Risk Assessment of Digital Assets – Insurance Applications in Cryptocurrencies and NFTs

Roberto Delgado Ferrezuelo
ro3187de-s@student.lu.se

Trygg-Hansa
Lund University

Academic supervisor: Paul Stankovski Wagner

Industry supervisors: Fredrik Thuring and Erik Rasmusson

Examiner: Erik Larsson

June 13, 2023

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

The aim of the project is to develop a framework for an insurance policy for digital assets. The project comprised several stages, starting with the identification of risks associated with these assets. Policyholders were then categorized into two groups based on a predefined rating factor. Subsequently, data about previous thefts was gathered, two different approaches were explored. In the first approach, patterns in cyberattacks targeting the selected assets are detected based on the risks previously identified, and a Python script is developed to automate the whole process. However, practical limitations surfaced, impeding the success of this approach, therefore a decision was made to pursue an alternative strategy for data collection, involving manual retrieval from trusted sources.

The collected data was used to fit various statistical distributions, enabling the prediction of the probability of policyholders experiencing loss of their digital assets. Additionally, a mathematical model was developed to provide a one-step forecast of the tokens prices, incorporating variables such as the floor price and token rarity. These predictions formed the basis for estimating the expected losses on a daily basis, which are utilized to calculate the company's potential liabilities.

A real-world scenario was simulated, where a user takes out an insurance policy to cover the risks for one of their items during the month of April 2023. The lump expected losses are calculated at the end of the month, assuming a daily exchange of money between both parties, and the final value is compared for both groups of policyholders.

Furthermore, an alternative approach was proposed, introducing a supplementary variable to the model based on the policyholder's behavior. The findings demonstrated consistency, as the expected losses fell within a reasonable range, with higher premiums for the riskier group of policyholders. However, it was observed that at a certain point, the perceived risk became higher for the safer group. Therefore, it is suggested to dynamically adjust the calculated parameters for the statistical distributions, taking this factor into account.

This pricing model serves as a preliminary framework for insurance policies and can be further refined through iterative improvements by incorporating historical claims data gathered by the insurance company. Ultimately, these enhancements aim to develop a comprehensive insurance policy offering.

This Master's thesis was written in collaboration with Trygg-Hansa through the Faculty of Engineering at Lund University.

Popular Science Summary

Back in 2008, a person or group of person under the pseudonym Satoshi Nakamoto unveiled a groundbreaking concept that would revolutionize the world: Bitcoin. Since then, Bitcoin has become an enduring buzzword, captivating the global stage. Its core objective was to establish a decentralized electronic cash system, liberating society from the grip of powerful entities that dominate financial services.

However, we won't delve deeply into Bitcoin itself. Instead, this paper focuses on the new applications that have emerged from the technology behind it —blockchain. If you haven't heard of the concept of blockchain, it can be described in simple terms as a decentralized and immutable ledger. In this new space, decentralization is the key.

As people began exploring Bitcoin and its possibilities, new applications started to emerge. Some of these applications simply run on top of the Bitcoin blockchain, while others decide to create their own separate blockchain. The most prominent example of the latter is Ethereum, which has expanded the possibilities of this peer-to-peer electronic cash system far beyond financial services. The properties of immutability and decentralization can be extended to other industries such as art, healthcare, supply chain, and more.

In particular, the art industry has been instrumental in the creation of assets that are the subject of study. Artists' minds are constantly brimming with creative ideas, always seeking ways to differentiate their art and come up with unique concepts never seen before. What if we combine that frenetic way of thinking with the expertise of a tech entrepreneur? That's exactly what happened in 2014 when Kevin McCoy and Anil Dash joined forces to create a new form of digital art —one that allows tracking the history of ownership, also known as provenance, which is highly valued by collectors.

Over the years, this concept continued to evolve, eventually leading to what we now know as NFTs (Non-Fungible Tokens). While the roots of NFTs can be found in the digital art industry, their potential applications go far beyond that. Many skeptics consider these digital assets as a Ponzi scheme, and you may have seen people making fun of collectors on social media by posting a screenshot of one of the famous apes accompanied by a comment like, "Sorry I stole your NFT". Somehow this is what this project is about, but for real thefts, not simple screenshots, of course. This paper will explore what these assets are, how they can be used, and

why taking a screenshot doesn't make you the real owner.

At the time of writing, the lowest price to buy one of these apes, for example, is around $80,000. The problem lies not in the price itself but in the concerns of users who want to join this community. While some can afford to pay $80,000 due to being avid fans of the collection, such as celebrities like Paris Hilton or Jimmy Fallon, the space is plagued by cyberattackers targeting less experienced users. This diminishes the attractiveness of the user experience and damages the reputation of the space. This is where insurance can come into play, offering the protection users need and providing them with advice on how to avoid being targeted.

For insurers, the challenge arises when creating new policies because they typically rely on past observations and historical data to predict the cost of claims for the company. In this relatively new space, where we are still striving to fully understand the technology, such extensive information is lacking. So, what can insurers do to dip their toes into this field? I believe there is no perfect answer to that question, so adapting to the limited information available seems to be the only option.

The NFT market is now sizable, with a variety of assets that could be covered by insurance policies. However, when we think about the future and the ongoing projects related to the metaverse and alternative realities, we realize that the range of insurable digital assets will likely experience a significant expansion. Therefore, it is interesting for insurers to start exploring this space, offering initial policies that will help gather valuable information and iteratively improve the initial models until a complete final version is achieved.

Acknowledgments

First, I would like to thank all the people supervising my work during all these months, to Fredrik and Erik for helping me with everything that was in their hands, to Paul for his good advice and to Pepe for, once again, making things a lot easier for me.

To my family as always, especially in this difficult year, to my friends back home and to my new international friends, just having met them has made the whole experience worthwhile.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Background and Motivation

The market for crypto and NFTs have boomed in the latest few years making digital objects valuable and sometimes extremely valuable. However, the mechanisms for protecting the digital object or its value have not seen the same development. The conceptual meaning of insurance is to allow individual agents to purchase an object associated with a certain risk and transfer that risk to an insurance company for a set premium. The legal agreement between the insurer and the policyholder outlining the specific circumstances in which the insurer will provide coverage for losses resulting from a predefined set of perils in exchange for the premium, is called the policy. By bundling together many different risks the insurance company reduces its risk volatility compared to the individual agents and makes a profit by setting the correct premium. This project aims at finding and analyzing similar insurance analogies in the crypto space and particularly for NFTs.

Several research papers have already delved into the current insurance landscape for digital assets. For instance, Adam Zuckerman's article [1] provides valuable insights and recommendations for insurers, highlighting potential areas of opportunity. However, these papers primarily adopt an informative approach. The objective of this project is to take a further step by establishing a quantitative approach, bridging the gap created by the scarcity of data.

By thoroughly exploring diverse information sources and meticulously documenting all available details regarding past theft incidents, it becomes feasible to construct a straightforward pricing model, categorizing policyholders based on their respective risk levels.

1.2 Aim and Scope

The potential of blockchain technology in the insurance industry is vast. This industry is notorious for its heavy reliance on paperwork when it comes to claim settlements. This often leads to human errors and lengthy resolution processes, ultimately resulting in financial losses for companies. Blockchain technology has the capability to revolutionize the entire value chain by providing a faster and more efficient method for settling claims. However, the scope of this project is

focused on exploring the insurability of digital assets, with a particular emphasis on NFTs.

The project aims to address four key questions:

1. What are the primary risks associated with digital assets?
2. How can the risk levels of policyholders be evaluated?
3. What is an appropriate pricing model, and how does it perform?
4. What data is required to train such pricing model, and where can it be sourced?

Of these questions, the final one presents the main challenge for this project. The text will discuss the two chosen approaches, examining their limitations, the degree of success achieved in their implementation, and their overall reliability.

1.3 Methodology

To address these four questions, an initial plan was established, and adjustments were made along the way.

The first and most time-consuming task involved gaining a comprehensive understanding of NFTs and their underlying technology. A thorough literature review was conducted, covering all relevant aspects. Since the NFT space is relatively new and technical terminology can sometimes hinder comprehension, it was essential to establish a strong knowledge foundation. Actively engaging with the space and immersing oneself in it proved to be an effective approach for gaining insights into its various components, providing the necessary groundwork to address the first question.

To identify the risks associated with these assets, reports from specialized blockchain analytics firms such as Elliptic or Chainalysis were utilized. Additionally, real-world cases were examined to explore common vulnerabilities and patterns in the attacks.

Once the risks and common attack vectors were identified and understood from a technical perspective, the focus shifted to evaluating the risk levels of policyholders, addressing the second question. The identification of risks revealed that the type of wallet used plays a crucial role in determining a user's exposure. Therefore, different wallet solutions in the space were explored, and a comprehensive analysis of the most common ones was conducted. This analysis enabled the categorization of individuals based on the risk level associated with the wallet they use for transactions.

Considering the time constraints and challenges involved, the fourth question was addressed before the third. Initially, an algorithm was developed to detect transactions following the identified patterns and store their hash along with other relevant information. The goal was to generate a large database with theft records for statistical significance. However, after investing substantial time in refining the code to filter out invalid transactions, it was decided to abandon this approach. The computational operations became increasingly large, and the time required to fetch blockchain blocks made it impractical for the project's duration and available

resources. Instead, reported theft cases were individually inspected, and relevant information about compromised wallets was manually stored in a data table. Multiple sources were utilized to obtain this information, which will be mentioned in the subsequent text.

Lastly, the optimal pricing model was determined. The initial idea was to find an analogy to other non-life insurance products using Generalized Linear Models and tariff cell analysis. However, due to the non-deterministic nature of NFT prices and the scarcity of available data, a different but similar approach was adopted. Instead of modeling claim severity and frequency to estimate the pure premium, the model utilized the probability of insured individuals losing their NFTs and a one-step prediction of the item's price. Commercial software, specifically Python and MATLAB, were employed for developing the model. First, the data gathered on previous thefts was modeled using a Python package. Then, a separate mechanism to price the NFTs was developed in MATLAB, considering factors such as the floor price and token rarity.

To address the second part of the third question, a Monte Carlo simulation was conducted. Different scenarios were simulated for two groups of policyholders. The expected outcome was to observe higher expected losses for the riskier group, suggesting higher premiums compared to the other. The Monte Carlo simulation also allowed for a comparison against analytically derived expressions to validate their accuracy.

Chapter 2
Web3

The World Wide Web has evolved since its inception, going through different stages. The first iteration, coined as the "Web 1.0" by Tim Berners-Lee in 1989, consisted of static websites owned by companies that provided a better access to information for users, but it lacked of interactivity.

In the second iteration or, "Web 2.0", there is a shift towards a more participatory network in which bidirectional communication flows are established, leaving behind the "push model" used in the Web 1.0. One of the main features of the Web 2.0 is the social networking, which allowed people from different parts of the world to be connected. The main problem of this iteration is the dependency generated in users that rely on centralized entities to act honestly as they have control over most of the internet infrastructure and users data.

The third iteration, commonly referred to as "Web 3.0" or "Web3", was introduced by Ethereum co-founder Gavin Wood as a solution to this problem. It leverages technologies like blockchain to distribute network access in a more equitable manner. According to the Ethereum description [2], Web3 is characterized by core principles such as decentralization, permissionless access, native payments, and trustlessness. This iteration is still under development and can also be interpreted from a machine-readability perspective, where data is represented in a format that machines can process.

As mentioned in the Twitter post [3], the three stages are commonly described as follows: "Web 1: Read, Web 2: Read-Write, Web 3: Read-Write-Own". Figure 2.1 provides a visual representation of the different iterations.

2.1 Blockchain

Blockchain is one of the underlying technologies that powers Web3, eliminating users' dependence on large corporations acting as intermediaries. The speaker in [4] describes it as the technology that enables a shift from the "Internet of information" to an "Internet of value", a democratized version where the asymmetry derived from the majority control of the global infrastructure by these authorities is reduced.

This concept was first implemented in 2008, when the whitepaper in [5] was published by an anonymous person or group of persons under the pseudonym Satoshi Nakamoto, which introduced to a new peer to peer electronic cash system called Bitcoin. Blockchain is the technology behind Bitcoin and it can be described

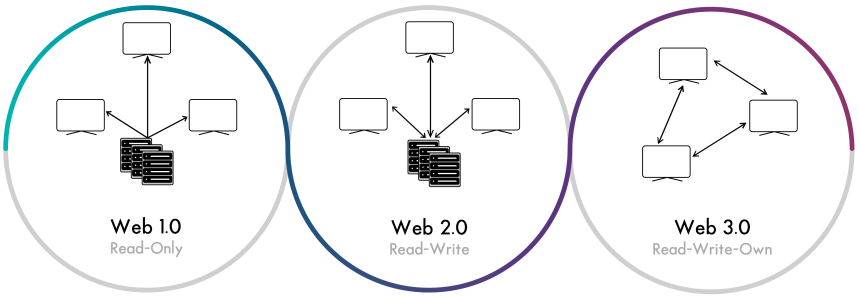

Figure 2.1: World Wide Web iterations

as an immutable distributed ledger where transactions are anonymously recorded. The anonymity is achieved by using public-key cryptography to generate a key pair that identifies the participants. The keys are stored in wallets and they can be non-deterministic, when private keys are generated randomly, or deterministic, which are commonly generated using the standards introduced in Bitcoin Improvement Proposals (BIPs) 39, 32 and 44 [7, 8, 9]. Public keys are derived from the private key using a cryptographic hash function such as the Elliptic Curve Digital Signature Algorithm. The public key is later hashed to create the public address. The private key is used to sign transactions, proving ownership of the assets being transferred, therefore it is kept in a secure location, while the public address is shared with the rest of the participants in the network so that they can send transactions to the wallet associated with the private key.

There are different blockchains and each of them is run by computers provided by volunteers around the world which are called nodes, each of these nodes has a copy of the ledger and for a transaction to be validated they have to agree based on a set of rules. A combination of cryptography and game theory is applied to avoid what is called the Byzantine Generals Problem [6], a dilemma in which isolated participants have to agree in a common decision but there is no guarantee that they will act on the group's best interest. Using consensus algorithms it is possible to create a Byzantine fault-tolerant system, a system in which trust among participants is not necessary, since it is in their own interest to act for the benefit of the group. There are different algorithms, but the most widely used are Proof of Work (PoW) and Proof of Stake (PoS). Typically, blockchains have a native currency that is used by these algorithms to incentivize participants to maintain the security and integrity of the network.

To send a transaction, users need to sign a digital message using their private key with the recipient's public address as the payee. The transaction is then broadcasted to the network of nodes who verify its validity and bundle it with other transactions into a block. Each block includes a header with information such as the timestamp, a reference to the previous block (thus forming a chain), the Merkle root, which is a hash of all transaction's hashes included in the block, and other parameters that can vary in the different blockchains.

2.1.1 Consensus Algorithms

PoW concept was first implemented in Bitcoin, it consists on a competition among a group of nodes, that are called miners, in finding a solution to a complex mathematical problem where the first in solving it is financially incentivized, reaching in this way a consensus on the state of the network and preventing what is called the double-spending, when a user tries to spend the same asset twice.

To find a solution to the problem, miners need to find a value, the nonce, that when hashed together with the rest of the components of the block's header, the resulting hash is below a certain target value that is dynamically set based on the total computational power of the network, whoever finds this value has to broadcast it to the rest of the nodes and, if accepted, he receives newly generated coins, also called minted coins. This hash serves as a unique identifier of the block and this value will be used as input in the next block to find the new solution, thus linking the blocks with each other and making it very difficult to manipulate a block as it would imply redoing all the subsequent work. Users willing to participate in the competition need to provide vast amounts of electricity and computational resources, also called the "stake" according to [10]. The stake discourages miners to act dishonestly as they would need to control the majority of the network, what is called the 51% attack, something that is highly expensive as the size of the networks continue to increase, making it the most cost-effective option to act according to the established rules. One of the main problems of the PoW mechanism is the high energy consumption, with most of the energy sourced from fossil fuels. At the time of writing, the Bitcoin Energy Consumption Index in [11] shows an annual carbon footprint of 52.10 Mt CO2. There are different alternatives for this algorithm that can considerably reduce the environmental impact, the most widespread solution is PoS.

In PoS, there are validators instead of miners and blocks are said to be forged or minted. To participate in the PoS validation process, nodes lock up a required amount of cryptocurrency in a wallet as a stake. An algorithm determines the next validator from a pool of candidates based on a number of considerations such as the node's hash value, which, according to the post in [12] is usually calculated by signing some network-related parameter using the private key, the amount of coins staked or the number of days the coins have been staked. Once the node is selected, it validates the transactions to be included in the block and adds it to the blockchain, receiving a share of the block's transaction fees as a reward (no coins are minted in PoS). If the network nodes detect a fraudulent transaction in one of the blocks, the validator who forged that block can be penalized loosing some of the cryptocurrency staked (higher than the transaction fees), also known as "slashing" or "burn" [13]. The 51% attack is highly impractical as it would imply to take control of the majority of the staked tokens which can be really expensive, for example, in the Ethereum network it would imply spending more than $110 billions, and even so, according to [14], the community can still use social recovery to restore the original state of the network.

Since the reward in the PoS mechanism is proportional to the amount of tokens staked, validators cannot benefit from economies of scale unlike it happens in PoW, where miners group together to form pools. It can be seen in Figure 2.2 how a few

mining pools have control over a big part of the Bitcoin network. The absence of

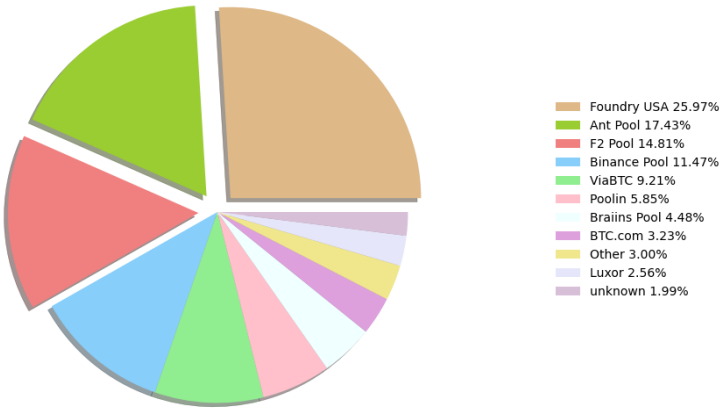

Figure 2.2: Hash rate distribution in Bitcoin network, Mar. 2022-Feb. 2023. Data source: [15]

.

such economies of scale and the lowering of the entry barriers, since there is no need to acquire expensive specialized equipment to participate in the validation, reduce the centralization risk. Moreover, as nodes are not competing to find the next block, one of the main benefits it brings to society is the energy saving. The Ethereum webpage in [16] shows a 99.988% reduction in the energy consumption since the Gasper (name of their PoS mechanism) implementation. However, it also has some setbacks, such as the possibility of validators forming oligopolies or the problem known as "nothing at stake". When new forks of the blockchain appear, the most profitable option for validators is to work on all of them as they do not incur additional costs, maintaining all these multiple versions can lead to vulnerabilities such as the double-spending attack mentioned above.

2.1.2 Private and Public Blockchains

Although the idea that fueled the growth and adoption of the blockchain technology is the elimination of the dependence on middlemen, the concept of private blockchain is starting to become widely adopted. Private blockchains do not align with the permissionless principle, leading to some reluctance from members of the public who view it as a fundamental characteristic, instead, the right to modify and add new entries into the ledger is reserved for only a few participants chosen by the entity running the network. The utilization of such blockchains has the potential to enhance the efficiency of antiquated processes in industries where the absence of competitiveness has hindered investment in process improvement, thereby enabling streamlining and optimization, Vitalik Buterin in [17] provides some interesting scenarios where it could be used as well as the advantages it

could bring to society, he also acknowledges the potential setbacks, such as public distrust and possible coercion.

