simplicity and usability to the interface when streaming money.

USDC, USDT and DAI, almost half of the balances invested in Sablier remained unwithdrawn (“sleeping”) (see Chart 4.1 right).

Although it was difficult to distinguish the purpose of most of transactions, the Sablier service seemed to be primarily used by early adopters which can be related also with supporting governance activities of the protocol such as vesting or air drops. The Polygon blockchain emerged as the preferred blockchain due to its faster and more cost-effective nature compared to Ethereum (at the time of the hackathon). While the usage of Sablier on Polygon was mainly for token vesting, Sablier on Ethereum focused on salary payments.

Chart 4.1

Selected indicators on Sablier

Source: Sablier blockchain data and authors’ calculations.

Unsupervised machine learning (ML) model was used to crunch the data and uncover clusters of transactions that can be subject to interpretation, e.g. timestamp clusters may reveal geographical trends. In addition, transaction types within the Sablier framework were classified. For simplicity, only on transactions on the Ethereum blockchain and all possible tokens were considered. To ensure consistency, additional tables were consulted. These tables map underlying assets to their corresponding conversion rates, like for example, applying a division by 10^18 for Wei units. Unfortunately, due to time constraints, the hackathon participants were not able to transform the value of tokens into dollar prices, so interpretation of the results should be cautious (in particular large amounts can reflect a token which comprises a high number of basic units (see chapter 2 on data), for instance 1 ETH would be expressed for instance as 10^18 while 1 USDT would be represented by 10^6). A k-means algorithm was employed to create clusters of transactions based on four features (see Table 4.2). This algorithm is an unsupervised machine learning model that groups data into clusters based on feature similarity. The essence of k-means lies in its capability to identify inherent patterns and segregate data points into coherent groups without any prior labelling.

Table 4.2 Features used in k-means algorithm to cluster Sablier transactions

Feature	Description
'deposit'	corresponds to the number of tokens or currency deposited in each transaction
'is_ day'	a binary indicator with value 1 if the transaction occurred between 6 am and 6 pm in the CET timezone
'is_ weekend'	a binary indicator with value 1 if the transaction occurred during weekends
'is summer'	a binary indicator with value 1 if the transaction occurred in summertime

By applying the algorithm to the Ethereum blockchain data, four distinct groups of transactions were identified. The first group encompassed the transactions happening only between 6 pm and 6 am in the CET time zone (i.e. at night in this time zone but day in other zones), and on weekends, thus potentially used during leisure periods. The second group covered transactions of a small number of tokens that occurred mostly at night in the CET time zone (day in other time zones) and only during summer. Such characteristics if happening in traditional payments could possibly be classified as a travel and tourism cluster, however in the context of the current low adoption rate of Sablier payments and a lack of longer times series (i.e more summers) also other interpretations are possible. The third group focused on mid-size volume transactions that happened only during daytime in the CET timezone and weekdays. This cluster could possibly capture salary payments which is the focus of Sablier on the Ethereum network. The last group featured very big volume transactions but with relatively low value prospectively indicating transfer of crypto-assets of speculative nature (e.g. meme tokens for which supply is often inversely correlated with their value). The analysis of these clusters showed that the majority of Sablier transactions on Ethereum seems to be concentrated in the salary payment cluster, which is indicative of regular, business-related activities. These activities likely include B2B transactions, services payments, and possibly payroll distributions, which aligns with the fact that Ethereum in Sablier has much more recipients than senders, and with the fact that Sablier describes itself as useful for continuous and autonomous payroll[49].

4.3.1 Commentary on risk

Due to the limited timeframe of the hackathon, the analysis of Sablier data was not intended to produce comprehensive results. Instead, the goal was to provide some early results and showcase the potential for future analysis. However, it was considered worthwhile to reflect on the potential extensions and risks requiring further analysis.

The Sablier protocol and payments protocols in general may bring potential benefits in terms of diminishing delivery versus payment risks (DVP), reducing the need to rely on financial intermediaries in the global payments infrastructure and making interactions overall simpler, faster and more efficient. In addition, there is the possibility to support micro or nano-payments which might enable value locked in different business areas of all geographies.

[49] https://blog.sablier.com/sablier-v1-0-is-live/

As DeFi follows traditional finance, risk management concerns in the context of DeFi payments protocol including Sablier still exist and need to be properly tackled. For instance, legal risk arises from the difficulty of identification of both parts of an unsigned contract which is typical from DeFi. Liquidity costs increase when collateral to ensure settlement is blocked on the contract. Permanent loss of blockchain signing key and access to funds is a new operational risk. To ensure that privacy is kept only strict minimum info on DLT account addresses is provided, which might however raise issues with respect to transparency and for ensuring the possibility of reverting transactions.