There is another solution that lies between the two options discussed so far, namely consortium blockchains. Here, the permissions to read and write in the ledger are restricted to a set of nodes instead of a single organization. Big insurance companies such as Allianz are adapting their processes using this type of blockchain-based solutions to settle faster and more efficiently international motor insurance claims. In the podcast in [18], Bob Crozier, Allianz's current Interim Chief Data Officer explains how the company is using the modular blockchain framework developed by the Linux Foundation, Hyperledger Fabric, to improve the intercompany billing process, from claim creation to settle status involving its different Europe's subsidiaries. By using a consortium blockchain they significantly cut down their frictional costs as well as the time required in the claim processing while keeping the deterministic finality (the time it takes for a transaction to be added to the blockchain, thus becoming irreversible), as opposed to the probabilistic nature of the permissionless blockchains for which it has been necessary to develop new solutions that allow the creation of more scalable networks such as the use of rollups in a separate layer in Ethereum. Rollups bundle many transactions and submit them back to the main network, distributing fees among all participants while also increasing finality without sacrificing security or decentralization, as outlined in [19]. Transactions data regarding the claims reside in the blockchain while personal information about the clients is placed in a separate relational database guaranteeing their confidentiality.

2.1.3 Smart Contracts

As previously explained, the key pair in a blockchain is stored in wallets, not the native currency itself and it is the private key what give access to the funds which reside inside the blockchain. The way funds are stored vary across the different networks, for instance, Bitcoin utilizes Unspent Transaction Outputs (UTXOs), while Ethereum employs account balances to keep track of cryptocurrency holdings.

There are different ownership mechanisms to regulate the assets spending, apart from public keys, they can be owned by scripts specifying a set of conditions under which they can be accessed. Ethereum developed a low-level bytecode language, Ethereum Virtual Machine (EVM), which builds on and extends the capabilities of Bitcoin's scripting language. According to the Ethereum whitepaper in [20], the EVM adds turing-completeness, value-awareness, blockchain-awareness, and state, thereby completing Bitcoin's programming language.

One of the most important features implemented using these added functionalities were the smart contracts. On Ethereum, smart contracts are distinct from Externally Owned Accounts (EOAs) in that they are governed by a piece of code rather than a private key. Smart contracts can interact with each other as well as with EOAs by encoding messages with the associated address as the receiving party of the transaction. They use the data contained in the message as input and translate it into opcodes, each of which corresponds to a specific action EVM can perform. The amount of gas consumed during the execution of these actions

varies depending on the complexity of the task. To cover the computational effort required for each action, a dynamic price must be paid for each unit of gas consumed. This price varies according to the current network congestion, meaning that during times of high demand, the cost of gas will increase to incentivize miners to prioritize transactions with higher gas prices, ensuring the stability of the network. Smart contracts are written in high-level programming languages such as Solidity or Vyper and deployed in the network paying the corresponding fees. When called, they are compiled into bytecode that can be executed by the EVM, determining the state transition of the network based on the logic programmed into the smart contract. According to [21], Ethereum can be viewed as distributed state machine governed by the rules defined by the EVM instead of a distributed ledger.

Smart contracts are the base of the assets for which the policy framework is being developed. They also bring many exciting opportunities to the insurance industry by enabling a shift from the traditional business processes to a new value chain where most of the manual tasks can be automated achieving faster and more accurate results, some of the main benefits and examples of the current insurance landscape will be provided in a later section.

2.2 NFTs

NFT stands for Non-Fungible Token and unlike cryptocurrencies like Bitcoin or Ethereum, they can not be swapped for each other as their value is unique. They rely on smart contracts to create a tamper-proof record of ownership and link users to the specific asset they possess. First, an overview of their history will be provided, using the article in [22] as a reference for the chronology of events.

The emergence of the initial idea behind NFTs came a long time ago with the publication of the paper in [23] by Meni Rosenfeld in 2012. This paper discusses the idea of adding metadata to Bitcoin transactions creating a system by which coins can be traced back to their genesis state allowing a distinction to be made depending on the history of transactions associated with them. Limitations in the Bitcoin scripting language posed a challenge to their development which spurred the creation of more flexible platforms with advance features that allow the implementation of complex asset management functionalities.

In 2014, the artist Kevin McCoy partnered with the entrepreneur Anil Dash, aiming to find a solution to the problem of provenance in digital art, the partnership resulted in what is considered to be the first ever created NFT, Quantum [24]. After delving into the potential of blockchain technology, the duo opted to utilize the Namecoin network, one of the earliest forks of Bitcoin, to deploy the artwork. After years since its deployment and with the increasing popularity of NFTs, the artist made some promotional efforts for the artwork, and eventually, Quantum was sold for a whopping $1.47 million in a Sotheby's auction.

Namecoin was initially developed to extend the functionality of the Bitcoin network by enabling data storage, leading to the creation of decentralized services such as a domain name system. However, the network's unique features caused a surge of legal issues following Quantum's auction. It requires users to periodically

create new transactions to update the encoded output with the asset's associated data to prevent it from expiring, something did not happen with Quantum as it can be seen in [25], where the output has not yet been redeemed. After the renewal period expired, another user claimed ownership rights of Quantum and filed a lawsuit against the artist, asserting that he was the rightful owner of the artwork. However, the lawsuit was recently dismissed by a federal judge in New York [26] who determined that the plaintiff was in possession of a different NFT since Quantum was later minted (similar to cryptocurrencies, NFTs can be minted and burned) on the Ethereum network [27].

Following the mint of Quantum, a first concept of platforms that allowed the creation of digital assets started to appear. 2016 was a significant year for the internet of memes, among which Pepe the Frog stands out. It is a creation of the artist Matt Furie and despite its notorious association with the alt-right movement, it played a pivotal role in the development of the NFTs. Creators started to mint variations of the meme on the Counterparty platform, a protocol running on top of Bitcoin that allowed users to trade digital tokens, thus becoming the first examples of digital assets being traded and valued as a unique, collectible item. Since Bitcoin was not tailored to that specific purpose, new alternatives began to emerge, Ethereum being one of the most prominent.

The shift to Ethereum and subsequent boom in the market started with the project known as CryptoPunks, created by the software developers Matt Hall and John Watkinson in 2017. It consists on a collection of 10,000 unique pixelated AI-generated images each of them with different traits. They used the smart contracts capabilities to create a code that allowed the buy and sell of the different punks among the network participants. There exists two different versions of the collection, the original, also referred to as V1 CryptoPunks, was released on 9 June 2017, it had some flaws in the code that allowed buyers to get back the money they paid for the tokens, meaning that the seller did not get any ether (native currency of the Ethereum network) for the sale. Therefore it was decided to create a new contract where the bugs were removed, the V2 Punks, and airdrop (term commonly used to to refer to the distribution of free NFTs to a group of people) them to the original claimants.

Similar to the BIPs, Ethereum has its own Ethereum Improvement Proposals (EIPs). Those submissions proposing a change related to the token ecosystem can become an Ethereum Request for Comment (ERC) if accepted by the community. ERCs provide a consistent interface for tokens, and the creation of CryptoPunks laid the groundwork for the now widely adopted ERC-721 standard, which has become the de facto standard for NFTs.

2.2.1 ERC-721 Standard

As stated in the Larvalabs (company founded by Hall and Watkinson) webpage in [28], the tokens did not fully conform to any existing standards, although they closely resembled an ERC-20 compliant token. They added some extra functionalities to enable the buy and sell of the tokens and created their own marketplace.

The ERC-721 standard, authored by William Entriken, Dieter Shirley, Jacob Evans, and Nastassia Sachs, offers users a smart contract template that enhances

network interoperability through a common interface that developers can adapt to their unique requirements. The specifications for this standard can be found in a Github repository in [29]. It is inspired by the ERC-20 standard, which was the first implemented standard, and it addresses some of its limitations, introducing a more complex interface that includes functions for creating, transferring, and querying unique tokens. The pair `(contract address, uint256 tokenId)` serves as a unique identifier for each token in a collection. The `contract address` refers to the smart contract where the collection is deployed, while the `tokenId` variable denotes the unique identifier of the item within the collection. Some of the common and most important functions typically included in ERC-721 smart contracts are the following:

1. `ownerOf(uint256 _tokenId)` - returns the owner of an NFT.
2. `balanceOf(address _owner)` - amount of tokens held by an owner.
3. `safeTransferFrom(address _from, address _to, uint256 _tokenId, bytes data)` - transfers the NFT only when called by the owner, an authorized operator or approved address and confirms that the address `_to` is capable of receiving the token.
4. `transferFrom(address _from, address _to, uint256 _tokenId)` - transfers the token, but the user is responsible for checking that the address `_to` is capable of receiving the token.
5. `approve(address _approved, uint256 _tokenId)` - approves another address to transfer the given token ID.
6. `setApprovalForAll(address _operator, bool _approved)` - sets or unsets the approval of a given operator to manage all the message sender's assets.
7. `getApproved(uint256 _tokenId)` - gets the approved address for a token ID, or zero if no address is set.
8. `isApprovedForAll(address _owner, address _operator)` - checks whether an operator is approved by a given owner.
9. `tokenURI(uint256 _tokenId)` - returns the URI with the token's metadata for a given ID.

In recent years, the Enjin development team, creators of a blockchain-based platform for gaming, have been working on an enhanced token standard known as ERC-1155. This standard builds upon the previous ERC-20 and ERC-721 standards, allowing for the creation of semi-fungible assets (SFTs), assets which possess some of the unique characteristics of NFTs, while also offering a degree of interchangeability.. Although the ERC-721 standard remains the most widely adopted option, the ERC-1155 standard provides several benefits in terms of scalability and space efficiency. Unlike ERC-721 contracts, which require a separate contract for each type of asset, ERC-1155 contracts enable multiple token IDs, each representing a distinct asset type, to be stored in the same contract. This reduces the amount of space required to store information on the network. Furthermore, it allows for batch transfers, where multiple items can be transferred simultaneously, improving the network scalability by reducing congestion and fees.

2.2.2 NFT Metadata Storage

An important aspect of the NFTs, sometimes misunderstood is the difference between the actual token and the media file to which the token is referencing. Retaking the previous explanation of CryptoPunks, its creators embedded a hash of the composite image with all the punks [30] in the smart contract code, allowing users to verify the authenticity of the tokens being bought. It existed some controversy around the index corresponding to each token in the composite image as it was not specified how they are sorted, from top to bottom, left to right..., to clarify it they published in their webpage a separate image of each token with the corresponding ID, however this meant that Larvalabs had the control over which index belonged to each asset and, as in many other scenarios in the crypto space, centralization is not universally embraced by users. In 2021, a Twitter post [31] announced that they decided to move the images and attributes on-chain, something that is not always feasible due to size limitations as it will now be explained. This example illustrates the distinction between the content being acquired and the token stored in the blockchain. In this case, the content is a 24 x 24 pixel image that is part of a larger composite image of 2400 x 2400 pixels. On the other hand, the token is a record stored in the blockchain that proves ownership of a specific item in the collection, identified by its unique ID.

There are various alternatives available for storing NFT metadata, but concerns have arisen about the safety of these solutions and their potential impact on market consolidation. Moxie Marlinspike, Signal founder, posted an article in [32], criticizing some of the aspects of NFTs which raised again a concern that has been existing in the space for a long time. He discussed how many of the top NFT collections store the metadata and the media file using centralized servers which can be easily accessible, allowing users to change the NFT's description, image, title, etc. Marlinspike went further and created his own NFT that displayed a different image depending on the IP or User Agent of the requester. This experiment highlighted the low credibility of collections that use centralized storage solutions. Additionally, he pointed out how his NFT was delisted from major marketplaces for an alleged "violation of some Terms Of Service" with the NFT automatically disappearing from his Metamask wallet due to its high dependence on APIs provided by large entities operating in the space, one of which was the marketplace that delisted his token.

These issues underscore the need for better solutions for storing NFT metadata that prioritize decentralization, security, and independence from centralized marketplaces. While centralized solutions may be more convenient, they come with significant risks, including potential loss of control and censorship. During these years there have been many improvements aiming to seek a solution for these concerns, the article in [33] provides a deep understanding on two classifications schemes based on how the NFT data is stored and its practical implications. In terms of risk management, the technical scheme takes precedence over practical considerations since it emphasizes specific details.

This classification scheme, also referred to as the "Michelin guide" in the Dom Hoffman's Twitter post in [34], categorizes NFTs on the Ethereum network into four groups and assigns a score to each. The lowest score is given to NFTs whose

smart contract returns a URI pointing to off-chain resources, which is the most common setup. Within this category, NFTs can be further divided into two groups based on whether the resources are stored in a centralized server or a decentralized file storage system. Figure 2.3 illustrates how a random user purchasing an NFT in this category can access the data. To purchase the token, the user sends the

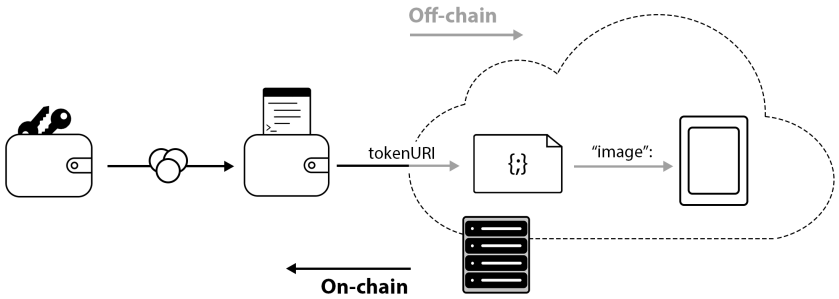

Figure 2.3: Metadata storage scheme for the lowest ranked NFTs in the Michelin guide

required number of coins to the smart contract address of the collection, with the token's unique ID determining its price. After the transaction is confirmed, a record is created with the user's address as the owner of the newly acquired token. The smart contract usually contains a link to a JSON file, accessible through the `tokenURI` function, that provides details on the token's attributes and points to the location of the media file.

Decentralized vs Centralized Storage

In the early days of NFT projects, centralized solutions were commonly used to store the metadata associated with the tokens. This approach involved storing all data in a single location, which provided the benefits of easy accessibility and centralized management. Additionally, centralized storage offered a high degree of customizability, enabling developers to tailor the storage solution to meet the specific needs of their project. As the NFT ecosystem progressed, decentralized storage solutions were created as an alternative to centralized storage. The Inter-Planetary File System (IPFS) emerged as a leading open-source project that provides a protocol for implementing this solution. IPFS is the most widely extended option to store NFTs' metadata nowadays. To better understand its components and how they interact with each other, the IPFS Camp Workshops in [35] and [36] are followed as well as the project site information in [37].

IPFS is not an implementation itself, rather it is a set of protocols designed to transfer and organize data in a decentralized manner. When an element is added to the system, it is split into smaller chunks, which can be of a fixed size or cleverly chunked (Rabin chunking), and then a Content Identifier (CID) is assigned

to each of these chunks. The CID is created by running the data through a hash algorithm and adding a metadata prefix that identifies the algorithm used, how the data is encoded, the version of the CID specification and the number-based encoding used for the string (in the CID version 0 most of this is implicit, with the resulting CID as a raw multihash with no added prefix). Once split, IPFS uses a set of specifications called InterPlanetary Linked Data (IPLD) to represent all that information and its relationships using a Merkle Directed Acyclic Graph (DAG). A DAG is way to represent data where nodes are connected to each other by their edges without forming a closed loop, Figure 2.4 shows the two currently supported layouts.

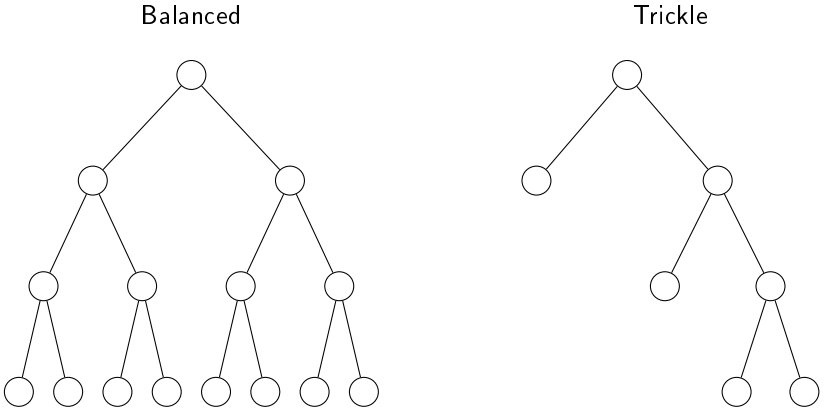

Figure 2.4: IPFS - DAG Layouts

In the Merkle DAG each node has a hash that is calculated based on its contents, therefore a slight modification on one of the chunks will propagate and create a complete different hash in the top node. Nodes are wrapped in something called the UnixFS wrapper, which includes metadata about the data such as its size, type, and other attributes. This allows IPFS to provide more granular control over how files, directories and their symbolic links are stored, accessed, and shared.

IPFS employs various mechanisms to locate a particular CID within the network. One such approach is Kademlia, which is a type of Distributed Hash Table (DHT) that maintains a record of peer IDs and the corresponding CIDs they can offer. Nodes can also use the Bitswap protocol, asking other members for CIDs and storing wantlists so that if they later receive the requested data can send it to the node who originally requested it. If a node do not have the computational resources required to use any of these mechanisms it can rely on an HTTP API, asking a delegated router to search for peers who have the CID on its behalf.

Once the peers in possession of the CID being searched are found, there are other systems used to distribute the content across the network of nodes. Apart from content routing, the Bitswap protocol can be used for this purpose. There are also nodes who offer HTTP Gateway APIs that allow other nodes not implementing any of the mentioned systems to fetch the data, being `ipfs.io` the official gateway maintained by the IPFS development team and the one used in

this text to refer readers to content stored using these protocols such as the image of Quantum.

IPFS has multiple implementations, each developed using different programming languages. For example, Kubo is an IPFS implementation written in Go, Nabu in Java, and iroh in Rust. This allow for IPFS to be used across various platforms and integrated into a wide range of applications. In addition to the different implementations, there are also related projects that build on IPFS's capabilities. One such project is Filecoin, which incentivizes users to rent out their unused storage space and creates a marketplace for storage, thereby improving long-term data availability. Another project is NFT.storage, which uses a combination of IPFS and Filecoin to provide long-term storage for NFTs.

Overall, it can be said that decentralized storage solutions provide a high degree of security as data is distributed across nodes instead of a single location. Reliability, as the data is addressed based on its content, therefore it can be easily checked whether it has been tampered. Accessibility, while data remains unchanged the identifier will continue to be the same, this prevents issues that can arise when hyperlinks become invalid or point to the wrong resource. It eliminates the possibility of a person or entity gaining greater control over the data by distributing it in a more equitable manner. Deduplication is also one of its key benefits, it refers to the ability of removing added data already existing in the system, enhancing its scalability and efficiency.

IPFS still faces an issue with content that is not pinned. Pinning is a process that prevents items from being removed during garbage collection as part of the caching mechanism. The next class in the classification scheme consists of NFTs whose data is permanently stored through the "calldata" of a transaction. Calldata is where the information from an external call to the contract is stored [38], which solves the pinning problem of IPFS, as the data becomes permanently available due to the nature of blockchain technology. However, this solution limits the token's functionalities since the data is only accessible from an external call, such as using a full node or delegating the call to a blockchain explorer like Etherscan. The data cannot be used by other functions contained in the same contract. One example of such a collection is 0xmons [39], which offers tools to store the acquired token in the calldata of a transaction, whose hash will be later retrieved by the smart contract code as the location of the metadata. The images in 0xmons are GIF files encoded in base64. This can be a cost-effective option since the cost of storing information in the calldata of a transaction is 39 times lower than that of storing it in the smart contract itself (16 gas per byte compared to 20,000 gas 32 bytes, respectively).

The third class in the classification scheme refers to assets whose data is stored in the contract, but requires a compiler to reconstruct the data from raw files. An example of this type of storage solution is 0xDEAFBEEF's Synth Poems, a collection of deterministic generative art. This means that the art is generated autonomously by a piece of code run in a computer, and the output will always be the same given the same input parameters. To enable users to retrieve the media file corresponding to a particular token, the author added the function

`getTokenParams` (shown in Figure 2.5) to the smart contract. When provided

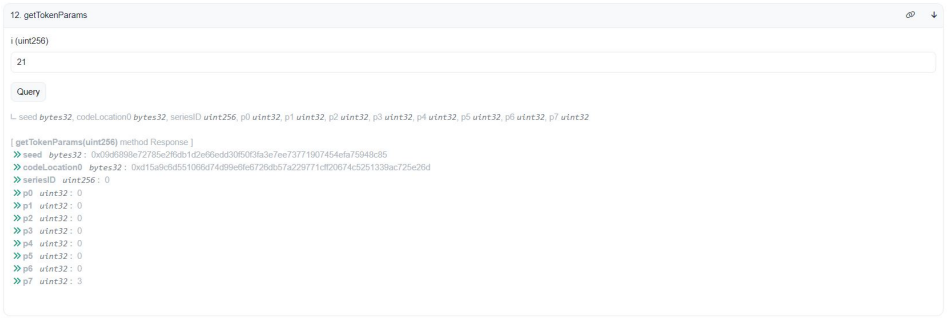

Figure 2.5: Metadata retrieved by Synth Poems' smart contract. Data Source: [41]

.

with a token ID, this function returns the hash of the transaction where the code written in C is stored, as well as a hexadecimal variable called the seed. The seed must be included in the raw code as the input to deterministically generate the corresponding media file, which consists of a one-minute audiovisual piece.

The most highly rated class of assets are those whose data is fully stored in the smart contract and can be reproduced within it without the need for any additional compilers. OnChainMonkey is an example of such a collection, as shown in Figure 2.6. When the `tokenURI` function is called, it returns a base64-encoded

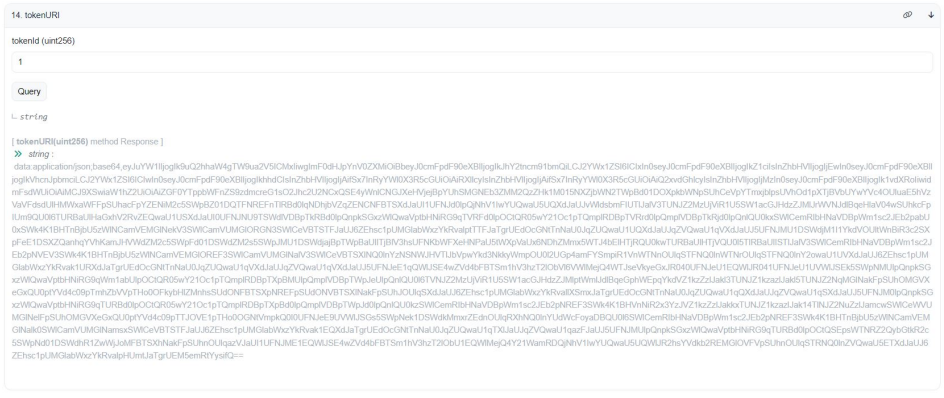

Figure 2.6: Metadata retrieved by OnChainMonkey's smart contract. Data source: [43]

.

string containing the token's metadata. Once decoded to UTF-8, the resulting JSON file contains the image description in SVG format, encoded once again in base64, this is usually done to handle special characters as discussed in the article in [42].

A quick examination of some of the top NFT collections traded in the largest

marketplace, OpenSea, provides meaningful insights about the storage solutions currently employed. Figure 2.7 illustrates the category under which 20 of the

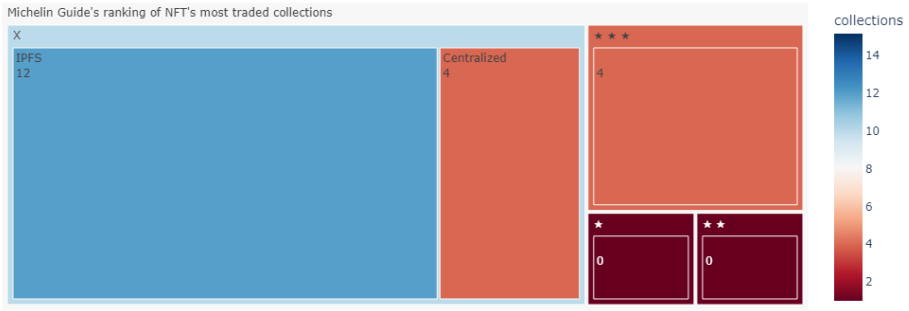

Figure 2.7: Tree map with the NFT metadata storage distribution for 20 of the most traded collections. Data source: [45]

biggest collections based on sales volume fall. For a list of the selected collections refer to [44].

IPFS is currently the preferred solution, and it appears that one- and two-starts collections are not among the most traded ones. The predominance of the 0-stars collections can be mainly attributed to the nature of the media files as discussed in the post in [46]. Fees in the Ethereum network are paid based on the amount of data being sent to the network, they are calculated as the product of the gas price at the time of transaction execution and the required gas (computational steps). 3-stars collections commonly use SVG files to store images within the contract. SVG stands for Scalable Vector Graphics and it is an XML-based format where the image is created using mathematical functions to represent the geometrical shapes that compose it. This format is easier to handle by the smart contract as the media file can be scaled to any arbitrary size, resulting in smaller files size. In contrast to vector-based SVG files, raster graphics are composed of a fixed grid of pixels, and the file size of an image is highly dependent on its resolution. This means that higher resolution images will require more pixels and therefore more storage capacity. Popular raster graphics formats include JPEG and PNG. However, they provide more granular control over colors, effects, and shapes than SVG. It is also worth noting that not all platforms support SVG natively, and additional software or plugins may be required to render these types of images. As a result, the decision to use different resolutions and graphics formats ultimately depends on the specific requirements of the project at hand. The development team must carefully consider factors such as the level of detail required for the media files, the computational resources needed to render them, the constraints of the EVM, and the limitations of the storage solution.

As an example, it can be considered Cryptoadz and Moonbirds, two popular collections published under the CC0 license [47]. As per the ERC-721 standard, the `tokenURI` function is required to return a URI that points to a JSON file

containing the token's description. Within this JSON file, there will be an object labeled as "image" or "image_data" that will provide the location of the media file. This feature allows the token to be utilized by various decentralized applications (dapps). For instance, NFT marketplaces like OpenSea have developed their own extension of the ERC-721 standard [48], enabling them to display the images in-app. However, in both the Cryptoadz and Moonbirds collections, a separate renderer smart contract was deployed to directly retrieve the media files. These smart contracts includes a function that, when called, returns a URI with the data scheme and a specific MIME type (Multipurpose Internet Mail Extensions) based on the media file's nature. In the case of Moonbirds, the URI format is "data:image/bmp;base64". The use of base64 encoding, as mentioned earlier, is particularly useful for handling special characters that could potentially affect the URI. Figure 2.8 showcases two tokens from each of these collections. The

Figure 2.8: CrypToadz #1044 (right) and Moonbird #1 (left). Images source: [49, 50]

Moonbirds collection uses BMP format to represent its images, while Cryptoadz uses GIF, both of which are types of raster graphics. By examining these images, it can be seen that all the necessary information is contained within a few pixels: 36×36 in the case of Cryptoadz and 42×42 in the case of Moonbirds. This makes it feasible to store the required layers to generate the images within the contract. These projects are also referred to as in-chain [51], as the images are rendered by the smart contract returning a bare-bones version of the image that does not require any additional computation to be displayed.

ecc0s, a 3-stars collection under public license, provides an example of how images can be generated in SVG format, as opposed to raster graphics. Figure 2.9 displays two of the collection's items and demonstrates how the resolution of the images is significantly higher compared to that of Moonbirds and Cryptoadz. The use of simple geometric shapes in ecc0s makes it easier to generate the images using a markup language. In NFT projects like this one, when the `tokenURI` function is called, it often returns an URI with the data scheme and the "application/json;base64" MIME type to encode and embed the JSON file with the token's description. Within this JSON file, one of the objects points to the media file, typically represented using the "image/svg+xml;base64" MIME type. Using

Figure 2.9: ecc0s #1 (left) and ecc0s #2 (right). Images source: [52]

this data structure enables the images to be displayed by web browsers and other software in a single HTTP request, rather than fetching it in multiple requests.

Hyperloot is another example of a collection published under the CC0 license whose metadata is stored off-chain due to the higher resolution of the images as it can be seen in Figure 2.10 where the original image is compared with one that has been resampled to 31×38 pixels using the approximation method in Photoshop. The representation is not 100% accurate as both images have had to be scaled to fit them within the page margins. Nevertheless, it provides a good visual representation of the idea to be conveyed. The increased level of detail of the

Figure 2.10: Hyperloot #1 original image (left) and resampled using a smaller number of pixels (right). Image source: [53]

images makes it highly expensive to generate them on-chain, as the minimum

number of pixels required to achieve that detail is considerably increased, in this case the images use 6,059,820 pixels, whereas in the Moonbirds and Cryptoadz collections they use about 1600. They could be generated using SVG at a lower cost, but generating intricate details or unique effects can be time-consuming and require extensive knowledge and experience.

While the technical aspects of NFT metadata are important for understanding how these digital assets are stored and traded, it is also essential to consider the legal and ethical implications of using and owning NFTs. One important issue to keep in mind is that purchasing an NFT does not always grant the buyer with Intellectual Property (IP) rights to the underlying digital asset. In some cases, the creators or owners of the digital asset retain ownership of the IP rights, even if the buyer holds the NFT as proof of ownership. In this case, the creators or owners of the images being used have waived their copyrights and related rights, allowing for their free and unrestricted use for informational purposes.

After the technical explanation about the underlying architecture of these assets, the chronology of events leading up to their widespread adoption will be continued. Following the enormous success of CryptoPunks, during the October 2017 ETHWaterloo, a hackathon bringing together many Ethereum experts from across the globe, a test version of the blockchain game CryptoKitties was developed, it was the first application to use the ERC-721 standard. The game consists of breeding cats whose appearance is determined by a number of attributes, the Cattributes, which can be inherit by the offspring. The cats are represented by ERC-721 tokens and can be obtained via breeding or acquiring them from sellers. The price of the NFTs is heavily influenced by their rarity, which in turn is determined by the perceived uniqueness and desirability of the item among users. This scarcity is a key driving factor in their value, as it is often based on the number of NFTs that share similar traits. NFTs with unique or uncommon traits are more likely to attract buyers' attention and command a higher price than those that do not.

The project was an enormous success, with the test version unveiled at the hackathon resulting in the sale of one of the earliest and most famous high-selling NFTs, Genesis, for a total of 246.9258 ETH ($113,082, considering the exchange rate at that time). The popularity of the game congested the network skyrocketing the gas fees, the monthly sales volume in December 2017 reached a total of 36,388 ETH, according to the information provided in [54]. This project set a significant precedent for NFT-based gaming, which is currently one of the most popular applications of NFTs. Its success was followed by the creation of new gaming and metaverse projects with Decentraland as one of the most prominent. Decentraland is a virtual world that uses both virtual reality and augmented reality technologies to create an immersive and interactive user experience. Decentraland operates on a unique governance structure that functions as a Decentralized Autonomous Organization (DAO). In this structure, decisions that will affect the virtual world are made through a process of decentralized decision-making and consensus-building among its members or token holders. This is achieved through the use of smart

contracts that encode the rules and decision-making processes of the DAO. The versatility of this governance model enables its application in diverse contexts, including the insurance industry. Later, it will be examined a prevalent instance of a DAO already functioning within this sector.

The platform has hosted big events, such as the recent Metaverse Fashion Week 2023 [55] with the presence of highly reputable firms, including Dolce & Gabbana, Tommy Hilfiger, and Adidas. The popularity of the project resulted in of the most expensive virtual lands sales ever made, when one of the subsidiaries of the company Tokens.com acquired the token Fashion Street Estate for a total amount of $2.4 million, transaction details can be found in the market tracker in [56].

Cryptokitties also served as the catalyst for the creation of OpenSea [57], which has now become the largest NFT marketplace. OpenSea provides a simple interface for users to trade a diverse range of digital assets in one place. To list an NFT for sale on OpenSea, users must first grant permission to the marketplace to manage their token through the `approve` or `setApprovalForAll` functions. The marketplace protocol, such as Seaport in the case of OpenSea, then takes over the management of the token and handles the listing process. OpenSea offers various mechanisms for selling items, including different timed auctions formats such as Dutch and English auctions, in addition to the traditional fixed-price listing. According to DappRadar [58], over 4 million traders have used OpenSea, making it the most widely used NFT marketplace. As of the time of writing, OpenSea accounts for 57.86% of traded volume in the top 25 NFT marketplaces, with a total volume of $35.48 billion. The emergence of these user-friendly NFT platforms have made NFTs more accessible to people with a less in-depth understanding of the underlying technology.

It was in 2021, when the NFT market experienced an unprecedented bull run, with skyrocketing demand and interest in NFTs. The involvement of major auction houses like Christie's and Sotheby's added a significant level of credibility to the burgeoning industry, attracting attention from a wide range of media outlets and stakeholders. As a result, the NFT market saw an influx of new investors, collectors, and creators, leading to a surge in sales and a deluge of innovative new projects. In March of the same year, the art world witnessed a groundbreaking moment in the history of NFTs with the sale of "Everydays: the First 5000 Days" by renowned artist Mike Winkelmann. The artwork was auctioned off at Christie's and fetched a record-breaking price of $69,346,250 [59], making it the most expensive NFT ever sold to a single collector. This unprecedented sale not only demonstrated the growing popularity and value of NFTs, but also marked a turning point in the traditional art market's acceptance of digital art as a legitimate and valuable form of artistic expression. While some critics speculate that this sale was a publicity stunt arranged between the collector and the artist to drive up the value of other tokens in the collection, the fact that such a high price was paid for an NFT is a clear indication of the growing interest and demand for these digital assets.

The surge in popularity and interest in NFTs continued throughout the end of 2021, driven by various factors. The lockdowns brought a new wave of individuals into the financial markets, with cryptocurrencies like Bitcoin reaching all-time highs. This, in turn, led to a significant increase in demand for NFTs, as illustrated

in Figure 2.11, which displays the sales and transaction volume history on the top two blockchains for NFTs. New blockchains started to appear in the scene and

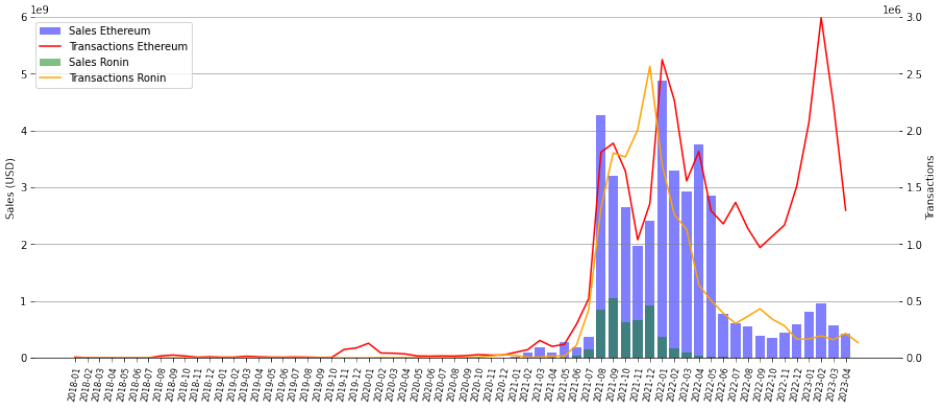

Figure 2.11: NFT sales and transactions volume history in the top two blockchains (excluded wash trades). Data source: [60]

.

many of the existing ones, started to implement their own NFTs such as Solana, Cardano and Flow among others.

Facebook's new strategic plan and rebrand to Meta also played an important role in the increasing demand of the NFTs during 2021. The company's vision of bringing the metaverse to the masses opens up new use cases for NFTs beyond the art industry. While some skeptics view the NFT market as a Ponzi scheme benefiting only early entrants, the concept of tracking ownership and provenance of digital assets has far-reaching applications in industries such as real estate and finance. Veracity Protocols is one such companies leveraging NFTs in conjunction with computer vision and machine learning algorithms to unlock the full potential of these assets. By creating a direct, immutable link between physical objects and their digital representation, they eliminate companies' dependence on unsecure links which can be removed, replaced or tampered with [61].

2.3 Vulnerabilities and Insurance Opportunities

The novelty of the technology underpinning digital assets has resulted in limited human understanding, leading to concerns about potential vulnerabilities that could be exploited by malicious actors. The intricate nature of the technology behind these assets has also deterred many from adopting them as readily as they would physical assets. Moreover, the escalating number and complexity of cyberattacks have hindered their widespread adoption. A study in [62] reported that out of the 1700 CISOs and IT professionals surveyed, 59% believed that cyberattacks are becoming increasingly sophisticated, and it is estimated that cybercrime will cost the world around $8 trillion.

The digital asset space has witnessed an increasing number of theft cases and

stolen value for both NFTs and cryptocurrencies in 2021 and 2022. Figure 2.12 illustrates the evolving landscape over two distinct but close periods of time. The left-hand graph demonstrates the total value stolen from the smart contracts of thirteen blockchains with high transaction traffic. The right-hand graph shows an increase in the number of NFTs stolen in 2022, despite a decrease in the average losses per item stolen.

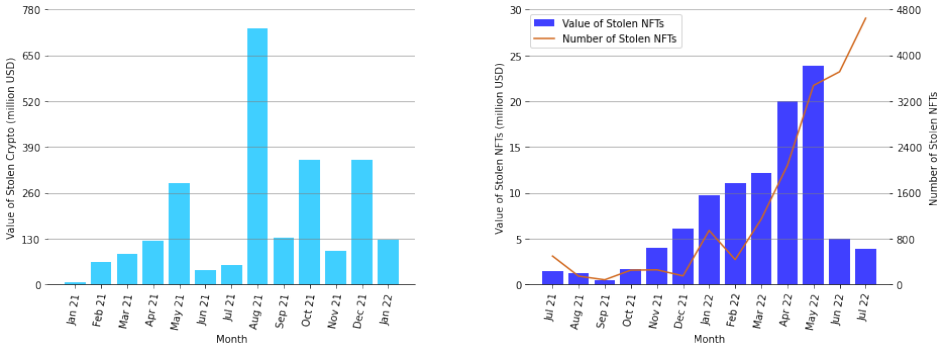

Figure 2.12: Cyrptocurrency lost to theft based on smart contract incidents on 13 different blockchains (left) and total value and number of stolen NFTs (right). Data source: [63, 64]

The lack of awareness among users regarding the potential risks associated with acquiring digital assets underscores the importance of having insurance coverage to safeguard their investments. Such coverage can bolster the reputation and credibility of these assets, paving the way for their expansion into other industries and driving the development of new applications to improve current business processes.

As reported in [65], currently only 1% of all crypto investments are covered under an insurance policy, which highlights the pressing need for insurers to start developing new policies in a market that is expected to grow at a CAGR of 11.1% and reach a value of $1.9 trillion by 2028 [66]. It is crucial for insurers to act quickly to provide insurance solutions that can mitigate the risks associated with digital assets and build confidence among users, thus promoting the long-term sustainability and growth of the cryptocurrency market.

The emergence of the Web3 economy presents a wealth of promising opportunities for the insurance industry, particularly with regard to the expanding range of insurable digital assets. This study delves into the rapidly expanding NFT market, with approximately $56.7 billion in total traded volume at the time of writing, based on the information provided in [60], making it a highly attractive market for early adopters who can develop scalable solutions and leverage their experience to redefine policy frameworks based on the increasingly available data and previous claims. As this market continues to evolve, those who are well-positioned to capitalize on these developments stand to reap significant profits.

Although the NFT market presents a compelling opportunity for insurers in the Web3 economy, it is not the only one. As previously mentioned, insurers can also leverage blockchain-based protocols to transform their traditional value chain,

reducing inefficiencies and unnecessary work. This presents a fascinating opportunity to streamline their operations and create more value for their customers in the Web3 ecosystem. The report in [67] provides some of the Web3 capabilities insurers could leverage to improve existing insurance products:

- *Smart contracts*: Insurers could embed policy agreements into code that is automatically executed when certain conditions are met. This provides a high degree of transparency as the code of the contract can be publicly accessible (depending on the type of blockchain as it was previously explained in the difference between private, consortium and public blockchains). By utilizing smart contracts, customers would have a better understanding on what type of coverage they are acquiring, avoiding the problem of vague descriptions buried under legal terms that are often difficult to understand. For instance, Coinbase's description of their coverage in [70] is quite brief and may not provide enough detail for customers seeking specific information. However, this also requires insurers to ensure that the code is understandable to customers who may not have extensive experience in the Web3 ecosystem.

- *Oracles*: Entities used to bring real-world data into the blockchain where the smart contract with the policy agreement is deployed. To avoid the problem of relying on a central authority providing the data, Decentralized Oracle Networks such as Chainlink Price Feeds [69] have been created. Oracles enable the creation of blockchain-based parametric insurance products which is a type of insurance where claims payouts are executed when pre-specified events are triggered such as natural disasters, whether events or market fluctuations.