There are also other risks related to regulatory obligations in traditional payment systems that are still not applicable to these types of protocols, namely related to Know-Your Customer (KYC), AML or fraud prevention rules. There are no obligations to define stop sending rules, seizure of fund protocols, customer notifications, handling of successions (heritage) or traceability of funds as in regulated payment systems. Finally, standards to account for adequate governance, cyber risk or reporting are not defined and not commonly adopted by different protocols.

5 Exploring Oracles

5.1 Features of blockchain oracles

Blockchain oracles are third-party services that enable DeFi applications to receive external data necessary for the execution of their smart contract. Smart contracts are programmed to self-execute actions based on predefined rules or triggers. Nevertheless, smart contracts do not possess inherent knowledge of real-world information like weather conditions or race results. They need an external source to provide this data for them to make decisions or execute functions. Oracles act as the “bridge” between on-chain (blockchain) of off-chain (external) systems.

The establishment and maintenance of trust is fundamental to the effective functioning of oracles[50]. Oracles embed varying degrees of trust depending on their level of centralisation; the ones with a high degree of centralisation rely on a limited number of data sources or have a single / limited number of entities responsible for transmitting information to the platform. This high level of centralisation leaves room for manipulation when the data sources are not trustworthy[51]. On the other hand, implementing fully decentralised oracles that can incorporate real-world information is difficult as it requires ensuring credible reporting in the absence of a single authoritative resource[52]. The fact that certain decentralised blockchain applications such as DeFi protocols must rely on off-chain data inputs to function, which opens a door to price manipulation is referred to as the “Oracle dilemma” or “Oracle problem”.

There are different types of blockchain oracles. In addition to centralised and decentralised ones as described above, there are software/hardware oracles, inbound/outbound oracles, contract-specific oracles and human oracles. By utilizing their programming capabilities, software oracles can extract information from predetermined sources such as online APIs, databases, exchanges, and other digital platforms. Hardware oracles facilitate the connection between blockchain networks and Internet-of-Things (IoT) devices, as well as other hardware such as sensor-equipped devices. Inbound oracles are responsible for bringing external information into the blockchain, whereas outbound oracles transmit blockchain data to external systems. Oracles can be tailored exclusively to the specific needs of individual smart contracts. Finally, human oracles rely on actual individuals to validate and provide information for smart contracts. Although there is a human element involved, rigorous systems are established. Aiming at achieving a high level of trust and reliability, the most prominent oracles are decentralized middleware[53] entities. This decentralised approach connects smart contracts to validated resources outside of

[50] Chanelle Duley, Leonardo Gambacorta, Rodney Garratt and Priscilla Koo Wilkens, The oracle problem and the future of DeFi, BIS Bulletin, Sep. 2023

[51] Auer R., Haslhofer B., Kitzler S., Saggese P., Victor F.,”The Technology of Decentralized Finance (DeFi)” 19 January 2023, BIS Working Papers No 1066, 13-14.

[52] Garratt, R, and C Monnet, “An impossibility theorem on truth-telling in fully decentralized systems”, BIS Working Papers, no 1117, August 2023

[53] Middleware is a type of computer software program that provides services to software applications beyond those available from the operating system. It can be described as "software glue".

their native blockchains, in order to acquire secure price feeds which are usually aggregated to provide a final reported price.

The following sections cover the challenge on oracles as was set in the ECB Hackathon. From the basis of the outputs generated in the Hackathon, an analysis of the functioning of oracles is provided - including with a case study on the business model of one oracle provider. From the risk perspective, cases of oracle manipulation and malfunctioning are identified, and ways to potentially alleviate vulnerabilities via regulation and good industry practices are discussed.

5.2 Hackathon challenge 3: oracles

The hackathon challenge was to analyse oracles using information from DeFi literature and published case studies. The hackathon participants were asked to review DeFi Oracles and elaborate on risks, indicators and data sources that may be of relevance to supervisors and central banks (See Table 5.1 covering analytical questions concerning challenge 3).

Table 5.1 Challenge 3. Analytical questions/indicators to cover in the analysis

1. Business models of blockchain oracles service providers (how do they make money?)
2. How are oracles linked to DeFi protocols or how do they feed their data to the DeFi protocols?
3. What are the risks associated with oracles?
4. Examples of malfunctioning oracles. Were malfunctioning oracles another of DeFi's flaws exposed during the UST failure?
5. How reliable are the data from oracles and how do they quality assure prepared their data? Is it checked beforehand? Is there any independent/third party providing assurance?
6. What is the concentration of oracle service or oracle providers? Are there oracles that are superior in the sense of having larger underlying data to generate/display prices, in terms of governance, or in the sense of having better mechanisms to avoid malfunctioning or cyber-attacks?
7. What are the available data sources concerning oracles?