- *Governance and utility tokens*: Tokens can be issued by companies to incentivize user contributions to the capital pool. These tokens allow stakeholders to participate in decision-making processes such as funds allocation, protocol upgrades, and investment decisions. Additionally, companies can use utility tokens to engage users in specific tasks such as claim evaluation or risk assessment, creating a sense of community. Governance tokens grant stakeholders the ability to vote on important decisions, ensuring that their voices are heard and valued. This level of participation and transparency can foster a sense of ownership and loyalty among users. Meanwhile, utility tokens can be used to incentivize users to perform specific actions, rewarding them for their contributions and encouraging continued engagement with the project. By leveraging a combination of both, companies can build a strong community around their project.

Nexus Mutual is one such example of insurance companies operating as a DAO, concept previously discussed when examining the governance structure of Decentraland, as part of the broader history of NFTs. As a mutual insurance company, Nexus Mutual is owned by its policyholders, rather than shareholders as is the case with traditional stock insurance companies [71]. This unique ownership structure makes it well-suited to operate as a DAO, given the decentralized and democratic nature of the organization.

According to information available on their website [72], Nexus Mutual offers its own NXM token, which provides users with various benefits such as on-chain governance, DAO governance, claim assessment, and staking. This token is backed by the capital pool created from all the ETH and DAI (two types of cryptocurrencies) invested by members. To ensure a reliable feed for the ETH/DAI price, which is necessary to maintain the minimum capital requirement for the platform's operations, Nexus Mutual utilizes Chainlink's Price Reference Contracts, a decentralized network of price oracles [73].

The backbone of the platform are its smart contracts, which require comprehensive security audits to ensure their reliability. The corresponding addresses of the smart contracts deployed on the Ethereum Mainnet can be found in [74]. Nexus Mutual policies are represented as NFTs, which contain the agreement details in their metadata. When a customer purchases a policy, a new NFT is minted and sent to the insured address. They currently offer coverage for a range of assets, including protocols deployed on EVM-compatible networks, validator node's stake, assets held in centralized crypto custodians, assets deposited into a vault strategy, and protection for cover providers.

Nexus Mutual's protocol provides a great example of how insurance companies can leverage the previously mentioned Web3 capabilities. The platform has already provided coverage for assets worth over $4 billion and has paid out more than $17 million in claims, as reported on their website [75]. These numbers demonstrate the potential profitability of the Web3 ecosystem and highlight the opportunities that will continue to emerge as human understanding of the technology evolves.

Chapter 3

Digital Assets Analysis and Elicitation

This chapter offers a thorough examination of the vulnerabilities that attackers commonly exploit in the realm of digital assets. Real-world examples are provided to enhance comprehension and establish a foundation for evaluating individual risk from an insurance standpoint. To facilitate the elicitation process, a Python script leveraging the identified attack vectors is developed. This script detects fraudulent transactions, capture their transaction hash, and store pertinent information. Furthermore, this chapter explores the limitations of such solution and presents an alternative approach.

NFTs can be classified according to their potential uses. Several websites provide different categorizations based on their own criteria. In this instance, OpenSea's classification will be followed, providing a brief explanation of each category:

Art. As discussed in their history, NFTs in the art category have become increasingly popular due to the need for a secure method of recording provenance and ownership of digital art. These NFTs have similar use cases to traditional art, such as collecting, exhibiting, and selling. In addition, a subcategory of art NFTs known as generative art has emerged, which involves art that is algorithmically generated by an autonomous system, it was already mentioned when giving the example of Synth Poems as a two-stars collection. Generative art NFTs have gained popularity due to their unique and unpredictable nature, with each piece being one-of-a-kind.

Gaming. Play-to-earn (P2E) games use a unique class of assets that offer players the opportunity to earn rewards as they progress through the game. These blockchain-based games have revolutionized the gaming industry by allowing players to monetize their in-game achievements. Cryptokitties is widely recognized as the first P2E game, while other popular titles like Axie Infinity and Gods Unchained have also gained immense popularity in recent times. With P2E games, players can earn valuable assets that can be sold for profit, adding a new dimension to the gaming experience.

PFPs. When most people think about NFTs, the first thing that comes to mind is Profile Picture NFTs. These digital assets are often used as avatars on social media platforms, especially among Twitter Blue subscribers. The popularity of

Profile Picture NFTs is on the rise, with famous collections such as the Bored Ape Yacht Club, CryptoPunks, and Doodles gaining massive traction in recent times. Beyond their aesthetic appeal, PFPs hold significant value as unique, one-of-a-kind assets that reflect the personality and tastes of their owners.

Photography. With the advent of NFTs, photographers now have a broader market to sell their artwork to, thanks to the exposure they get through various NFT marketplaces. Although photography may not be the most prominent category in the NFT space, it has been attracting new users and garnering significant attention. Collections such as Where My Vans Go have achieved remarkable sales volumes, exceeding 4000 ETH.

Domain names. They serve a similar purpose to traditional Domain Name Services, providing human-readable addresses that are easier to remember than long hexadecimal strings, making it simpler for users to verify that money is being sent to the right address. Ethereum Name Service domains are the most popular in this category, with names like paradigm.eth selling for over $700,000.

Music. NFTs are transforming the music industry by offering tokenized versions of artists' songs. Unlike the traditional purchasing model where the buyer pays for a license to listen to the song, NFT buyers purchase ownership rights of one of the minted tokens. This model creates a more equitable relationship between artists, labels, and streaming platforms, which in the traditional Web2 model, take a significant cut of artists' profits and creativity freedom. Moreover, fans play a more participative role, receiving royalties on streaming rights in some cases or even exclusive access to concerts or merchandise. By giving fans a direct stake in the success of the artist, NFTs have opened up new avenues for creative expression and monetization, offering a more democratic and transparent model for the music industry.

Sport collectibles. Tokens capturing memorable moments in the history of sports or featuring well-known celebrities. It represents a shift from the traditional sports card market, which has experienced a boom in recent years. By incorporating Web3 capabilities, these digital cards enable buyers to track the full history of each and ensure its authenticity. Examples of these can be found on NBA Top Shot, a marketplace featuring numerous tokens displaying basketball video clips.

Virtual worlds. Assets that represent ownership of lands, wearables, properties, and other items in alternate realities. As mentioned earlier, people can purchase virtual plots of land in Decentraland for vast amounts of money. This category of assets opens up new and exciting applications in the insurance industry, where analogies such as insuring a house or a vehicle could be adapted to this new alternate dimension. The introduction of NFTs in virtual worlds offers a unique opportunity to create new digital economies and redefine the way we people interact in virtual spaces.

It is crucial to distinguish between the different categories of NFT assets be-

cause they pose varying levels of risks. For instance, PFPs are among the most popular NFT collections and have been targeted in multiple theft cases in the past, making them a high-risk category. As a result, this study focuses on PFPs since they provide valuable insights into mitigating potential risks and offer a wealth of historical data about previous heists.

3.1 NFT Attack Vectors

NFTs can be compromised in various ways, with varying levels of sophistication. To gain a better understanding of the potential attacks, some of the most relevant from an insurance perspective will be mentioned following the extensive guide provided by Elliptic, a blockchain analytics firm, in [64].

3.1.1 Phishing Scams

Phishing scams encompass various types of attacks that share similar characteristics. Typically, these attacks involve malicious actors trying to deceive users into authorizing transactions that result in fund theft or compromising sensitive information through the replication of pop-ups resembling legitimate entities. In some cases, the attacker may try to set their own address as the operator in the `setApprovalForAll` function, giving them control over the user's tokens, this family of attacks was defined by Microsoft in early 2022 as "ice phishing" in their post in [76], and it is one of the most common vectors of attacks. In other cases, hackers simply infect with a malware the user's computer. Phishing attackers use a combination of engineering and psychological techniques to develop elaborate plans that can be difficult to detect. Some of the common forms these attacks can take will be discussed.

Domain Squatting and Impersonation

Cybercriminals often create counterfeit websites that mimic authentic ones, using search engine optimization techniques to boost their rankings. Figure 3.1 illustrates a recent search in Google for one of the most popular collections, showcasing an example of this practice.

Social Media Compromises

NFT collections teams and marketplaces often create their own social networks to communicate with customers and provide updates on the project's roadmap. Discord is one of the most commonly used social media platforms for this purpose. However, these communities are also vulnerable to malicious users who take advantage of the need for communication by posting fake links that can harm unsuspecting users.

Malicious users can employ various techniques, including social engineering, to manipulate a member of the project and gain access to sensitive information such as login credentials for their Discord account. Once a hacker gains access to a

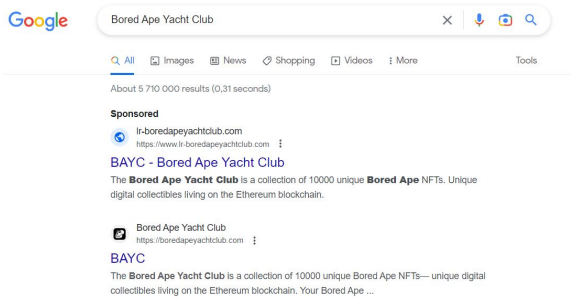

Figure 3.1: Fake website mimicking the original BAYC's website in Google Chrome

server, they can pose as the legitimate account owner and post malicious links to a wide audience.

To prevent such attacks, it is crucial for those managing the server to ensure that their security measures are effective and that there are no exploitable bugs. An example of an attack derived from a faulty tool occurred in the OpenSea Discord server. Collector Jeff Nicholas, told in the post in [77], how he brought a Zendesk ticket to the Discord channel to expedite a process (as many other reputable collectors used to do), but was then contacted via private message by a hacker impersonating an OpenSea help center staff member using permissions only granted to moderators. The hacker guided Nicholas to display a QR code of his MetaMask wallet, which was subsequently drained. This incident highlights the need for improved security measures, particularly when using external tools such as ticketing systems, and for increased awareness among community members regarding potential threats.

Another common failure among NFT collection developers is the display of broken links on their servers. This can provide an opportunity for hackers to reuse the link and create fake servers associated with it, which are then filled with malicious links.

Phishing Emails

Phishing emails are a common type of attack that can be similar to social media compromises. However, the scope of these attacks is often smaller as the targeted audience is typically segregated rather than concentrated on a single platform. To carry out a phishing attack, the hacker needs access to the victim's email address.

An example of a phishing attack occurred during OpenSea's migration to a new protocol that required users to migrate their listings within a short time frame to avoid paying gas fees [78]. This created an opportunity for hackers to exploit users' Fear Of Missing Out (FOMO) by sending them emails that appeared to be from the OpenSea team, providing instructions on how to migrate their assets. These emails included a malicious link that, when clicked, could lead to a variety of harmful consequences.

Airdrops Phishing Scams

The example of Cryptopunks airdrop was already shown in the text, it refers to the practice of distributing items from a collection for free to users as a way to promote the collection or reward loyal members. However, this practice has also been exploited by hackers who create fake websites with simple interfaces that trick users into claiming the airdrop. These fake websites can lead to malicious transactions and result in significant financial loss for the victim.

One way hackers promote these fake websites is by airdropping items into random users' wallets and including a reference to the website where they can mint the actual token or receive some sort of reward. Additionally, they may also impersonate legitimate airdrops by replicating their websites, a practice similar to domain squatting.

Trojan Horse NFTs

In some cases, unexpected airdropped items, instead of pointing to an external website, contain a malicious code that can execute harmful actions on the user's device or wallet. One example of this is the vulnerability discovered by the software company Check Point in OpenSea's platform, as detailed in their article in [79]. The vulnerability allowed for the embedding of malicious code into an SVG file that, when displayed in a web browser, would prompt victims to sign a malicious transaction under the OpenSea's operation domain.

These are the most common forms in which phishing scams can be seen, however as attackers become more sophisticated, it is likely that new and advanced techniques will emerge in the future, it is therefore necessary for insurers to educate users on how to avoid falling victim to these tricks by following strict safety checks.

In addition to phishing scams, there are other forms of attacks that can lead to the theft of NFTs. Although these attacks may not be as frequent as phishing scams, they still pose a significant threat to the security of the NFT, accounting for a high number of reported cases.

3.1.2 Swap Scams

In addition to NFT marketplaces, there are platforms that allow users to swap NFTs with each other. Hackers can exploit poorly designed user interfaces to pass off worthless NFTs as if they were from the original collection, tricking genuine users into trading with them. One example is the KiwiSwap platform, which used a flawed verification mechanism for official NFTs, displaying a green checkmark alongside the image to let users confirm they were receiving the original token. However, this allowed a malicious actor to create knock-off NFTs displaying the same checkmark and trade them with victims, as described in the post on [80].

3.1.3 Recovery Scams

Recovery scams are particularly insidious as they prey on the vulnerability of users who have already suffered a security breach. In these types of attacks, hackers create bot accounts on social media platforms like Twitter that automatically reply to users who have posted about losing access to their wallets, seed phrases or tokens. The bots offer to help the victim recover their funds and often include a link to a website that appears to provide recovery services. However, the website is actually a front for the hacker to steal the victim's funds.

3.1.4 NFT-based Protocols Exploits

These are probably the most sophisticated attacks, usually executed by experienced users with extensive knowledge of the space. Due to the complexity of these attacks, preventing them can be challenging since they can take various forms depending on the vulnerabilities they exploit. However, based on past experiences, it is possible to identify certain groups based on similar patterns.

Marketplaces Protocols Exploits

As seen in the previous numbers when discussing OpenSea's influence in the space, there are some NFT marketplaces who have become the go-to platforms for buying and selling NFTs, resulting in a certain level of centralization in the ecosystem. Collections restore to these entities to promote their collection in exchange for part of the earnings that are paid as a percentage of the sells. OpenSea currently charges users a 10% on the minting earnings [81] and 2.5% on secondary sells [82]. The life cycle of an NFT project can be separated in two parts: the minting process, where users create the NFTs (a record on the blockchain stating they own an item with a certain ID), and the secondary sales, where users trade it with other assets. The minting process is fairly standardized, and developers can create smart contract code for the primary sales event, thereby avoiding marketplace fees. However, some developers still choose to use these marketplaces, as they provide access to a wider audience and benefit from the marketplace's trust and reputation as a battle-tested solution.

During the minting phase, collections receive all the revenue from the sale, which is usually set at a lower price due to the random assignment of tokens. In some cases, the reveal of the media files is postponed so that users do not lose their interest in the minting phase as the rarest tokens are bought. In contrast, the secondary sales revenues are set as a percentage of the sales, the royalties, which, for example at OpenSea are capped at 10%. The small size of these roayalties incentivize users to trade their tokens, making it an appealing option for collection creators as it helps to attract a broader audience. The relatively modest profits from secondary sales make it appealing for collection creators to utilize these "centralized marketplaces". Even though they are mostly run by smart contract code, they have a certain level of control over the collections, such as the ability to pause the sale of an item, exclude an entire collection from their platform and in some platforms, they even control the assets being listed for sale as it is the case of the custodial option in marketplaces like Nifty Gateway or Binance.

Given the high volume of transactions managed by NFT marketplaces, these entities must deploy strong safety measures and security checks, as they are prime targets for attackers. Elliptic's report highlights two different marketplaces with faulty tools that allowed users to exploit them for profit, including OpenSea. In this particular attack on OpenSea, users were able to exploit a flaw in the platform's functionality. The blockchain operates as an append-only ledger, which means that when users wish to create new listings, they must first cancel any existing ones. However, OpenSea allowed users to modify the price of their listings without incurring gas fees. This resulted in the old listings remaining active on the blockchain, even though the platform displayed the new price on the front-end. Consequently, users were able to purchase items at the previous lower price, leading to situations where NFTs were sold at a staggering 99% below their floor price, as reported in [83].

Airdrop Exploits

While most secondary sales of NFTs take place on major marketplaces, some developers choose to reward their loyal members by promoting their collections and offering free incentives, implementing their own solutions to do so. One common method to achieve this is to allow current owners of items within the collection to claim rewards. However, developers must exercise caution when writing the conditions in the smart contract that users must meet to claim their rewards. This is to prevent malicious users from taking advantage of these airdrops.

The article in [86] describes how an attacker stole unclaimed items from five users of the Bored Ape Yacht Club (BAYC) collection. In March 2022, the BAYC team decided to airdrop 10,094 ApeCoin, their own cryptocurrency, to all BAYC holders. However, due to a flaw in the code that did not check how long users had owned the items, an attacker was able to use a flash loan, a type of uncollateralized loan that can be borrowed and repaid within a single transaction, to borrow five items that were deposited in a vault on NFTX. NFTX is a platform that creates liquidity for NFTs by allowing users to earn yield from protocol fees.

In NFTX, users deposit an NFT in the corresponding collection vault and receive a token in exchange, whose value is determined by the balance of ETH and NFTs in the liquidity pool. Each time a trade is made in the vault, users providing liquidity by staking NFTs (inventory providers) or ETH and NFTs (liquidity providers) are rewarded with a share of the fees that have been paid. To execute the flash loan, the attacker purchased an ape that was listed for sale, needed to pay the protocol's fees. Then, in the same transaction, borrowed five items whose reward had not been claimed by their owners and subsequently claimed the reward and returned the items to the vault, netting a profit of approximately $350k. The whole transaction details can be found in Etherscan in [87].

This incident highlights why users often prefer to use established NFT marketplaces such as OpenSea rather than white-labeled marketplaces. Even large teams of developers, such as those behind Yuga Labs' collections (which includes BAYC), may not have the same level of experience in creating robust solutions as marketplaces that have been battle-tested by managing a high volume of daily transactions.

Cross-chain Bridges Exploits

Bridges provide a mean for interconnecting different blockchain networks, but their maturity level is not yet sufficient, and they remain vulnerable to various types of attacks. Bridges have been a common target of many of the biggest cryptocurrency thefts over time. According to the article in [84], bridge attacks accounted for 70% of all cryptocurrency losses in the past year alone.

Network congestion on blockchain networks such as Ethereum has led to the development of new solutions, commonly referred to as domains, including layer 2 scaling solutions and new layer 1 blockchains. These domains offer faster transaction confirmation times and lower fees than their predecessors. To facilitate the transfer of assets between these new solutions and existing ones, bridges have been developed. However, the implementation of bridges presents a challenge known as the "interoperability trilemma", which is discussed in the article referenced in [85]. This trilemma refers to the trade-off between security, generalizability, and extensibility when implementing mechanisms for transferring assets between domains. Achieving high levels of security, generalizability, and extensibility simultaneously is difficult, and typically, one of these capabilities must be sacrificed to achieve the other two.

One commonly used mechanism for transferring assets between domains is the lock-mint/burn-release process. In this approach, the assets being transferred are locked in the source domain, and an equivalent amount is minted in the destination domain. To reverse the process, the destination's minted tokens are burned, and the locked assets can be redeemed. However, native crypto assets residing on one blockchain, such as Bitcoin, cannot be used on other chains, such as Ethereum. To solve this problem, a wrapped version of the token is created to meet the destination standards. Bridges typically rely on a relayer to handle communication between domains, and there are several implementation options. Trusted bridges, are the most common choice, they utilize a federation of off-chain relayers that validate and verify transactions. To achieve consensus on which transactions to include in the bridge, relayers may use a multisignature (multisig) mechanism that requires a certain number of signatures from a pre-selected group of validators. While this solution supports the exchange of arbitrary cross-domain data and is compatible with all domains, it is censorship-prone due to the interoperability trilemma. If the majority of nodes in the federation is compromised, funds can be stolen.

While most bridge attacks to date have targeted fungible tokens, new bridges are now facilitating the transfer of NFTs between domains. As a result, it is essential to explore what are the safest options.

3.2 Policy Rating Factors

The subsequent sections will concentrate on the statistical analysis that actuaries must undertake to develop a pricing model.

To establish the premium for a policyholder, a set of rating factors is utilized to categorize them according to their insurability risk. The rating factors are to be selected based on the study of the NFT attack vectors. The study reveals that

phishing attacks are currently the most prevalent threat, highlighting the significance of the wallet type chosen by users in minimizing the attacker's window of opportunities. As a result, a classification scheme has been developed to categorize insureds based on the wallets used. This classification primarily focuses on solutions within the Ethereum blockchain, considering that it is the primary platform for NFT trades, but it can also be extended to other domains.

Initially, a classification can be established based on the type of account where assets are deposited. While EOAs are the native accounts on the Ethereum blockchain, recent developments have emerged to adapt smart contracts' behavior to function similarly to wallets. These are commonly referred to as smart contract wallets, with Argent being an example. The logic associated with these wallets is more intricate since smart contracts cannot independently initiate transactions; they require triggering by an EOA.

In the case of Argent wallets, when a user creates an account, both a smart contract and an EOA are automatically deployed. The private key of the EOA, securely stored in the user's device, communicates off-chain with a relayer responsible for on-chain interactions with the smart contract containing the wallet logic [88]. These wallets enhance the capabilities of conventional EOAs by introducing new features, such as setting guardians. Guardians are a designated set of accounts with specific permissions over the smart wallet. They can perform actions like locking and unlocking the wallet or initiating a recovery procedure in case the user loses the device with the EOA associated with the smart contract.

If a user misplaces the device with the EOA registered as the owner of the smart wallet, they can request one of the designated guardians to lock the account, preventing unauthorized access to the funds. Account recovery is also possible, allowing the user to set a new device as the wallet owner, subject to approval from a specified number of guardians. Additionally, ownership of the wallet can be easily transferred without interruptions by obtaining signatures from the required number of guardians and the current owner of the account. Furthermore, certain functionalities like implementing a prolonged waiting time for spending a significant amount of assets can be incorporated into these wallets. They are sometimes referred to as "vaults" and are considered one of the safest mechanisms for long-term asset storage. However, when users wish to trade their NFTs, they must transfer them to an EOA since this is the wallet supported by major marketplaces.

Secondly, wallets can be classified based on how the keys are stored. This classification does not impact smart contract wallets since the private key can be replaced in case the user's device, where it is stored, becomes compromised. EOAs can be broadly categorized as hot and cold wallets (there is also an intermediary group called warm wallets, but it will not be discussed here). Hot wallets store the private key online, which is highly convenient for users requiring frequent daily transactions. They can simply access the wallet extension like MetaMask and authorize transactions within seconds. However, hot wallets are more vulnerable to theft as the window of potential vulnerabilities is wider. On the other hand, cold wallets store the private key offline on a separate device. Users must connect this device with the private key every time they want to perform a transaction. A commonly used solution for cold wallets is Ledger. Cold wallets that are specifically used to store NFTs, without interacting with any other party apart from the wallet

used to list the token for sale, are sometimes also referred to as vaults.

Lastly, wallets can be classified based on how they are custodied. There are custodial wallets, where users entrust the management of their keys to the entity providing the wallet service. This is commonly observed in centralized exchanges such as Coinbase or Binance. On the other hand, there are non-custodial wallets where users have complete responsibility for how their keys are stored. Custodial entities typically implement robust security measures, safeguarding private keys in physically secure locations. However, they remain enticing targets for attackers. Hackers may also attempt to circumvent the multi-factor authentication process required for users to access their funds and transfer them to their own wallets.

Custodial wallets are a popular choice for many users due to their simplicity and user-friendly experience, particularly for those with less technical expertise. These wallets offer a streamlined onboarding process, making it easier for users to get started and manage their funds.

Smart contract wallets can be implemented in various ways, each with different security features. These features may include the number of guardians, waiting time for transaction execution, withdrawal limits, and more. Due to these variations, it is challenging to classify them into a single group. However, considering their high level of security and the limited available data on historical thefts associated with smart contract wallets, categorizing them as a type of cold wallet is a prudent assumption to mitigate unexpected losses. Similarly, NFTs stored in custodial wallets, such as those obtained from the Binance NFT marketplace, can also be categorized as cold wallets due to the lack of data and the safety measures implemented in these solutions.

This classification simplifies the rating factors to how keys are stored, allowing policyholders to be categorized into hot and cold wallets. Given the relative newness of the space, insurers may initially develop slightly overpriced models as a precautionary measure. These models can be further refined to create more accurate and competitive policies as more data becomes available.

3.3 Data Collection

In order to tackle the challenge presented by the absence of publicly accessible registries containing data on NFT theft cases, two distinct approaches have been explored in an effort to find a viable solution. Both approaches focus on gathering data specifically from the Ethereum network, which has experienced a significant number of cyberattacks in recent years.

3.3.1 Process Automation Based on Patterns Identification

The initial strategy involves utilizing pattern identification from prevalent attacks to create a Python script that stores the hash of fraudulent transactions in a dictionary, along with the potential type of attack that compromises the wallet.

There are two approaches to implementing this automated process. The first and most versatile method involves running a full node and locally storing a copy of the blockchain. This allows for quick retrieval of all the necessary information. The second option is to utilize public APIs provided by blockchain explorers like

Etherscan. This option is more feasible as it doesn't require storing the entire blockchain, which can be cumbersome due to its large size. According to [89], the blockchain's size is 972.55 GB at the time of writing for a full node. To overcome the storage limitations, a free API key was requested from Etherscan, and a Python script was implemented to fetch data from the available API endpoints within the free plan. Some information was directly extracted from the retrieved JSON files, while others required parsing using BeautifulSoup objects before being stored in the respective dictionary objects. The script can be found in [93], which primarily aims to identify two types of thefts: phishing and compromised private keys. The patterns shown in Figure 3.2 were identified for most of these thefts.

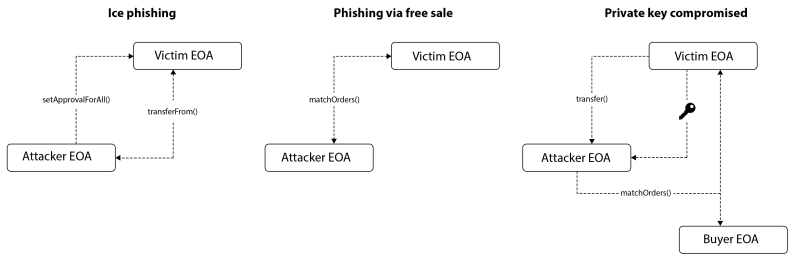

Figure 3.2: Common strategies employed to steal NFTs

As previously explained, ice phishing refers to cases where a user is deceived into signing a malicious transaction that designates the hacker's address as the operator of tokens in a collection using the `setApprovalForAll` function. Subsequently, the hacker transfers the tokens to his own address. In many of these cases, a recurring pattern emerges: the transaction is initiated by the same address that receives the tokens from the victim's wallet. An example of such transactions is depicted in Figure 3.3.

Figure 3.3: Example of Ice Phishing. Data source: [94]

Another method of deceiving users involves tricking them into signing a transaction where the offer side is empty, and the victim's tokens are listed in the consideration side. Figure 3.4 showcases an example of a highly sophisticated theft, wherein a user was lured into listing his 14 BAYC items for a mere 0.00000001 ETH. Further details about this scam can be found in [96].

When a private key is compromised, a common behavior observed is the immediate sale of the token to one of the existing bids, followed by the transfer of the proceeds from the sale to an external account controlled by the attacker.

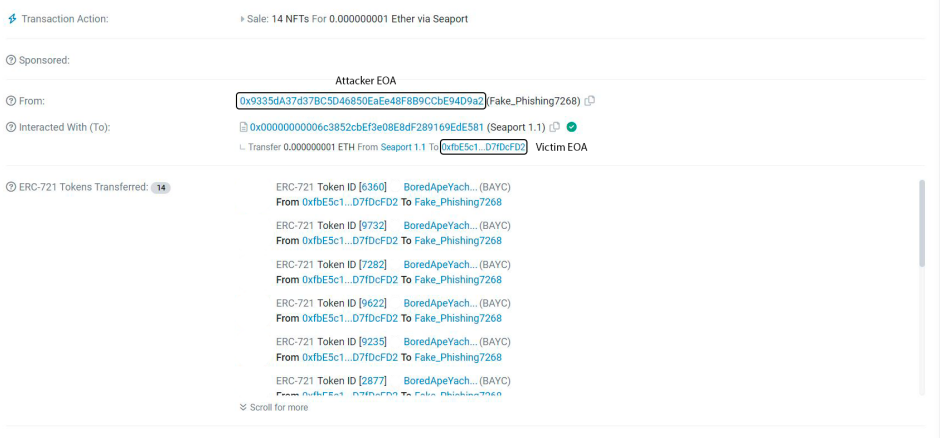

Figure 3.4: Example of Phishing via free sale. Data source: [95]

Figure 3.5 illustrates an example of such transactions.

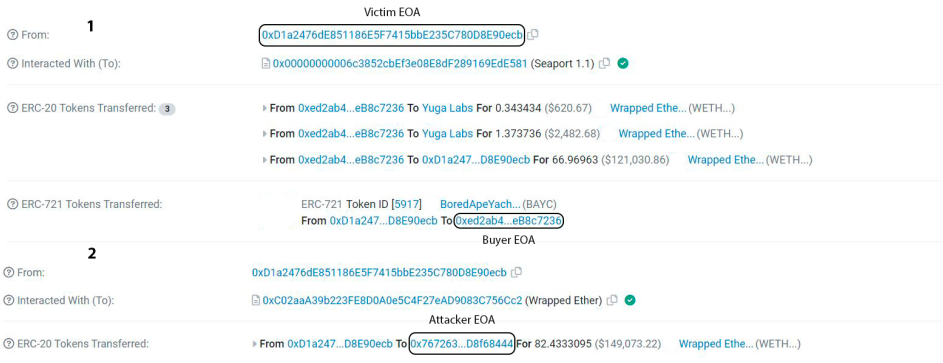

Figure 3.5: Transactions executed by an attacker who gained access to the victim's private key. Data source: [97, 98]

The intention behind using this approach was to identify transactions that follow the specified patterns, manually examine them, and iteratively refine the code until achieving an automated mechanism with a desirable level of accuracy. However, only a trial version was developed due to the API's call rate limitations, making it challenging to fetch a high volume of transactions within a reasonable timeframe. As stated on their website [99], there is a specified limit of 5 calls per second in the free plan. However, during code execution, the observed limit was 100 calls per minute, as depicted in Figure 3.6 that illustrates the calls made per minute after an hour of running the program. This discrepancy could be attributed to other operations performed by the script, network latency, server load, among other factors. Since the time required to discover a significant number of cases became impractical, the decision was made to collect the data manually.

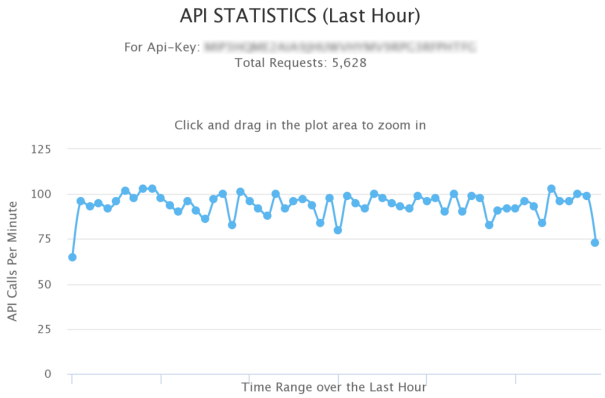

Figure 3.6: API calls per minute after one hour running the Python script

3.3.2 Manual Collection

Manually collecting data is a laborious process and not the ideal choice for applications like insurance pricing, where the accuracy of the model heavily relies on historical observations. However, due to the aforementioned practical limitations, this method is employed in this policy framework. The advantage of manual data collection is that, when utilizing reliable sources, it allows for the inclusion of a wide range of attacks in the database with a high level of confidence in the accuracy of the information being added. Automating the process would require numerous iterations to attain the same level of accuracy for all recorded cases.

The information can be sourced from various websites and platforms. Numerous newspapers in the crypto space provide updates on relevant events, including major thefts targeting renowned collectors. However, these articles often lack technical explanations as they aim to reach a broader audience.

During the exploration of different articles reporting thefts found on the internet, it was discovered that the most detailed and accurate information can be obtained from crypto sleuths on Twitter. Among them, ZachXBT [90] and PeckShieldAlert [91] have proven to be valuable sources.

Additionally, the project "Web3 is Going Just Great" [92], developed by Molly White, has been found to be highly useful in aggregating many reported thefts in one place. The platform offers a simple and easily readable interface, where each case is structured as a chronological thread. These threads often include references to Twitter posts where authors like ZachXBT and PeckShield provide technical explanations and relevant details. By leveraging these sources, a comprehensive and up-to-date understanding of reported thefts can be obtained.

All registered cases involve thefts associated with ten of the most popular collections, with nine of them being PFPs and one falling under the virtual worlds category. For each case, a new entry has been added to a Python dictionary corresponding to the respective collection. These entries include essential information such as the date of the theft, token minting and purchase, floor price of the collec-

tion at the time of the theft, type of attack, and the token's ID. Subsequently, all the dictionaries were sent to a local MySQL database using the Python DB API. The entire process can be found in the previously mentioned Jupyter notebook [93].

A total of 124 cases have been registered. Due to the lack of information regarding the type of compromised wallet in many instances, the approach taken is to record the specific attack that compromised the wallet and classify them based on whether the vulnerability can be exploited in a hot wallet, cold wallet, or both. The data table can be found in [100] as a CSV file. It is made up of all the tokens belonging to the selected collections, those with a '1' in the stolen column represent the registered thefts. In addition to the columns mentioned earlier, there is an additional column called 'rarity_score' that plays a crucial role in estimating the price of the token, as further elaborated in the following section. Figure 3.7 depicts the dimensional data model diagram for the SQL tables used.

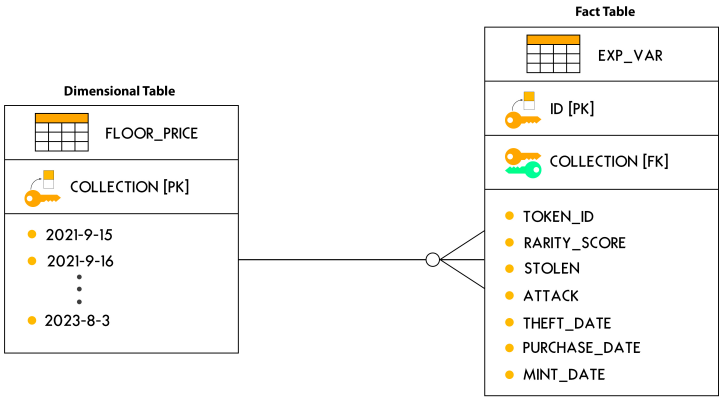

Figure 3.7: SQL dimensional data model diagram

Chapter 4
Statistical Modeling

The objective of the statistical analysis is to calculate an estimate of the anticipated losses associated with an insured item over a specific time period, commonly known as the "risk premium". The insurer charges the policyholder based on this estimate to cover the risk for the upcoming days. In this context, a one-step prediction approach is adopted to forecast the risk premium, assuming a daily exchange of money between the insurer and the policyholder. While this approach can be extended to longer time frames, the accuracy of the model diminishes for longer-term predictions.

The risk premium is calculated using the following formula:

$$\text{risk premium}_{t+1} = f(x = \text{days elapsed}) \times \text{item price}_{t+1} \tag{4.1}$$

In this equation, t represents the current point in time when the prediction is being made. The first term on the right side of the equation represents the probability of the item being lost after a certain number of days have passed since its purchase by the policyholder. The second term is the one-step prediction of the price.

Hence, the analysis can be divided into two distinct parts. Firstly, a probability distribution will be fitted to the observations in the dataset to model the likelihood of losing an item. Subsequently, another model will be developed to forecast the price of the token using time series analysis techniques.

4.1 Probability of Item Loss

To determine the probability of item loss, the activity of the token has been tracked from the date it was stolen until the purchase date by the owner whose wallet was compromised. The number of days elapsed has been recorded for the theft cases where it was possible to identify the purchase date. However, in some instances, identifying the purchase date was challenging due to the token being transferred across multiple wallets. Therefore, it cannot be conclusively confirmed that the last buyer of the token is the same person whose wallet was compromised.

To analyze the duration the token remained in the user's wallet, a histogram has been generated to visualize the distribution of the elapsed days. The Python package `Fitter` has been utilized to identify the optimal distribution that best fits the data. This package explores a total of 123 distributions available in the `scipy`

package and ranks them based on various criteria, including mean square error, Akaike Information Criterion, and Bayesian Information Criterion.

As mentioned earlier, the policyholder's charges are determined based on the rating factor, which in this case is the type of wallet. Since specific information about the compromised wallets was not available, the dataset has been divided into two, based on the type of attacks that compromised the wallets.

The first dataset is made up of all the attacks that could compromise a hot wallet. In this case, all the attacks found are capable of compromising a hot wallet, which is reasonable considering their vulnerability. However, it is important to note that not all types of attacks, such as physical manipulation of storage devices for cold wallets, would necessarily affect hot wallets. The second dataset, which can be considered a subset of the hot wallets, only includes attacks that could steal funds from a cold wallet.

The histogram with the number of days elapsed until each token was stolen is shown in Figure 4.1.

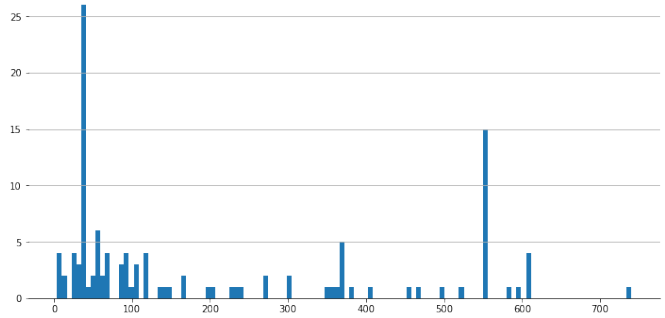

Figure 4.1: Histogram with the number of days elapsed until a token was stolen. Data source: [100]

To prevent overfitting issues when adjusting the distributions, a jitter parameter has been introduced for bins with more than 10 observations. The jitter values were randomly generated from a uniform distribution ranging from -10 to 10 days.

Figure 4.2 illustrates the shape of the histogram for the hot wallets, along with the top 5 ranked distributions based on the 'sumsquare_error' value. The accompanying table provides details on the criteria utilized for ranking these distributions. Following a thorough analysis, the alpha distribution has been chosen as the most suitable option. According to [101], this distribution is commonly employed in wear tool problems and lifetimes modeling, making it a promising fit for this particular application. Referring to [102], the probability density function (PDF) of the alpha distribution can be expressed as follows:

$$p(x;\alpha) = \frac{1}{x^2 + \Phi(\alpha)\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(a-\frac{1}{x})^2} \tag{4.2}$$

The aforementioned process was also conducted for thefts where the attack could potentially compromise a cold wallet. This particular dataset consists of 97 observations in total. Figure 4.3 presents the fit of the top 5 distributions for

	sumsquare_error	aic	bic	kl_div	ks_statistic	ks_pvalue
nct	0.000420	1527.538344	-1474.917981	inf	0.136014	0.022128
alpha	0.000436	1504.720238	-1475.300197	inf	0.124994	0.044381
norminvgauss	0.000438	1482.848301	-1469.922458	inf	0.142010	0.014781
invweibull	0.000442	1498.740019	-1473.624631	inf	0.130523	0.031534
genextreme	0.000442	1498.740011	-1473.624616	inf	0.130523	0.031534

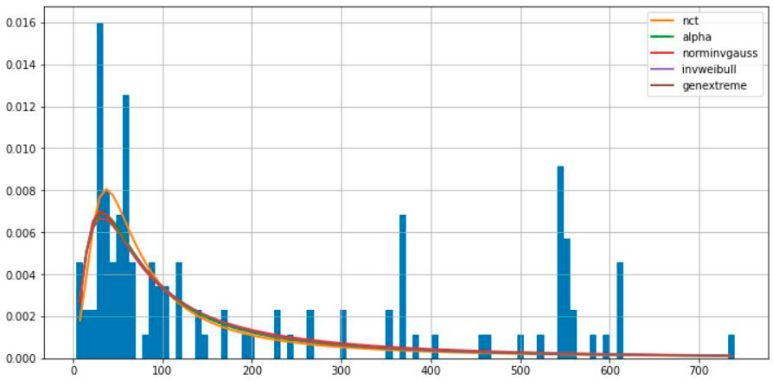

Figure 4.2: Shape of the histogram with the top 5 ranked distributions for attacks compromising hot wallets, along with the corresponding criteria values. Data source: [100]

this dataset. Notably, the alpha distribution emerged again as one of the best fits, leading to the decision to utilize it for this group as well.

The optimized parameters and fitted distributions for each group of policyholders are summarized in Table 4.1.