Source: ECB

5.3 Findings of the hackathon and review of case studies of malfunctioning or manipulation

In order to provide insights on the questions raised in the challenge, the hackathon participants studied the various types of DeFi oracles, their functioning and characteristics, as well as the prominent oracle providers and related risks. This part thus presents the way oracles work, the types of oracles, the case of Chainlink, which is one of the most widely used oracles, types of risks and malfunctions and cases where oracles malfunctioned or were subject to manipulation.

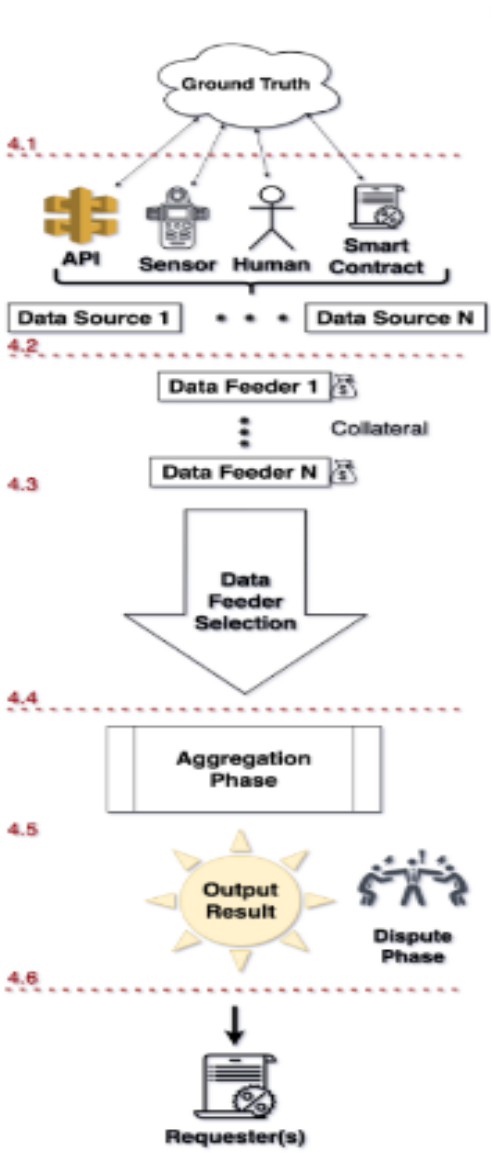
5.3.1 Oracle types and functioning

The study of oracle solutions in the DeFi ecosystem reveals a variety of implementation designs. The two main alternatives are the centralised and the decentralised model, with the latter generally considered more secure and more aligned with DeFi's philosophy (no reliance on a single data provider). The decentralised Oracle model uses multiple nodes, and its operation is based on a number of general steps[54] which are depicted in Figure 5.1 and analysed in the following.

First, the oracle receives a price request from the blockchain and queries external data sources for the desired information. Specific entities, the data feeders, gather and report the data from a data source to the oracle system. There is a data feeder selection process to select legitimate and credible data feeders. After this process the reported data is aggregated, then a dispute phase may verify the result and finally the oracle report price is sent to the blockchain. After the information is within the closed blockchain system, it can be used in a variety of ways, typically by triggering a smart contract. An oracle may have both on-chain and off-chain components, i.e. the infrastructure involved in the various steps can be inside or outside of a blockchain.

[54] Shayan Eskandari, Mehdi Salehi, Wanyun Catherine Gu, and Jeremy Clark. 2021. SoK: oracles from the ground truth to market manipulation. Proceedings of the 3rd ACM Conference on Advances in Financial Technologies. Association for Computing Machinery, New York, NY, USA, 127–141. DOI:https://doi.org/10.1145/3479722.3480994

Figure 5.1
Oracle functioning



Source: Shayan Eskandari, Mehdi Salehi, Wanyun Catherine Gu, and Jeremy Clark. 2021. SoK: oracles from the ground truth to market manipulation. Proceedings of the 3rd ACM Conference on Advances in Financial Technologies. Association for Computing Machinery, New York, NY, USA, 127–141

Oracles are essential parts of the DeFi ecosystem with a strong impact on the performance of DeFi and the prevention of fraud. It is thus important for oracles to provide credible and secure data. To achieve this, it is desirable oracles meet requirements of timeliness, accuracy, and transparency, as described below. Oracle updates should be frequent to keep information up to date. However, decisions on the frequency of updating should take into account market volatility, so that higher volatility would lead to higher frequency, preserving thus the need to maintain the cost-effectiveness of oracles. Self-evidently, the data reported by oracles should accurately reflect the requested information. It is important to preserve accuracy and achieve thus resilience to misreporting.

Finally, in the context of transparency, the governance of Oracles, and also the underlying information sets should be clearly understandable and traceable based on public sources.

An oracle's performance against these desired characteristics depends on how the various components of an oracle (as per the aspects captured in Figure 5.1) combine into a system. The joint implementation of the oracle components will determine the operational robustness of the oracle and its resilience to attacks. Depending on their strengths in these respects, oracles can be vulnerable to manipulation via various types of attacks which exploit specific oracle characteristics and thereby interfere with and manipulate DeFi applications.

5.3.2 Potential oracle manipulation

Oracles are subject to a broad set of security-related issues. Such vulnerabilities arise for a variety of reasons; including their criticality for the functioning of DeFi pricing mechanisms, their positioning at the intersection of on-chain and off-chain information, and the difficulties to track attackers in the pseudonymous DeFi ecosystem.

Different attack types arise in each of the steps depicted in Figure 5.1, necessitating in each case well-chosen mitigation mechanisms. For example, as various types of data sources can be used in the first step including databases, sensors, humans and smart contracts, it is advisable to use a combination of data sources – thereby reducing reliance on a single input.

For the data-feeders selection process it is also important to choose only credible sources. This process can be either centralised or decentralised. In the decentralised case, determination of oracle inputs can be based on voting, whereby token holders vote on the number of data feeders and on who these data feeders will be. Alternatively, the oracle design can be based on staking, whereby data feeders post collateral (stake) against the data they provide. In all cases, risks of manipulation remain. For example, within a voting-based governance framework, token ownership may be concentrated in the hands of a few individuals – thereby giving those individuals excessive influence on the selection of data feeders in the distributed selection process.

Table 5.2 presents the main types of manipulation or malfunction of oracles according to the research in the hackathon challenge, together with indicative mitigation measures. These examples have been derived from studying related research work[55] and from our research on a number of incidents related to oracle manipulation and malfunction - each of which resulted in investors losing significant amounts of crypto assets.

[55] Oracle Manipulation - Smart Contract Security Field Guide (scsfg.io)

Table 5.2. Types of Oracle manipulation/vulnerability and mitigation measures

Type of oracle / oracle components	Manipulation/Malfunction	Mitigation
On-chain	Spot Price Manipulation (on-chain data fetching)	average of time-weighted prices or delaying in order to earn time and correct any mistakes (expensive)
Off-chain	Off-chain Infrastructure bugs/errors	Secure Coding Practices are needed (e.g. OWASP[56])
Centralized	Data feeders credibility: single point of failure	move to decentralized data feeders selection
Decentralized	Low data feeders credibility: Freeloading (copy values without validation) and mirroring	need for incentives, staking (reputation effect)
Decentralized with governance scheme	An attacker can take control of enough tokens to pass proposals on the parameters of the scheme, such as the list of approved data feeders	Even distribution of tokens, non- specialized tokens for oracle governance

5.3.3 Case study of Oracle design - Chainlink

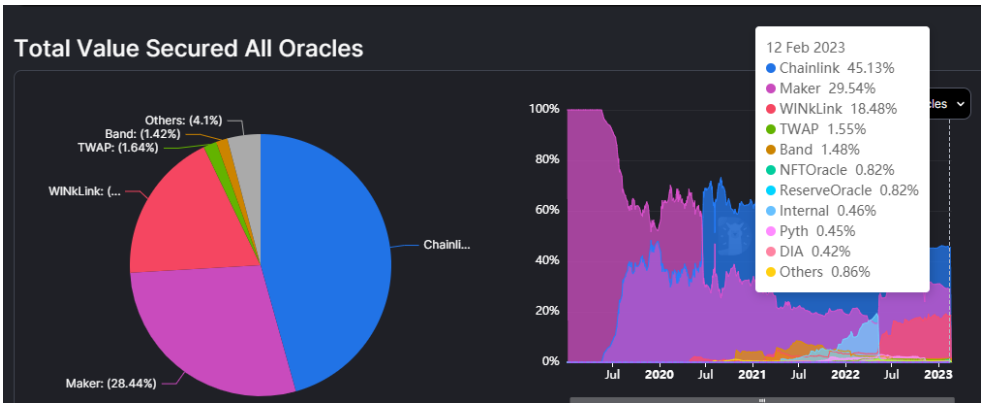
This subsection presents the business model of one of the most well-known oracle providers - Chainlink. It elaborates on the functioning of Chainlink and discusses the potential implications for risk. In order to gain a deeper understanding of oracle risks, three oracle incidents are examined.