Group	Distribution	Shape	Location	Scale
Hot wallet	Alpha	$\alpha = 0.0003$	-27.2	84.6
Cold wallet	Alpha	$\alpha = 0.002$	-21.9	82.5

Table 4.1: Optimal parameters for the distributions

The observed results reveal a slightly higher area under the density function for hot wallets during the initial month compared to cold wallets. Specifically, the cumulative density functions (CDFs) exhibit values of 0.14 for hot wallets and 0.11 for cold wallets within the first 30 days, as demonstrated in Figure 4.4 for visual comparison. This disparity was anticipated since the probability of a hot wallet being compromised is generally higher, leading to victims being hacked or scammed within shorter time frames. However, it is important to note that the difference is not significant, and it could be attributed to the small size of the dataset being used. A more comprehensive data collection process would likely provide a clearer distinction between the two wallet types.

Consequently, policyholders in the hot wallets group would pay higher premiums during the initial months. However, over time, the area under the PDFs will

	sumsquare_error	aic	bic	kl_div	ks_statistic	ks_pvalue
nct	0.000454	1525.189991	-1106.170091	inf	0.149848	0.028732
alpha	0.000481	1505.293914	-1105.275784	inf	0.153552	0.023269
norminvgauss	0.000482	1475.968223	-1100.529254	inf	0.159536	0.016368
genextreme	0.000484	1496.639041	-1104.799344	inf	0.146505	0.034605
invweibull	0.000484	1496.639063	-1104.799328	inf	0.146505	0.034604

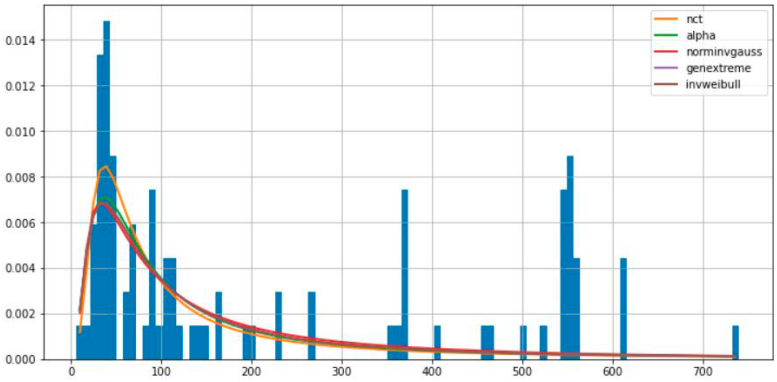

Figure 4.3: Shape of the histogram with the top 5 ranked distributions for attacks compromising cold wallets, along with the corresponding criteria values. Data source: [100]

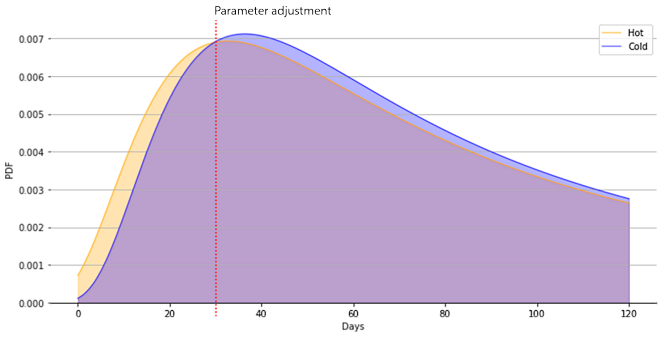

Figure 4.4: Comparison of the CDFs of both groups for the first 120 days

become more similar for both groups, eventually approaching a value of 1 for an infinite number of days. This suggests that periodic recalibrations in the parameters can be implemented by insurers to prevent overcharging cold wallet policyholders. By continually evaluating and updating the parameters, insurers can ensure a fair and balanced premium structure for policyholders in different wallet groups.

To calculate the first term in the right side in Equation 4.1, it is still necessary to determine the probability of an item from a collection being stolen. For this estimation, the number of BAYC items marked as suspicious on OpenSea is chosen. Various reasons, apart from being stolen, may result in items being flagged as suspicious, such as involvement in wash trading activity or their acquisition through stolen funds. Additionally, only those reported will be flagged, and it is likely that not all stolen NFTs are reported. Therefore, the compensation for the unreported ones is considered to be provided by items marked as suspicious due to other non-theft-related activities. Given that BAYC is one of the most popular collections and thus a prime target for cyberattackers, the chosen value will be somewhat higher compared to other collections, making it a safe assumption.

A helpful resource for obtaining this information is the crypto data aggregator @beetle, which has created the dashboard referenced in [103]. As of the time of writing, the dashboard indicates a total of 153 NFTs marked as suspicious out of a collection size of 10,000 items. Therefore, it will be assumed that 153 out of every 10,000 items are lost due to theft.

The probability of loosing an item at a certain point in time can then be calculated as the product of the value obtained from the corresponding PDF and the probability of it being stolen, which is $153/10,000$ in this case. This conservative assumption aligns with the concept of initially developing slightly overpriced policies before gathering more information and offering more competitive prices.

4.2 Price Modeling

The second part of the analysis deals with the token price forecast. The value of a NFT is influenced by various factors, including provenance, utility, scarcity, and users' self-identification, among others. While quantifying most of these factors can be challenging, it is still feasible to construct a simple model using measurable elements.

The model proposed here for pricing NFTs is based on two primary factors: the collection's floor price and token rarity. The floor price represents the minimum amount required to acquire an item from a given collection, while token rarity measures the uniqueness of a particular token. Token rarity has played a significant role in the history of NFTs, as demonstrated by the impact it had on Cryptokitties, where users engaged in breeding activities to obtain new and distinct collectible cats, highlighting the value placed on uniqueness within the NFT ecosystem.

In each collection, every token possesses unique traits described in the JSON file obtained through the `tokenURI` function. Several websites offer APIs that provide information on the number of tokens within a collection that share a specific trait. By utilizing this data, it becomes possible to assign a score to each

token based on its similarities with others in the collection. Tokens that exhibit very few shared traits with others are considered more unique, leading to a higher perceived value among collectors. Consequently, it is reasonable to augment their price by a certain percentage above the floor price, based on their level of rarity.

To model and forecast the floor price, time series analysis can be applied to the gathered data from the website referenced in [105]. The data has been stored in a separate dimensional table, following the structure illustrated in Figure 3.7. In this table, each column represents a specific date, and each row contains the corresponding floor price for a particular collection. The data collection process was completed on March 8, 2023, resulting in a historical dataset covering the period from September 15, 2021, to March 8, 2023, with daily frequency.

The process for modeling the floor price of the BAYC collection, described in this section, can be applied to other collections as well. Implementation details and the Matlab script can be found in the provided reference [104].

The first step is to split the data into three groups: the training dataset, which comprises 70% of the observations, the validation dataset with 10%, and the test dataset. This division allows for proper evaluation and validation of the model.

Next, it is essential to visually inspect the raw data for any outliers or irregularities. Plotting the raw data helps in identifying any unusual patterns or extreme values. The corresponding plot can be found on the left side of Figure 4.5. To address sudden peaks in the data, a median filter was employed. This filter

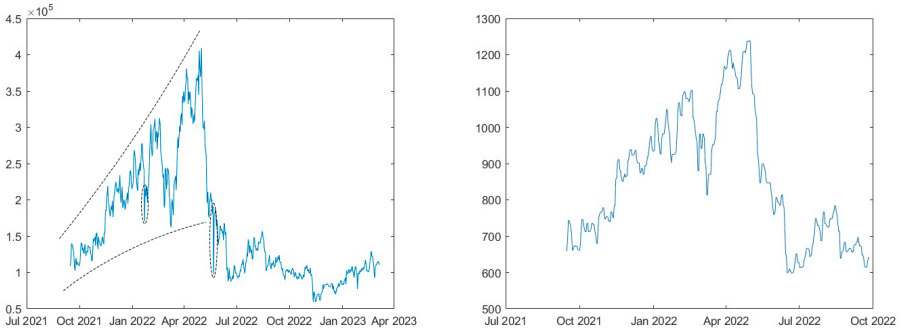

Figure 4.5: Line chart of the raw data (left) and the cleaned training dataset (right)

smooths the signal by taking the median of three adjacent measures at each data point. Additionally, it can be observed that the variance is not stable, with a significant increase occurring in mid-2022. To stabilize the variance, a Box-Cox transformation was applied.

Upon examining the normality plot, a value of -0.5 for λ was suggested. However, implementing this value compressed the range of the values, making it challenging to develop an accurate model capable of predicting small movements in the signal. Consequently, a value of 0.5 for λ was chosen.

The right side of Figure 4.5 depicts the signal after applying the median filter and the power transformation to the training dataset. These preprocessing steps

result in a smoother and more stationary signal, facilitating the development of accurate models for floor price forecasting.

Following the data transformation, the next step involved identifying trends and periodicities. To assess any potential seasonal components, the periodogram of the transformed data with the mean subtracted was examined. This is depicted in the left side of Figure 4.6. To analyze the periodogram, a Hanning window

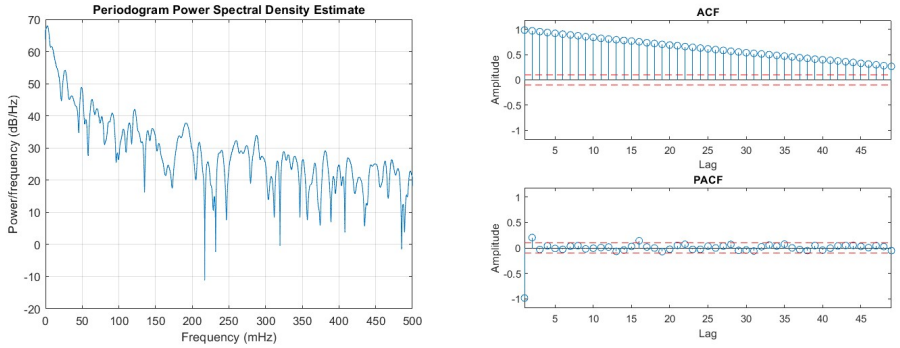

Figure 4.6: Periodogram using a Hanning window and 16,384 discrete Fourier transform points (left) and ACF and PACF of the data (right)

with the same length as the data was applied. The absence of prominent peaks in the periodogram suggests that there are no recurring patterns or cyclic behavior influencing the floor price.

To assess the need for differencing in the data, the autocorrelation function (ACF) and partial autocorrelation function (PACF) were examined. The ACF and PACF plots, displayed on the right side of Figure 4.6, provide insights into the correlation structure of the data.

Observing the ACF plot, it can be noted that the autocorrelation decays slowly, indicating a long-term dependence between observations. Additionally, the PACF plot exhibits a significant spike at lag 1, suggesting a strong component that requires differencing.

To address this, the first-order differencing operator, denoted as ∇_1, was incorporated into the modeling process. This differencing operation aims to eliminate the long-term dependence and make the data more stationary.

The subsequent step involves model selection. Two different models were fitted, beginning with the explanation of the autoregressive (AR) model, which was later compared against its time-recursive version. The differentiated signal's resulting ACF and PACF indicate a strong AR(1) component, as illustrated in the left side of Figure 4.7. To incorporate the AR(1) component, the prediction error method estimator in MATLAB was utilized, resulting in updated ACF and PACF plots shown on the right side of the same figure.

To assess the presence of outliers, the trimmed ACF (TACF) was plotted together with the ACF using a trim factor of 0.02. The objective was to determine if any remaining outliers were corrupting the ACF. However, since both func-

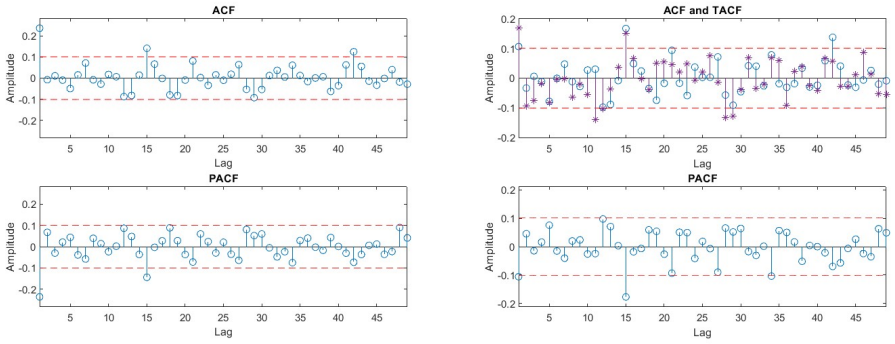

Figure 4.7: ACF and PACF of the differentiated signal (left) and residuals of the model (right)

tions showed a close resemblance, it was concluded that no outliers needed to be removed.

By applying the "1 out of 20" rule, which allows for a small percentage of residuals to fall slightly outside the confidence intervals, the residuals can be considered as white noise. This indicates that they exhibit randomness and do not possess any discernible patterns. As a result, a significant amount of information from the original signal has been extracted and captured in the model.

The whiteness of the residuals was further examined using the Monti test, which suggested that they can be considered white. However, it is important to note that this test's reliability is contingent on the normality assumption of the PACF. In this case, the PACF was found to deviate from normality according to the Jarque-Bera test. Consequently, the results of the Monti test should be interpreted cautiously and may not be fully reliable.

The resulting AR model for the transformed signal is as follows:

$$A(z) = 1 - 1.236z^{-1} + 0.2365z^{-2} \tag{4.3}$$

For the second model, a Kalman filter was implemented with the AR polynomial initialized using the previously calculated values. The filter was applied to the entire signal, and the resulting parameters show slight differences compared to the original model. A visual comparison of the parameters in both models is presented in Figure 4.8.

For large-step predictions in the first model, solving the Diophantine equation is necessary. However, in this case, where it is assumed that the exchange of money between parties occurs on a daily basis, the one-step prediction can be directly formed using the AR model. This simplification allows for straightforward forecasting of the floor price.

The results obtained from both models, AR and time-recursive, have been plotted together in Figure 4.9.

The variances of both residuals were computed in the test dataset as a performance indicator to determine the best fit for the data. In this case, the AR model exhibits a slightly lower variance compared to the Kalman filter. Furthermore,

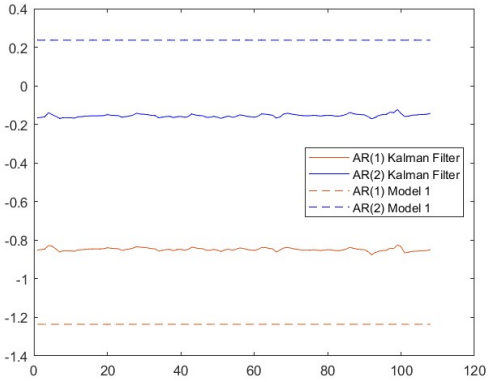

Figure 4.8: AR coefficients comparison in the test dataset

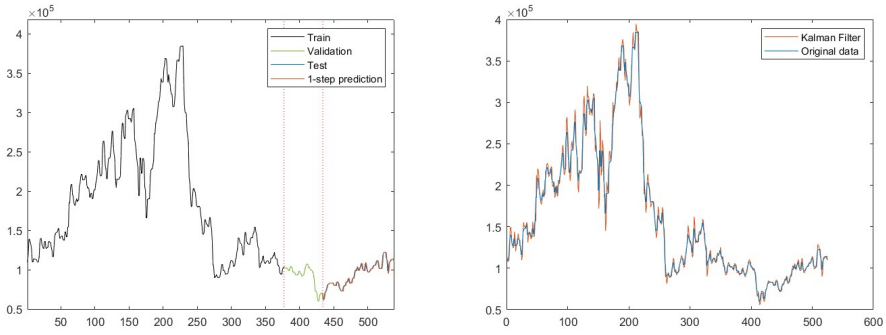

Figure 4.9: Comparison of the original data and the one-step predictions using the AR model (left) and the Kalman filter (right)

the whiteness tests, such as the Monti test, indicated that the residuals of the AR model can be characterized as white noise. In contrast, the residuals of the Kalman filter model show some patterns or structure, indicating that not all the information has been fully captured. Figure 4.10 provides a visual comparison of the ACFs for both sets of residuals.

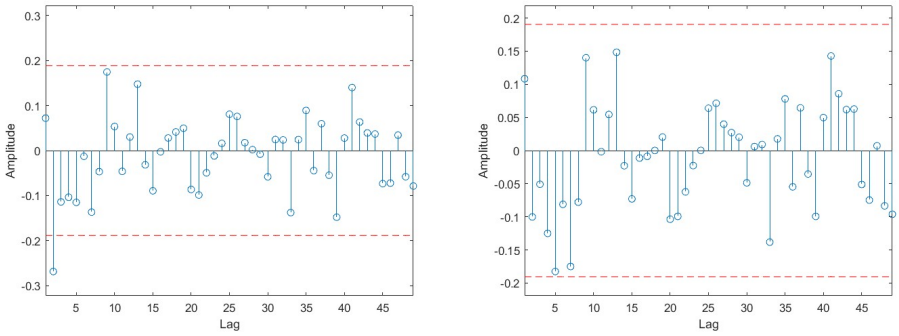

Figure 4.10: ACF of the residuals for the Kalman filter model (left) and the AR model (right)

Based on these evaluations, the AR(2) model is deemed more suitable for the BAYC collection, as it demonstrates better performance in terms of variance and whiteness of residuals. Therefore, the AR(2) model will be utilized to form the one-step linear predictions for the floor price of the BAYC collection.

To incorporate rarity into the pricing mechanism, a rule needs to be established, wherein tokens with a higher degree of rarity receive an increase in their price. To calculate the rarity scores for each token, the API provided by the website in [106] was utilized. These scores were then stored alongside the remaining token data in the fact table, as depicted in the diagram shown in Figure 3.7.

The values were calculated using a linear approach, where the percentage of items sharing each of the token's traits was added together. Tokens with lower values indicate a higher degree of scarcity, as there are fewer tokens with those specific traits. This scarcity generates a perception of uniqueness, which in turn increases the token's price.

To determine the prices of the tokens, it is utilized exponential interpolation. In this pricing approach, the token with the least rarity (highest score) is assigned a 0% increase above the floor price. Conversely, the token with the lowest rarity score is assigned the maximum increase in price, using the BAYC sales prices as a reference point.

To further investigate the price differences, an analysis of the last sales for the rarest and least rare tokens was conducted. Two tokens, namely item #3009 (the 5th rarest) and item #8272 (the 2nd least rare), were selected to determine the disparity in prices that people are willing to pay. These tokens were sold within a close time frame, allowing for a meaningful assessment of the price differences.

The activity record on OpenS ea revealed that item #3009 was sold at a price

almost three times higher than the other token. While it is important to note that the price agreed upon in a sale is influenced by various unquantifiable external factors, this information provides a valuable initial reference for the pricing mechanism. Therefore, it has been decided to assign a factor of three times the floor price increase to the rarest token.

By incorporating this price disparity observation, the pricing system can account for the varying preferences and perceived value of tokens with different levels of rarity. The mathematical expression of the pricing algorithm is as follows:

$$\text{item price}_{t+1} = \hat{fp}_{t+1}(1 + \frac{2}{1 - e^{-k}}(e^{-k \cdot score} - e^{-k})) \tag{4.4}$$

Here, $\hat{fp}_{t+1}$ represents the one-step prediction of the floor price based on the available information up to time t. The variable *score* corresponds to the normalized value of the calculated rarity score for the token under consideration. The parameter k is used to control the steepness of the interpolation curve.

Setting an appropriate value for k is crucial, as it influences the rate at which token prices change, especially for tokens with highly specific and unique traits. To illustrate the impact of different k values, Figure 4.11 displays the shape of the curves for various values of k. After testing different values, it has been determined

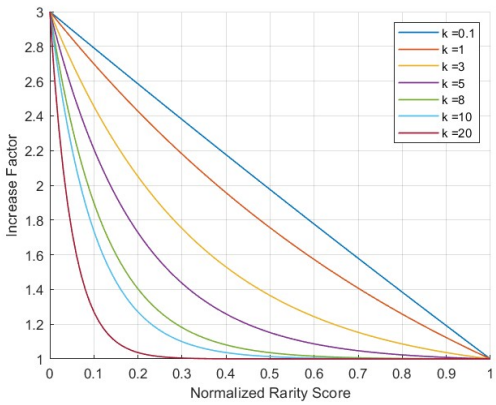

Figure 4.11: Interpolation curves for different values of k

that a value of 5 yields accurate results for different tokens. Hence, the decision has been made to proceed with this value for k.