Chainlink is one of the most well-known providers of oracles, with tens of billions of US dollars in Total Value Secured (TVS)[57] (see Chart 5.1). During the hackathon analysis, Chainlink maintained the dominant position as the most heavily used Oracle across the DeFi landscape. Chainlink seamlessly integrated hundreds of protocols from different DeFi categories, including the Aave lending protocol.

[56] The Open Worldwide Application Security Project (OWASP) is an online community that produces freely available articles, methodologies, documentation, tools, and technologies in the fields of IoT, system software and web application security (OWASP Secure Coding Practices - Quick Reference Guide | Secure Coding Practices | OWASP Foundation)

[57] Total value secured (TVS) refers to the aggregate total value locked (TVL) within all smart contracts that depend on a given decentralized oracle network's proper operations.

Chart 5.1
Distribution of TVS across Oracles



Source: DefiLlama

Architecture. Chainlink's architecture involves multiple layers of decentralised aggregation and security mechanisms integrated in the architecture and business model. Data is provided by various data sources, including both decentralised and centralised exchanges. Professional data aggregation firms known as "data feeders" (e.g. CoinMarketCap, CoinGecko, Tiingo) collect this raw market data and generate e.g. a volume-weighted average price of a crypto-asset combining data from selected crypto-exchanges. Weighting of different exchanges is based on certain objective metrics, e.g. market depth, latency, and spread - thereby smoothening out potential market anomalies. The data feeders make available their aggregated data typically for a fee via Application Programming Interfaces (APIs) and Service Level Agreements (SLAs), thereby establishing financial incentives to maintain and provide accurate data. The second layer of the architecture consists of nodes. Each node fetches data from various data feeders and responds with a median value. Trigger parameters are used for increasing the update frequency when there is market volatility. Finally, independent Chainlink nodes form decentralized Oracle networks which produce Oracle reports and store them in the smart contract.

The multilayer, decentralized architecture, the use of weighted averages of prices at the different layers and the use of SLAs are considered as security mechanisms that prevent a single point of failure and the impact of malicious or erroneous reported prices by single entities. They also give incentives to entities to provide accurate data in a timely manner.

Business model. Chainlink's native token LINK is used to pay node operators for their data provision and validation work, and for staking. Specifically, Chainlink's node operators set prices for price data requests. The price depends on:

(a) the fee (gas[58]) that the nodes must pay for answering data requests to the blockchain. This can get very expensive on certain blockchains. For example, the

[58] Gas refers to the unit that measures the amount of computational effort required to execute specific operations on the blockchain network. The gas fee is the amount of gas used to do some operation, multiplied by the cost per gas unit. In Ethereum mainnet, gas fees have to be paid in ether (ETH). Gas prices are usually quoted in Gwei, which is a denomination of ETH (1 GWEI=0.000000001 ETH).

same job may be charged 3 LINK per transaction on the Ethereum mainnet versus 0.15 LINK on Polygon.

(b) the fluctuations of LINK/ETH (or LINK/MATIC), which is a kind of "foreign exchange variance" in the business model of Chainlink. The node operator is paid in LINK, but the costs of the node are in ETH or the native token of other blockchains calling for data like MATIC.

Chainlink's node operators need to provide price feed updates irrespective of the gas price. So, for certain periods of time those jobs can run at a loss.

Partnerships. Chainlink is involved in various collaborative partnerships with a view to enhancing resilience. For example, they are collaborating with the Institute of Electrical and Electronics Engineers Computer Society Blockchain and Distributed Ledger Standards Committee (IEEE BDL) to develop global standards around the adoption of decentralized oracle networks.[59] They are also collaborating with Swift and several financial institutions so that they can provide a single point of access to multiple networks using existing infrastructure.[60]

5.3.4 Case studies of oracle manipulation

In order to gain a deeper understanding of oracle risks, the hackathon participants elaborated on incidents involving oracle manipulation or malfunctioning. They focused on explaining the vulnerabilities that were exploited, the impacts of these exploitations and the lessons learned. They chose three case studies that revolve around the malfunction or manipulation of oracles, each with different outcomes, lessons learned, and mitigating actions.

5.3.4.1 Case study A: Hardcoded price of LUNA in Chainlink leads to false price exploitation in Blizz platform (2022)

In May 2022, during the UST[61] failure, the Chainlink oracle was used to feed LUNA's on-chain price to value LUNA collateral pledged on various DeFi platforms. Price feeds hardcoded the price of LUNA at USD 0.10 and stopped updating LUNA's price when the Terra ecosystem was suspended. The price of LUNA then dropped below the hardcoded USD 0.10 to USD 0.01 and eventually to zero, but this was not reflected on platforms that were using the hardcoded price, such as Blizz Finance. As such, people who noticed the flaw were able to buy large amounts of LUNA at the market price (USD 0.01), post it as collateral and borrow funds from Blizz at a value of USD 0.10.

Lessons learned: the source of the problem in this case study was an error in the off-chain part of the oracle infrastructure. The oracle functioned as intended but its

[59] Chainlink Collaborates With IEEE BDL to Develop International Oracle Standards

[60] Swift unlocks potential of tokenisation with successful blockchain experiments | Swift

[61] Terra system stablecoin that could be exchanged by a floating quantity of the sister token LUNA

design needed a more secure development practice. Such issues in the off-chain parts of oracles could be alleviated by best coding practices, such as OWASP (see section 5.3.3) and robust testing including a “malicious actors simulation”, before going into production. Furthermore, auditing and independent review could help oracles to ensure that they remain protected against new threats.

5.3.4.2 Case study B: The Synthetix sKRW incident (2019)

Synthetix is a DeFi derivatives platform which supports trading in various crypto-assets. In order to estimate prices for the supported assets and use them for its services, Synthetix (at the time of the studied incident, i.e. in 2019) was using a custom oracle implementation which aggregated multiple related price feeds involving off-chain price data providers. The aggregated prices were then fed via smart contracts into the on-chain ecosystem at fixed intervals. Synthetix clients then used these prices to take long or short positions against supported assets.

On 25 June 2019, one of the price feeds that Synthetix relied upon mis-reported the price of the Korean Won to be 1000 times higher than the true rate. Due to a lack of necessary controls in the oracle pricing system, this price was accepted by the system and was posted on-chain - thereby allowing a trading bot to quickly trade in and out of the sKRW market,and reaping a profit of over USD 1 billion.

Lessons learned: similarly to the previous case study, while on-chain aggregation and price reporting worked correctly, the incident arose from the inadequate design of the off-chain component. The incident is thus an example of an off-chain component malfunction affecting on-chain oracle data feeds.

5.3.4.3 Case study C: Chainlink node attack via gas fee (2022)

In this case, an attacker began sending valid price feed requests to Chainlink node operators, which resulted in operators having to pay a lot of gas fees for responding (Ethereum transaction fees). The attacker exploited the high fees on the network by driving up the gas costs of these nodes and then minting Chi tokens in the decentralized exchange aggregator 1inch. Chi is a tokenized form of gas and was normally used to defray high gas costs. It is also the most liquid gas token, specifically it was liquid in Mooniswap - 1inch's Automated Market Maker protocol. It seems that this was the reason that Mooniswap was chosen for the attack, with the attacker then able to capitalize on the minted Chi tokens gained and sell them for ETH. The attacker exploited the way in which nodes respond to queries and involved the use of a token tied to network transaction costs.

In addition to the attacker benefiting as explained, a broader disruption was caused leading to additional risks. During the period of the attack, nine node operators experienced a draining of their ETH balances, meaning that they could no longer fulfil price requests during the attack period. This diminishing of the pool of node operators lasted for approximately two hours. However, the unaffected node

operators could continue feeding data - so the impact of the reduced population of node operators was not high on this occasion.

This incident can be categorised as a spam attack, which, overall, had three consequences: it drained resources of certain nodes, caused the unavailability of these nodes in the oracle mechanism, and allowed the attacker to generate profits.

After the strange pattern of the gas token being minted was noticed and reported, the incident was handled by the Chainlink security team via a process known as “whitelisting”. In the whitelist approach, node operators rank the most valuable data requests, coming from the most active DeFi protocols — for example, Aave and Synthetix — and fulfil only their requests while blocking all other, non-whitelisted, requesters. However, whitelisting was considered a temporary solution, whereas for a permanent solution, Chainlink would need to find "common ground with actual data consumers"[62]

Lessons learnt: the attack was based on a vulnerability of the oracle mechanism which was related to the possibility for the attacker to easily send multiple price requests combined with the specific gas mechanism that allowed the minting of specific tokens. A careful assessment of the intermediate steps of an oracle mechanism, such as the prioritization of price requests and the vulnerability assessment of the gas mechanism may help in preventing similar attack scenarios. Auditing and independent review may also help in this direction.