Chapter 5
Results

This chapter presents a real case scenario demonstrating the application of the pricing model to calculate the monthly risk premium for a policyholder seeking an insurance policy for an asset in their collection. Additionally, it explores various risk prevention measures that insurers can consider implementing as requirements for policyholders. Furthermore, potential enhancements to the developed model are discussed, aiming to further improve its performance.

5.1 Estimation of the Risk Premium

To estimate the risk premium, the Jupyter notebook in [107] was developed to simulate a realistic scenario where a policyholder obtains an insurance policy to cover the risks associated with one of the items from the BAYC collection during April 2023. The risk premium was calculated for cases where the insured utilized either a hot or cold wallet, allowing for a comparison between the two groups at the end of the month.

The theoretical results were validated by comparing them with Monte Carlo simulations, which involved generating one million different scenarios to calculate each value of the risk premium. The theoretical results were derived using the mathematical expression presented in Equation 4.1. For the simulated results, a Bernoulli distribution was utilized, with the probability of success determined based on the wallet type. The alpha distribution with the corresponding parameters outlined in Table 4.1 was used for each wallet type. Additionally, the item selection was conducted randomly using a uniform distribution. Consequently, running the code repeatedly will generate calculations for different tokens. To replicate calculations for a specific item, its ID must remain constant.

It has been represented in Figure 5.1 a comparison of the daily values of the risk premium for each wallet type using Monte Carlo simulations and the analytical expressions. Simulated and analytical results exhibit remarkable consistency, providing validation for the developed mathematical expressions. The risk premium for hot wallets initially surpasses that of cold wallets, but as the month progresses, the two converge. This behavior can be attributed to the shape of the PDFs, as detailed in Figure 4.4. To ensure a fair pricing mechanism, it is suggested to periodically update certain parameters of the alpha distributions. This adjustment is necessary because cold wallets pose lower risks compared to hot wallets, and it

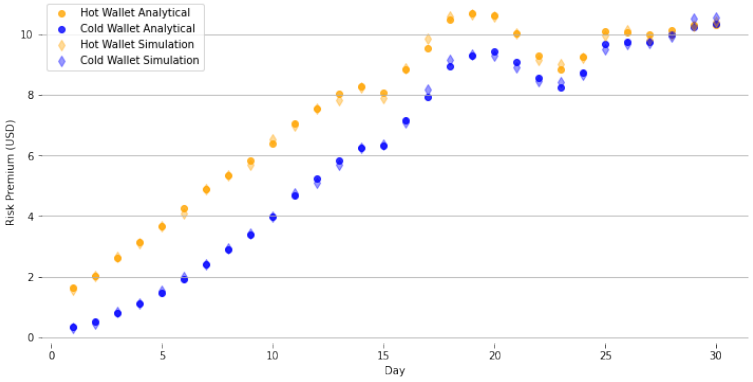

Figure 5.1: Risk premium calculation for the month of April 2023 using simulation and analytical expressions

would not be fair for users opting for safer options to pay higher premiums.

For this specific random token with ID 6444, the lump risk premium at the end of the month amounts to $227 for hot wallets, whereas it is $184 for cold wallets. Insurers can then determine the premium based on their business needs, incorporating an appropriate percentage increase over the calculated risk premium.

5.2 Risk Prevention Measures

It is important for insurers to advice their customers on the best security practices to follow, especially when entering emerging markets like the NFT market, where risks are still being discovered. Developing accurate mathematical models to estimate risk transfer from policyholders to insurers is essential, but it is also worth considering a shift towards a paradigm of risk prevention rather than relying solely on risk transfer.

To ensure precautionary measures, insurers could establish basic requirements for users seeking insurance policies, particularly during the initial years. For instance, users might be required to install phishing site detection software, such as AegisWeb3 developed by PeckShield, or utilize a cold wallet, provided by the insurance company, where the token has to reside in order to be eligible for the compensation, except in some specific circumstances that would require to deposit it in another place. Furthermore, it is important to clearly define the risks covered by the insurance policy and explicitly exclude certain risks, avoiding vague descriptions as mentioned earlier in the text. For example, over-the-counter trades could be excluded from coverage, encouraging users to exercise extreme caution when engaging in such transactions.

Implementing these risk prevention measures would benefit both parties involved. Users would face fewer risks, resulting in potentially lower premiums and increased coverage availability. Insurers, on the other hand, would experience fewer claims, leading to improved loss ratios — the ratio of claim amount to earned premiums. It would also allow insurers to segment clients more effectively, as their

risk profiles would be better defined. This way, insurers could make more precise underwriting decisions and develop more accurate pricing models.

The implementation of these measures can be further promoted through economic incentives, as discussed in the next section. By providing rewards or discounts for policyholders who actively adhere to risk prevention measures, insurers can encourage and reinforce positive behaviors that mitigate potential risks.

5.3 Usage-based Insurance for NFTs

In recent years, a new form of insurance known as usage-based insurance has emerged, particularly in the realm of car insurance. Unlike traditional methods that rely solely on standard rating factors, this innovative approach incorporates the user's driving behavior into the premium calculation. The objective is to establish a fairer pricing mechanism wherein individuals who adopt safer measures are duly rewarded, while those who do not face penalties through dynamic adjustments in their premiums based on these measured factors. Furthermore, this approach effectively serves as an educational tool by economically incentivizing users to integrate risk prevention measures into their daily routines.

However, one of the drawbacks associated with usage-based insurance is the requirement for insurers to install devices in the insured vehicles to measure various factors and assess the riskiness of the user's driving habits. This implementation raises potential privacy concerns, particularly regarding the constant monitoring of drivers through GPS location tracking. Questions may arise regarding the legality and ethical implications of continuously monitoring a driver's whereabouts to determine the safety level of the roads they travel on or the prevalence of accidents in those areas.

Analogously, a similar concept can be applied to digital assets, and insurers can reap significant benefits from the underlying technology. One notable advantage is the elimination of concerns surrounding user privacy violations, thanks to the inherent properties of the blockchain. Being permissionless, open, and transparent, the blockchain allows insurers to track the complete history of a covered token's activities, enabling them to assess the safety of a user's transactions.

In the current policy framework, the determination of premiums for different groups has relied solely on the type of wallet used. However, when considering the risks associated with digital assets, it becomes evident that there are numerous other significant factors that have not been utilized due to limited available data. These additional parameters play a crucial role in assessing risk, and insurers can incorporate them to create more robust pricing mechanisms. Some of them will be explored.

5.3.1 Approved Addresses

As previously mentioned, one common way in which users can lose control of their tokens is by inadvertently granting access to attackers, a practice known as ice phishing. To mitigate this risk, various entities, such as the blockchain explorer Etherscan, provide tools [109] that display all approved addresses for managing different types of tokens associated with a particular public address and

offer the functionality to revoke active allowances. These tools analyze the entire transaction history linked to the address and store the addresses that have been approved through transactions containing the `setApprovalForAll` function. By using the Etherscan approval checker, users can conveniently verify the approved addresses for their tokens.

There are specific addresses that require approval, often associated with smart contracts used in DeFi applications. Examples include Seaport in OpenSea or BenDAO's token managers, which allow users to stake NFTs or use them as collateral for borrowing ETH. These approvals are commonly observed, and their reliability can be evaluated by examining the source code of the respective smart contracts.

However, assessing approved EOAs poses a challenge due to the anonymous nature of the blockchain. It is not always possible to determine if a token owner has intentionally granted access to a particular EOA. Nonetheless, by continuously evaluating all approved addresses, it is possible to detect if a blacklisted address associated with phishing or other incidents is granted approval. This proactive approach can help prevent theft by alerting the user to potentially malicious activity.

To mitigate risks associated with unknown or untrusted approved addresses, insurers can consider implementing a premium increase if the number of approved addresses with unknown reputations exceeds a certain threshold. This slight premium adjustment would differentiate insured individuals who grant access only to reputable and trusted actors from those with a higher number of unknown or potentially risky approvals.

5.3.2 Sweeping the Floor

The term "sweep the floor" in the NFT ecosystem refers to a practice wherein the floor price of a collection is artificially inflated by repeatedly buying and selling tokens to wallets owned by the same collector. This creates a false perception of increased demand and interest in the project, with the intention of attracting traders using algorithms that detect profitable collections.

However, this technique can become problematic, particularly in collections with a small number of items, where the floor price can be easily manipulated. Users can collude to inflate the floor price to a level where the compensation an insurer would have to pay may exceed the previous value of the token. In such cases, it becomes advantageous for users to willingly give away the token in exchange for economic compensation.

To address this fraudulent activity, insurers can actively monitor the activity of a collection. When indications of wash trading aimed at sweeping the floor are detected, it is crucial to impose a penalty on the compensation to be perceived by the user.

5.3.3 Social Media Accounts Protection

Instances of hackers gaining unauthorized access to official social media accounts and using them to post phishing links have led to significant losses. While this is a factor beyond the control of individual users, an analogy to the car man-

ufacturing industry can be drawn.

Car manufacturers have recognized the importance of implementing safety mechanisms in cars, not only for driver protection but also because it translates into better insurance prices for end users. According to an article cited in [108], the average cost of full coverage car insurance in the USA in 2023 is $1601. The type of car being insured partially influences this price, with insurance companies offering discounts for vehicles equipped with advanced safety measures.

Manufacturers have then been driven to develop new and sophisticated safety measures to capture buyers' attention and compete with other brands. This competition raises industry standards, resulting in better products for consumers. Insurance companies play a role in driving this improvement by incentivizing the implementation of advanced safety measures.

Applying this analogy to the NFT ecosystem, slight adjustments in the risk premium value of collection items could be considered based on the safety measures implemented by the developers for their social media accounts. Teams that prioritize security measures, such as implementing multi-factor authentication, restricting access to a limited number of individuals, utilizing anti-phishing software, or adding additional verification requirements for users, would receive favorable treatment. This would influence users' decisions when choosing tokens from different collections with similar item prices.

By incorporating these factors into the policies, teams would increase their efforts to develop robust protection measures, addressing one of the weakest points that has resulted in users losing their tokens.

In the current landscape, user vigilance is crucial when engaging with others to avoid being scammed. This constant need for caution has undermined the overall user experience. Many creators have even added alerts to their nicknames on different platforms, such as "WILL NEVER DM YOU FIRST", as a precautionary measure against phishing attempts as shown in Figure 5.2 where both founders

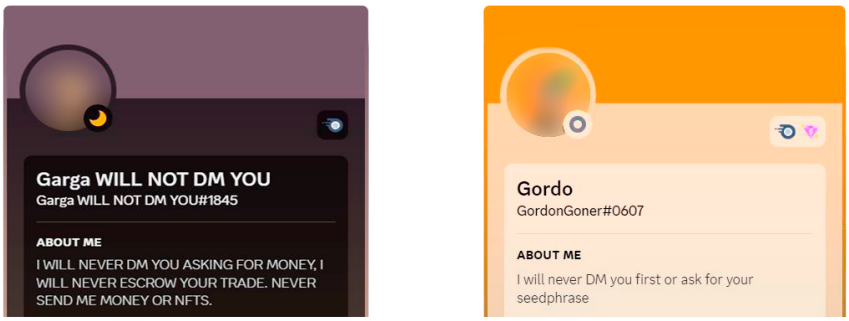

Figure 5.2: Discord nickname and personal information of the BAYC founders

of the BAYC collection added a description in their profile stating that users will never be contacted by them via private message. In the event that such contact does occur, it is highly likely that their account has been compromised. Apart from that, the BAYC collection serves as a benchmark for implementing additional safety measures. They have specific channels in their server dedicated to posting

the project's official links and informing users about ongoing scams. Making an effort in this direction, the impact on user experience could result into a more appealing space for current and new participants.

It is important to note that even with all the implemented measures by the teams, it is the end user who must carefully evaluate their interactions and follow strict verification procedures to avoid falling victim to the many attackers constantly looking for their prey. As an example, Figure 5.3 displays a message received within five minutes after joining the BAYC Discord server from an unverified member promoting a fake free mint website for a collection whose minting stage ended a long time ago, likely leading to a wallet drain. There is a common saying in the space that users should always keep in mind before pulling the trigger: "If it seems too good to be true, it probably is".

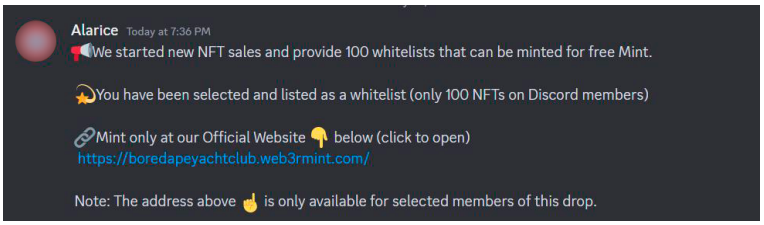

Figure 5.3: Message received in Discord after joining the BAYC's server from an unverified member

5.3.4 Marketplaces and Bridges

Lastly, it is proposed to consider monitoring the platforms where users deposit their tokens. In Chapter 3, a subsection was dedicated to discussing exploits in various NFT-based protocols, including marketplace code exploits and cross-chain bridge exploits.

The insurer could adjust premiums based on the reputation of the marketplace where policyholders list their items for sale. Hackers have discovered ways to exploit protocols in marketplaces, ranging from major platforms like OpenSea to smaller ones like TreasureDAO. However, it is likely that well-established marketplaces with larger teams of developers and extensive third-party audits will be less susceptible to exploits resulting from faulty tools.

A similar approach can be applied to bridges. Bridging NFTs, while not yet a common practice, may become more prevalent in the coming years as the space continues to expand. Insurers can choose to penalize the use of insecure bridges by token owners when shifting between domains. An example of poor safety measures could be a trusted bridge where the federation of relayers consists of very few nodes, most of them owned by the same entity providing the bridging services. An incident last year highlighted the consequences of such vulnerabilities, which resulted in one of the biggest thefts in the history of the space [110].

There are many other aspects insurers could assess when determining the pre-

mium rates. Here some of the considered to be most important are highlighted. This approach aims to enhance the users experience by incentivizing them to adopt basic safety measures, leading to a lower number of compromised wallets and avoiding common attacks that can be easily prevented. Moreover, it promotes fair pricing, making insurance more affordable for conscientious users who are less exposed to risks, thereby expanding the potential customer base.

The final model could be represented as follows:

$$\text{risk premium} = \text{base risk premium} \pm \text{UBI rate} \tag{5.1}$$

The base risk premium is calculated using the mathematical expressions developed in Chapter 4, while the UBI rate is an additional parameter that imposes a maximum percentage increase or decrease on the base risk premium, based on users' behavior. It is important to define the allowable range for this percentage increase or decrease, as it provides users with a reference point for comparing different policies. A clear and transparent description of the UBI rate ensures that users can easily compare prices across different companies and have confidence in the long-term cost of their chosen policy.

Chapter 6

Future Work and Conclusions

Reflecting on the four key questions raised in the introduction, it can be concluded that the project has successfully addressed its objectives. However, it is important to acknowledge that from the outset, the project aimed to create a framework for a future insurance policy. The statistical significance of the analysis conducted relies heavily on the quality and quantity of the data utilized. Unfortunately, due to constraints in time and resources, the issue of insufficient data has not been fully addressed as would be necessary for a fully developed policy. It is worth noting that in certain jurisdictions, insurers are legally obliged to present statistical data to support new rating structures.

This paper can serve as a foundation for insurance companies to take their initial steps in this promising market and gain an early advantage in the competition to gather information. Over time, those who embark on this endeavor sooner will be better positioned to capitalize on the benefits arising from a sector projected to experience exponential growth in the coming years.

List of Acronyms

ACF	Autocorrelation Function
API	Application Programming Interface
AR	Autoregressive
BAYC	Bored Ape Yacht Club
BIP	Bitcoin Improvement Proposals
BMP	Bitmap
CDF	Cumulative Distribution Function
CID	Content Identifier
DAG	Directed Acyclic Graph
DAO	Decentralized Autonomous Organization
DHT	Distributed Hash Table
DeFi	Decentralized Finance
EIP	Ethereum Improvement Proposals
EOA	Externally Owned Account
ERC	Ethereum Request for Comment
ETH	Ethereum
EVM	Ethereum Virtual Machine
FOMO	Fear Of Missing Out
GIF	Graphics Interchange Format
HTTP	Hypertext Transfer Protocol
IP	Intellectual Property
IPFS	InterPlanetary File System
IPLD	InterPlanetary Linked Data
JPEG	Joint Photographic Experts Group

JSON	JavaScript Object Notation
MIME	Multipurpose Internet Mail Extensions
NFT	Non Fungible Token
P2E	Play To Earn
PACF	Partial Autocorrelation Function
PDF	Probability Density Function
PNG	Portable Network Graphics
PoS	Proof Of Stake
PoW	Proof Of Work
SFT	Semi-Fungible Token
SVG	Scalable Vector Graphics
TACF	Trimmed Autocorrelation Function
UBI	Usage-Based Insurance
URI	Uniform Resource Identifier
UTXO	Unspent Transaction Output
XML	Extensible Markup Language

References

[1] A. Zuckerman. 'Insuring Crypto: "The Birth of Digital Asset Insurance'." December 29, 2020. University of Illinois Journal of Law, Technology and Policy. Available at SSRN: https://ssrn.com/abstract=3756619orhttp://dx.doi.org/10.2139/ssrn.3756619 (accessed May 25, 2023)

[2] "Introduction to Web3." ethereum.org. https://ethereum.org/en/web3/ (accessed Mar. 15, 2023).

[3] him.eth [@himgajria]. Web 1: Read Web 2: Read-Write Web 3: Read-Write-Own. *Twitter*. May. 29, 2020. Available: https://twitter.com/himgajria/status/1266415636789334016 (accessed Mar. 15, 2023)

[4] D. Tapscott. "How the blockchain is changing money and business." Sep. 16, 2016. [Online video]. Available: https://www.ted.com/talks/don_tapscott_how_the_blockchain_is_changing_money_and_business (accessed Mar. 21, 2023)

[5] S. Nakamoto. "Bitcoin: A Peer-to-Peer Electronic Cash System." Oct. 31, 2008. Available: https://bitcoin.org/bitcoin.pdf (accessed Mar. 22, 2023)

[6] "Byzantine Fault Tolerance Explained." academy.binance.com. https://academy.binance.com/en/articles/byzantine-fault-tolerance-explained (accessed Mar. 23, 2023)

[7] M. Palatinus, P. Rusnak, A. Voisine and S. Bowe. "Mnemonic code for generating deterministic keys." GitHub repository. Sept. 10, 2013. Available: https://github.com/bitcoin/bips/blob/master/bip-0039.mediawiki. (accessed: Mar. 25, 2023)

[8] P. Wuille. "Hierarchical Deterministic Wallets." GitHub repository. Feb. 11, 2012. Available: https://github.com/bitcoin/bips/blob/master/bip-0032.mediawiki. (accessed: Mar. 25, 2023)

[9] M. Palatinus, P. Rusnak. "Multi-Account Hierarchy for Deterministic Wallets." GitHub repository. Apr. 24, 2014. Available: https://github.com/bitcoin/bips/blob/master/bip-0044.mediawiki. (accessed: Mar. 25, 2023)

[10] "What Is a Blockchain Consensus Algorithm?" academy.binance.com. https://academy.binance.com/en/articles/what-is-a-blockchain-consensus-algorithm (accessed Mar. 26, 2023)

[11] "Bitcoin Energy Consumption Index." digiconomist.net. https://digiconomist.net/bitcoin-energy-consumption (accessed Mar. 26, 2023)

[12] "In Proof of Stake, what is the hash value parameter in randomized block select?" ethereum.stackexchange.com. https://ethereum.stackexchange.com/questions/126504/in-proof-of-stake-what-is-the-hash-value-parameter-in-randomized-block-select (accessed: Mar. 29, 2023)

[13] Slance. "What is Proof of Stake - Explained in Detail (Animation)." Nov. 23, 2021. [Online Video]. Available: https://www.youtube.com/watch?v=YudpU58uYuM&ab_channel=Slance (accessed Mar. 29, 2023)

[14] "Proof Of Stake (POS)." ethereum.org. https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/ (accessed Mar. 29, 2023)

[15] "Pool Distribution." btc.com. https://btc.com/stats/pool?pool_mode=year (accessed Mar. 29, 2023)

[16] "Ethereum's energy expenditure." ethereum.org. https://ethereum.org/en/energy-consumption/ (accessed Mar. 29, 2023)

[17] V. Buterin. "On Public and Private Blockchains." blog.ethereum.org. https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains (accessed Mar. 31, 2023)

[18] B. Crozier, Speaker. "How Allianz took a blockchain platform from pilot to 1 million transactions." *CIO Priorities 2022*. 2022. Info-Tech Research Group [Podcast]. Available: https://www.infotech.com/research/how-allianz-took-a-blockchain-platform-from-pilot-to-1-million-transactions. (accessed Apr. 19, 2023)

[19] "Layer 2." ethereum.org. Available: https://ethereum.org/en/layer-2/. (accessed: Apr. 21, 2023)

[20] V. Buterin. "A Next-Generation Smart Contract and Decentralized Application Platform." Ethereum Whitepaper. 2014. Available: https://ethereum.org/en/whitepaper/. (accessed: Apr. 20, 2023)

[21] "Ethereum Virtual Machine (EVM)." ethereum.org. 2021. Available: https://ethereum.org/en/developers/docs/evm/. (accessed: Apr. 20, 2023)

[22] A. Hamilton. "The Beginning Of NFTs - A Brief History Of NFT Art." Mar. 6, 2023. Available: https://www.zenofineart.com/blogs/news/the-beginning-of-nfts-a-brief-history-of-nft-art. (accessed: Apr. 23, 2023)

[23] M. Rosenfeld. "Overview of Colored Coins." Dec. 4, 2012. Available: https://bitcoil.co.il/BitcoinX.pdf. (accessed: Apr. 20, 2023)

[24] K. McCoy. "Quantum." May. 3, 2014. [Online image]. Available: https://ipfs.io/ipfs/QmPkJoCk1vZ7wGMhfDer9fwpQtRGWwPn7NrocaCn7JS2SM. (accessed Apr. 22, 2023)

[25] "Details for Transaction fa8b9a6ad4d266f...d3bb48f8d." chainz.cryptoid.info. Available: https://chainz.cryptoid.info/nmc/tx.dws?1217290.htm. (Accessed: Apr. 22, 2023)

[26] E.Lee. "Lawsuit Against Sotheby's and Kevin McCoy Dismissed." nftnow.com. Mar. 21, 2023. [Online]. Available: https://nftnow.com/news/lawsuit-against-sothebys-and-kevin-mccoy-dismissed/. (accessed: Apr. 22, 2023)

[27] "Contract 0xE81a4543...9E578F8771D9." etherscan.io. Available: https://etherscan.io/address/0xe81a45439ff9bc5841202ce4b2049e578f8771d9. (accessed: Apr. 23, 2023)

[28] "CryptoPunks." larvalabs.com. Available: https://www.larvalabs.com/cryptopunks. (accessed: Apr. 23, 2023)

[29] W. Entriken , D. Shirley, J. Evans and N. Sachs. "Non-Fungible Token Standard." GitHub repository. Jan. 24, 2018. Available: https://github.com/ethereum/EIPs/blob/master/EIPS/eip-721.md. (accessed: Apr. 24, 2023)

[30] Larvalabs. "CryptoPunks Composite Image." [Online image]. Available: https://www.larvalabs.com/public/images/cryptopunks/punks.png. (Accessed: Apr. 23, 2023)

[31] CryptoPunks [@cryptopunksnfts]. The Cryptopunks are now fully on chain!. *Twitter*. Aug. 18, 2021. Available: https://twitter.com/cryptopunksnfts/status/1428099416326557696 (accessed Apr. 23, 2023)

[32] M. Marlinspike. "My first impressions of web3." moxie.org. Available: https://moxie.org/2022/01/07/web3-first-impressions.html (accessed May 06, 2023)

[33] Takens Theorem. "Souls of Immortal NFTs." medium.com. Available: https://medium.com/etherscan-blog/souls-of-immortal-nfts-de212a840de5 (accessed May 06, 2023)

[34] dom [@dhof]. the michelin guide to ""on chain"" nfts. *Twitter*. Jun. 30, 2021. Available: https://twitter.com/dhof/status/1410060181849919489 (accessed May 06, 2023)

[35] IPFS. "How IPFS Deals With Files - IPFS Camp Workshop." *YouTube*. Sep. 17, 2019. Available: https://www.youtube.com/watch?v=Z5zNPwMDYGg&ab_channel=IPFS. (accessed: Apr. 26, 2023)

[36] IPFS. "How IPFS Deals With Files - IPFS Camp Workshop." *YouTube*. Sep. 17, 2019. Available: https://www.youtube.com/watch?v=Z5zNPwMDYGg&ab_channel=IPFS. (accessed: Apr. 26, 2023)

[37] "How IPFS works." docs.ipfs.tech. Available: https://docs.ipfs.tech/concepts/how-ipfs-works/#subsystems-overview. (accessed: Apr. 26, 2023) https://www.youtube.com/watch?v=Z5zNPwMDYGg&ab_channel=IPFS. (accessed: Apr. 26, 2023)

[38] "What is calldata?" ethereum.stackexchange.com. Available: https://ethereum.stackexchange.com/questions/52989/what-is-calldata. (accessed: May 06, 2023)

[39] "0xmons v2 Cthulhu: On-chain Encoding." blog.0xmons.xyz. Available: https://blog.0xmons.xyz/79081566310. (accessed: May 06, 2023)

[40] "About 0xDEAFBEEF." deafbeef.com. Available: https://www.deafbeef.com/about.htm. (accessed: May 06, 2023)

[41] "Smart contract details 0xd754937672300Ae6708a51229112dE4017810934." etherscan.io. Available: https://etherscan.io/address/0xd754937672300ae6708a51229112de4017810934#readContract. (accessed: May 06, 2023)

[42] S. de la Rouviere. "Flavours of On-Chain SVG NFTs on Ethereum." blog.simondlr.com. Available: https://blog.simondlr.com/posts/flavours-of-on-chain-svg-nfts-on-ethereum. (accessed: May 06, 2023)

[43] "Smart contract details 0x960b7a6bcd451c9968473f7bbfd9be826efd549a." etherscan.io. Available: https://etherscan.io/address/0x960b7a6bcd451c9968473f7bbfd9be826efd549a#readContract. (accessed: May 06, 2023)

[44] R. Delgado. "NFTMetadataStorage.xlsx." GitHub repository. May 06, 2023. Available: https://github.com/rdf5/insurancenft/blob/main/NFTMetadataStorage.xlsx. (accessed: May 06, 2023)

[45] "Collection stats." opensea.io. Available: https://opensea.io/rankings?sortBy=total_volume. (accessed: May 06, 2023)

[46] nick.eth [@nicksdjohnson]. Everyone is making NFTs that generate their artwork... *Twitter*. Aug. 27, 2021. Available: https://twitter.com/nicksdjohnson/status/1431144024052690944 (accessed May 06, 2023)

[47] "CC0 'No Rights Reserved'." creativecommons.org. Available: https://creativecommons.org/share-your-work/public-domain/cc0/. (accessed: May 06, 2023)

[48] "Metadata Standards." docs.opensea.io. Available: https://docs.opensea.io/docs/metadata-standards. (accessed: May 08, 2023)

[49] "Smart contract details 0xe8d8c0a6f174e08c44ab399b7ce810bc4dce096a." etherscan.io. Available: https://etherscan.io/address/0xe8d8c0a6f174e08c44ab399b7ce810bc4dce096a#readContract. (accessed: May 06, 2023)

[50] "Smart contract details 0xb1bEfc9E7B76C1e846EBBf3e6E1Ab029C86e7435." etherscan.io. Available: https://etherscan.io/address/0xb1bEfc9E7B76C1e846EBBf3e6E1Ab029C86e7435#readContract. (accessed: May 06, 2023)

[51] "Moonbirds art, preserved on the Ethereum blockchain forevermore." proof.xyz. Available: https://www.proof.xyz/moonbirds/in-chain. (accessed: May 06, 2023)

[52] "Smart contract details 0x94cB646dD34b3B0fF7C116208F7f7fF7Ac216079." etherscan.io. Available: https://etherscan.io/address/0x94cB646dD34b3B0fF7C116208F7f7fF7Ac216079#readContract. (accessed: May 07, 2023)

[53] Hyperloot #1 JPG file. metadata.hyperlootproject.com. Available: https://images.hyperlootproject.com/nft/1.jpg. (accessed: May 06, 2023)

[54] "Cryptokitties sales history." kittyhelper.co. Available: https://kittyhelper.co/sales-history/?period=custom&d1=2017-12-01&d2=2017-12-31&sort=1. (accessed: Apr. 27, 2023)

[55] "Tradition and Innovation Collide: Decentraland Metaverse Fashion Week 2023." decentraland.org. Feb. 27, 2023. Available: https://decentraland.org/blog/announcements/tradition-and-innovation-collide-decentraland-metaverse-fashion-week-2023. (accessed: Apr. 29, 2023)

[56] "Decentraland EST 4339." nonfungible.com. Available: https://nonfungible.com/market-tracker/decentraland/EST/4339. (accessed: Apr. 29, 2023)

[57] "OpenSea - Our Story." opensea.io. Available: https://opensea.io/abou. (accessed: May 08, 2023)

[58] "NFT Marketplaces." dappradar.com. Available: https://dappradar.com/nft/marketplaces?period=all. (accessed: May 08, 2023)

[59] "Beeple (b. 1981), Everydays: The First 5000 Days." onlineonly.christies.com. Available: https://onlineonly.christies.com/s/beeple-first-5000-days/beeple-b-1981-1/112924. (accessed: Apr. 29, 2023)

[60] "Blockchains by NFT Sales Volume." cryptoslam.io. Available: https://www.cryptoslam.io/blockchains. (accessed: May. 1, 2023)

[61] P. Smith, Presenter J. Krcmar, Speaker. "Session 2." *NFT Educational Series*. Jun. 30, 2021. Institutes RiskStream Collaborative [Podcast]. Available: https://vimeo.com/572091172. (accessed: May. 1, 2023)

[62] Mimecast. "The State of Email Security 2022." Available: https://www.mimecast.com/state-of-email-security/. (accessed: May. 2, 2023)

[63] R. de Best. "Total value of cryptocurrency lost to and recovered from theft and other attacks between March 2020 and February 2022." Feb. 3, 2022. Available: https://www.statista.com/statistics/1285057/crypto-theft-size/. (accessed: May. 2, 2023)

[64] E. Arda, M. Nadini, C. De Pow and T. Annison. "NFTs and Financial Crime." Available: https://hub.elliptic.co/reports/nfts-and-financial-crime/. (accessed: May. 2, 2023)

[65] B. Lindrea. "Crypto insurance a 'sleeping giant' with only 1% of investments covered." cointelegraph.com. Sep. 12, 2022. Available: https://cointelegraph.com/news/crypto-insurance-a-sleeping-giant-with-only-1-of-investments-covered. (accessed: May. 2, 2023)

[66] Fortune Business Insights. "Global Cryptocurrency Market, Insights and Forecasts, 2017-2028." fortunebusinessinsights.com. Available: https://www.fortunebusinessinsights.com/industry-reports/cryptocurrency-market-100149. (accessed: May. 2, 2023)

[67] P. Ricard, J. Zwick, U. Koyluoglu, A. Flint and C. Freeman. "Will Web3 Reinvent Insurance?." oliverwyman.com. Available: https://www.oliverwyman.com/our-expertise/insights/2022/jul/oliver-wyman-will-web3-reinvent-insurance.html. (accessed: May. 3, 2023)

[68] Coinbase. "Insurance Coverage." coinbase.com. Available: https://www.coinbase.com/legal/insurance. (accessed: May. 3, 2023)

[69] Chainlink. "What Is a Blockchain Oracle?." chain.link. Available: https://chain.link/education/blockchain-oracles. (accessed: May. 3, 2023)

[70] Coinbase. "Insurance Coverage." coinbase.com. Available: https://www.coinbase.com/legal/insurance. (accessed: May. 3, 2023)

[71] J. Kagan. "Mutual Insurance Company: Definition and How They Invest." investopedia.com. Available: https://www.investopedia.com/terms/m/mutual-insurance-company.asp. (accessed: May. 3, 2023)

[72] Nexus Mutual. "Documentation." nexusmutual.io. Available: https://docs.nexusmutual.io/. (accessed: May. 3, 2023)

[73] K. Petrie. "Nexus Mutual is now using Chainlink's price reference data contracts for decentralized valuations of the multi-currency capital pool." medium.com. Available: https://medium.com/nexus-mutual/nexus-mutual-is-now-using-chainlinks-price-reference-data-contracts-for-decentralized-valuations-6a62c5d4e030. (accessed: May. 3, 2023)

[74] Nexus Mutual. "Smart Contracts Details." nexusmutual.io. Available: https://api.nexusmutual.io/sdk/. (accessed: May. 3, 2023)

[75] Nexus Mutual. "Cover Underwritten and Claims Paid." nexusmutual.io. Available: https://nexusmutual.io/. (accessed: May. 3, 2023)

[76] C. Seifert. "'Ice phishing' on the blockchain." microsoft.com. Available: https://www.microsoft.com/en-us/security/blog/2022/02/16/ice-phishing-on-the-blockchain/. (accessed: May 10, 2023)

[77] jeffnicholas.eth [@_jeffnicholas_]. Today has been rough. *Twitter*. Aug. 25, 2021. Available: https://twitter.com/_jeffnicholas_/status/1430323445057744897. (accessed May 09, 2023)

[78] A. Sarkar. "OpenSea planned upgrade stalls as phishing attack targets NFT migration." cointelegraph.com. Available: https://cointelegraph.com/news/opensea-planned-upgrade-stalls-as-phishing-attack-targets-nft-migration. (accessed: May 09, 2023)

[79] D. Barda, R. Zaikin and O. Vanunu. "Check Point Research Prevents Theft Of Crypto Wallets On OpenSea, The World's Largest NFT Marketplace." research.checkpoint.com. Available: https://research.checkpoint.com/2021/check-point-research-prevents-theft-of-crypto-wallets-on-opensea-the-worlds-largest-nft-marketplace/. (accessed: May 10, 2023)

[80] quit [@0xQuit]. Today, bored ape holder "s27" lost their bubble gum ape and matching mutants. *Twitter*. Apr. 5, 2022. Available: https://twitter.com/0xQuit/status/1511198290565509120. (accessed May 10, 2023)

[81] "Part 3: Set your drop earnings." docs.opensea.io. Available: https://docs.opensea.io/docs/part-3-set-your-drop-earnings. (accessed: May 11, 2023)

[82] "What are OpenSea's fees?" support.opensea.io. Available: https://support.opensea.io/hc/en-us/articles/14068991090067-What-are-OpenSea-s-fees-. (accessed: May 11, 2023)

[83] W. Gottsegen. "Uncanceled Listings." coindesk.com. Jan. 24, 2022. Available: https://www.coindesk.com/layer2/2022/01/28/openseas-week-from-hell/. (accessed May 11, 2023)

[84] N. Bambysheva and M. G. Santillana. "Over $3 Billion Stolen In Crypto Heists: Here Are The Eight Biggest." forbes.com. Available: https://www.forbes.com/sites/ninabambysheva/2022/12/28/over-3-billion-stolen-in-crypto-heists-here-are-the-eight-biggest/?sh=3cb28490699f. (accessed: May 11, 2023)

[85] A. Bhuptani "The Interoperability Trilemma." blog.connext.network. Available: https://blog.connext.network/the-interoperability-trilemma-657c2cf69f17. (accessed: May 11, 2023)

[86] Amber Group. "Reproducing the $APE Airdrop Flash Loan Arbitrage/Exploit." medium.com. Available: https://medium.com/amber-group/reproducing-the-ape-airdrop-flash-loan-arbitrage-exploit-93f79728fcf5. (accessed: May 12, 2023)

[87] "Transaction Details." etherscan.io. Available: https://etherscan.io/tx/0xeb8c3bebed11e2e4fcd30cbfc2fb3c55c4ca166003c7f7d319e78eaab9747098. (accessed: May 12, 2023)

[88] J. Niset. "Argent Smart Wallet Specification." GitHub repository. Apr. 23, 2021. Available: https://github.com/argentlabs/argent-contracts/blob/develop/specifications/specifications.pdf. (accessed: May 14, 2023)

[89] "Ethereum Chain Full Sync Data Size." ycharts.com. Available: https://ycharts.com/indicators/ethereum_chain_full_sync_data_size (accessed: May 16, 2023)

[90] ZachXBT [@zachxbt]. *Twitter*. Available: https://twitter.com/zachxbt. (accessed: May 26, 2023)

[91] PeckShieldAlert [@PeckShieldAlert]. *Twitter*. Available: https://twitter.com/PeckShieldAlert. (accessed: May 26, 2023)

[92] M. White. "Web 3 is Going Just Great." web3isgoinggreat.com. Available: https://web3isgoinggreat.com/. (accessed: May 26, 2023)

[93] R. Delgado. "Stolen NFTs Database.ipynb." GitHub repository. May 17, 2023. Available: https://github.com/rdf5/insurancenft/blob/main/Stolen%20NFTs%20Database.ipynb. (accessed: May 17, 2023)

[94] "Transaction Details." etherscan.io. Available: https://etherscan.io/tx/0xafc951c5aadc63dbff23ca7b628b36d9faf35ae38b484f91db1134b5558cb01d.

[95] "Transaction Details." etherscan.io. Available: https://etherscan.io/tx/0xd82484e970a1a0a065f4e710da84990df5cee35e2305fcf88db44271a24c5ceb.

[96] Serpent [@Serpent]. Analysis of how a scammer stole 14 BAYCs worth over 852 ETH. *Twitter*. Dec. 17, 2022. Available: https://twitter.com/Serpent/status/1604074440941506560. (accessed May 17, 2023)

[97] "Transaction Details." etherscan.io. Available: https://etherscan.io/tx/0xf1877ae321b3e9dbf871d4f026df434fe12fc1ad3f64ce61e97789bc2e33ad07. (accessed: May 17, 2023)

[98] "Transaction Details." etherscan.io. Available: https://etherscan.io/tx/0xd554a83a3e4ff332620048c747647051d6d01ed465dd84ba55a2b9d918b80cc1. (accessed: May 17, 2023)

[99] "Etherscan API Plans." etherscan.io. Available: https://etherscan.io/apis. (accessed: May 17, 2023)

[100] R. Delgado. "exp_var.csv." GitHub repository. May 17, 2023. Available: https://github.com/rdf5/insurancenft/blob/main/exp_var.csv. (accessed: May 17, 2023)

[101] S. Glen. "Alpha Distribution." statisticshowto.com. Available: https://www.statisticshowto.com/alpha-distribution/. (accessed: May 20, 2023)

[102] "scipy.stats.alpha." docs.scipy.org. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.alpha.html. (accessed: May 20, 2023)

[103] beetle. "Total OS Marked Stolen Assets in Top Collections." dune.com. Available: https://dune.com/beetle/opensea-stolen-assets-top-pfp-collections. (accessed: May 20, 2023)

[104] R. Delgado. "BAYC_Floor_Price_Modeling." GitHub repository. May 22, 2023. Available: https://github.com/rdf5/insurancenft/tree/main/BAYC_Floor_Price_Modeling. (accessed: May 22, 2023)

[105] "NFT Price Floor." nftpricefloor.com. Available: https://nftpricefloor.com/en. (accessed: May 26, 2023)

[106] "rarity.tools." rarity.tools.com. Available: https://rarity.tools/. (accessed: May 26, 2023)

[107] R. Delgado. "Monte Carlo Simulation.ipynb." GitHub repository. May 23, 2023. Available: https://github.com/rdf5/insurancenft/blob/main/Monte%20Carlo%20Simulation.ipynb. (accessed: May 23, 2023)

[108] P. Gusner and L.Masterson. "Average Cost Of Car Insurance 2023." forbes.com. Available: https://www.forbes.com/advisor/car-insurance/average-cost-of-car-insurance/. (accessed: May 23, 2023)

[109] "Ethereum Token Approval." etherscan.io. Available: https://etherscan.io/tokenapprovalchecker. (accessed: May 27, 2023)

[110] P. Jha. "The aftermath of Axie Infinity's $650M Ronin Bridge hack." cointelegraph.com. Available: https://cointelegraph.com/news/the-aftermath-of-axie-infinity-s-650m-ronin-bridge-hack. (accessed: May 24, 2023)