5.3.5 *Policy reflections on oracles*

In considering the potential policy relevance of Oracles, and the options that may exist to reduce associated risks, it is worth to consider what analogies exist for the function of DeFi Oracles in the traditional financial sector. Table 5.2 depicts these analogies.

Table 5.2. Similarities between DeFi Oracles and Oracles in the traditional finance

DeFi	Traditional finance	Observations
		Similar financial stability risk dynamics in algo trading to DeFi / Smart Contracts. Automated asset liquidations create downward spirals.
	Bloomberg + algo trading	
		Similar concerns on manipulability of LIBOR and Oracles.
	LIBOR + loan contract	
		Similar concerns on financial stability risk dynamics (fire-selling of bonds when firms lose investment grade status)
Oracle + Smart Contract + information from websites (e.g. https://coinmarketcap.com/) and other data sources		Similar concerns about vulnerability to manipulation and bad incentives - based on the experience of AAA ratings “bought” on subprime during the financial crisis.
	Credit Rating + investment fund rules	

This perspective of comparing the challenges associated with DeFi to the challenges of maintaining robust informational benchmarks within traditional finance provides a starting point for thinking about potential avenues to increase DeFi resilience. In this

[62] https://www.theblock.co/post/76986/chainlink-nodes-attack-eth

context, it is notable that certain challenges and failures have arisen within traditional finance regarding the operation of informational benchmarks, such as LIBOR and credit ratings. The challenges faced within the traditional finance benchmarks have arisen as a result of perceived sub-optimalities in their governance and design. In turn, these sub-optimalities have resulted in opportunities for occasional benchmark manipulation – and thereby a reduction in market trust. To address the problems that have arisen in these respects, the market and rule-makers have collaborated to alter the governance and design of benchmarks with a view to demonstrate transparently their robustness and subject them to regulation[63] and adequate supervision[64] by public authorities as for the case of credit ratings[65] to provide an additional safeguard.

Within the world of DeFi, similar efforts to increase trust in Oracles by demonstrating robust governance structures that protect market actors against the risks of market manipulation can be beneficial. Options to achieve this goal include the following:

- Secure Coding Practices – oracle vulnerability to manipulation can be reduced by using best IT development and security practices, e.g. as suggested in Open Worldwide Application Security Project (OWASP)

- Use of Robust Risk Management Frameworks in compliance with international security regulations and guidelines (e.g. DORA, BIS[66] and NIST[67]). Vulnerability assessment of oracles should be integrated in risk management processes.

- Transparent auditing and independent review could help oracles to ensure that they remain protected against new threats.

- Decentralised design structures – Increasing decentralization in the various steps of oracles increases resilience to manipulation by colluding entities.

- Regulation & standardisation of oracles could help in protecting the financial ecosystem from manipulation and attacks. Research work of Zetzche et al. (2020)[68] which proposes embedding regulatory requirements (e.g. transparency, disclosure, compliance) into the design of DeFi protocols could also be considered in this regard.

---

[63] Benchmark Regulation (2016)

[64] ESMA as Supervisor of Benchmark Administrators (2022)

[65] ESMA is the single direct supervisor of Credit Rating Agencies (CRAs) within the EU since 2011.

[66] Bank for International Settlements (bis.org)

[67] Cybersecurity Framework | NIST

[68] Zetzsche, Dirk Andreas and Arner, Douglas W. and Buckley, Ross P., Decentralized Finance (DeFi) (September 30, 2020). Journal of Financial Regulation, 2020, 6, 172–203, Available at SSRN: https://ssrn.com/abstract=3539194 or http://dx.doi.org/10.2139/ssrn.3539194

6 Concluding remarks

The Defi hackathon was successful in analysing more detailed data and developing expertise in central banks as they seek ways to understand, analyse and monitor the DeFi phenomenon. Two DeFi protocols, Aave and Sablier, and in particular direct blockchain data were examined, allowing hackathon participants to explore various types of related relevant information in this context. Furthermore, the DeFi hackathon was valuable in developing analytical and collaboration skills of the participants coming from various central banks who did not know each other beforehand and during the event worked in diversified teams in terms of skills.

While blockchain data is openly accessible, it remains relatively opaque and cumbersome to process and analyse. For the DeFi hackathon, a blockchain indexer was used in preparation to the event, to extract and reformat the data for the above-mentioned two protocols distributed across several blockchains. One of the main challenges the participants faced in analysing such processed blockchain data was the lack of comprehensive information about the content of the prepared tables, the variables in the tables and the linkages between them.

Constrained by the 48-hour-timeframe of the hackathon, the participants did not strive for comprehensiveness but for demonstrating the possibilities and potential in granular analyses performed directly on raw data, rather than relying on third party data provision.

The indicators developed concerning Aave focused on gaining insights into borrowers, loan features (including flash loans), and deposit pools across various blockchains. Flash loans obtained special attention by the hackathon participants as one of the most creative tools in the DeFi industry.

The analysis of Sablier data was based on a simpler data model and led to the exploration of clustering techniques to provide insights on the use of streaming payments. It was noted that the use of the protocol varied among blockchains in accordance with the different real use cases.

The study of oracles revealed insights into the challenges around establishing, maintaining, and monitoring trust and efficiency in data feeds in a decentralised and digitised financial system.

These main outcomes provide a contribution for supervisors and central bankers in their oversight or monitoring capacities to continue work on DeFi towards a deeper understanding of this novel space.

Although the size of DeFi is still not comparable to corresponding traditional finance activities, it is worth analysing further, including innovations on DLT platforms that aim at better scalability and interoperability. New ecosystems could emerge based on recognised standards among players, and these may well require different sets of regulatory and supervisory actions going forward.

7 References

1. Auer R., Haslhofer B., Kitzler S., Saggese P., Victor F.,"The Technology of Decentralized Finance (DeFi)" 19 January 2023, BIS Working Papers No 1066, 13-14.

2. Chanelle Duley, Leonardo Gambacorta, Rodney Garratt and Priscilla Koo Wilkens, The oracle problem and the future of DeFi, BIS Bulletin, Sep. 2023

3. Gudgeon L., Perez. D., Werner S., Knottenbelt W. J. "DeFi Protocols for Loanable Funds: Interest Rates, Liquidity and Market Efficiency"arXiv:2006.13922v3 (q-fin.GN)15 Oct, 2020, 3.

4. Igor Makarov and Antoinette Schoar, Cryptocurrencies and Decentralized Finance, BIS Working Papers No 1061, Dec. 2022

5. Kaihua Qin, Liyi Zhou, Pablo Gamito, Philipp Jovanovic, Arthur Gervais An empirical study of DeFi liquidations: incentives, risks, and instabilities; IMC '21: Proceedings of the 21st ACM Internet Measurement Conference November 2021Pages 336–350https://doi.org/10.1145/3487552.3487811

6. Shayan Eskandari, Mehdi Salehi, Wanyun Catherine Gu, and Jeremy Clark. 2021. SoK: oracles from the ground truth to market manipulation. Proceedings of the 3rd ACM Conference on Advances in Financial Technologies. Association for Computing Machinery, New York, NY, USA, 127–141. DOI:https://doi.org/10.1145/3479722.3480994

7. https://docs.aave.com/developers/guides/liquidations

8. https://github.com/aave/aave-v3core/blob/master/techpaper/Aave_V3_Technical_Paper.pdf

9. https://blog.wehodl.finance/aave-v1-v2-and-v3-a-comparison-of-three-generations-of-defi-lending-9573eec663e2

10. Oracle Manipulation - Smart Contract Security Field Guide (scsfg.io)

11. https://www.theblock.co/post/76986/chainlink-nodes-attack-eth

12. Garratt, R, and C Monnet (2023): "An impossibility theorem on truth-telling in fully decentralized systems", BIS Working Papers, no 1117, August.

13. Qin, K., Zhou, L., Livshits, B., Gervais, A. (2021). Attacking the DeFi Ecosystem with Flash Loans for Fun and Profit. In: Borisov, N., Diaz, C. (eds) Financial Cryptography and Data Security. FC 2021. Lecture Notes in Computer Science(), vol 12674. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-64322-8_1

14. Xie, Y., Kang, X., Li, T., Chu, CK., Wang, H. (2022). Towards Secure and Trustworthy Flash Loans: A Blockchain-Based Trust Management Approach. In: Yuan, X., Bai, G., Alcaraz, C., Majumdar, S. (eds) Network and System Security. NSS 2022. Lecture Notes in Computer Science, vol 13787. Springer, Cham. https://doi.org/10.1007/978-3-031-23020-2_28

15. Zetzsche, Dirk Andreas and Arner, Douglas W. and Buckley, Ross P., Decentralized Finance (DeFi) (September 30, 2020). Journal of Financial Regulation, 2020, 6, 172–203, Available at SSRN: https://ssrn.com/abstract=3539194 or http://dx.doi.org/10.2139/ssrn.3539194