Mapping the NFT revolution: market trends, trade networks and visual features

Matthieu Nadini,[1,2] Laura Alessandretti,[3] Flavio Di Giacinto,[1,4]
Mauro Martino,[5] Luca Maria Aiello,[6] and Andrea Baronchelli[1,2,7,*]

[1] *Department of Mathematics, City University of London, EC1V 0HB, London, UK*
[2] *The Alan Turing Institute, British Library, 96 Euston Road, NW12DB, London, UK*
[3] *Technical University of Denmark, DK-2800 Kgs. Lyngby, DK*
[4] *Department of Neuroscience, Catholic University of the Sacred Heart, Rome, IT*
[5] *IBM Research, Cambridge MA, USA*
[6] *IT University of Copenhagen, DK*
[7] *UCL Centre for Blockchain Technologies, University College London, London, UK.*

(Dated: June 4, 2021)

Abstract

Non Fungible Tokens (NFTs) are digital assets that represent objects like art, videos, in-game items and music. They are traded online, often with cryptocurrency, and they are generally encoded as smart contracts on a blockchain. Media and public attention towards NFTs has exploded in 2021, when the NFT art market has experienced record sales while celebrated new star artists. However, little is known about the overall structure and evolution of the NFT market. Here, we analyse data concerning 6.1 million trades of 4.7 million NFTs generating a total trading volume of 935 millions US dollars. Our data are obtained primarily from the Ethereum and WAX blockchains and cover the period between June 23, 2017 and April 27, 2021. First, we characterize the statistical properties of the market. Second, we build the network of interactions and show that traders have bursts of activity followed by inactive periods, and typically specialize on NFTs associated to similar objects. Third, we cluster objects associated to NFTs according to their visual features and show that NFTs within the same category tend to be visually homogeneous. Finally, we investigate the predictability of NFT sales. We use simple machine learning algorithms and find that prices can be best predicted by the sale history of the NFT collection, but also by some features describing the properties of the associated object (e.g., visual features of digital images). We anticipate that our analysis will be of interest to both researchers and practitioners and will spark further research on the NFT production, adoption and trading in different contexts.

* To whom correspondence should be addressed: abaronchelli@turing.ac.uk

I. INTRODUCTION

"WTF are NFTs? Why crypto is dominating the art market" is the title of the February 21, 2021 episode of *The Art Newspaper* podcast [1], signalling both the impact of Non Fungible Tokens (NFTs) on the art world and the novelty they represent for most of the general public. The revolution is not confined to the art market. While NFT adoption in gaming has already reached a certain maturity, for example concerning the trade of in-game objects, different other industries, especially involved with the production of digital content such as music or video, are experimenting with the technology. Overall, in the first four months of 2021 the NFT volume has exceeded 2 USD billions, ten times larger than the entire trading volume in 2020 [2].

So, what's an NFT? An NFT is a unit of data stored on a blockchain that certifies a digital asset to be unique and therefore not interchangeable, while offering a non-duplicable digital certificate of ownership for the NFT [3]. More broadly, an NFT allows to establish the "provenance" of the assigned digital object offering indisputable answers to such questions as who owns, previously owned and created the NFT, as well as which of the many copies is the original. Several digital objects can be associated to an NFT including photos, videos, audio, and other types of digital files, and NFTs are now being used to commodify digital objects in different contexts, such as art, gaming and sport collectibles. Most NFTs are part of the Ethereum blockchain but other blockchains can implement their own versions of NFTs [4].

The first example of NFTs used to represent digital art concerns CryptoKitties, a blockchain game on Ethereum that allows players to purchase, collect, breed and sell virtual cats [5]. In December 2017, the game congested the Ethereum network [6]. By many considered a chief example of the irrationality driving the cryptocurrency market in 2017 [7], Cryptokitties remained the only popular example of NFTs for almost two years. In December 2020, the market of NFT art started to grow again [2] and attracted a huge attention in March 2021, when the artist known as Beeple sold an NFT of his work sold for $69.3 million at Christie's [8]. The purchase resulted in the third-highest auction price achieved for a living artist, after Jeff Koons and David Hockney [9]. Several other record sales followed: three Cryptopunks, which are randomly generated set of 10,000 unique digital characters, were sold at $7.5, $1.54, and $1.3 million dollars, respectively; the first tweet was sold at $2.9 million dollars; and the Auction Winner Picks Name, an NFT with music video and dance track, sold at $1.33 million dollars [10]. NFTs profitability has attracted several celebrities, who created their own NFTs, as well as the most popular sports, with collectibles of NBA and famous football players that are currently sold for hundreds of thousands dollars [11].

Researchers have just started looking at NFTs, often focusing on specific aspects. For example, the study of copyright regulations was explored in ref. [3], technical details, such as, NFTs components, protocols, standards, and desired properties in ref. [12], while possible new blockchain-based protocol to trace physical goods in ref. [13]. An overview of the implications that NFTs have on art is done in references [14, 15], where the blockchain-based NFTs are presented as an evolution of the 1917 Artist's contract, introduced in 1971 to share secondary sale royalties with the artist. Other studies have focused on a limited number of similar NFTs, such as CryptoKitties [16, 17], or NFTs offered on Decentraland [18], SuperRare [19, 20], or a combination of Decentraland, Cryptopunks, and Axie [21]. Main takeaways of previous findings are that the digital abundance of NFTs in digital games has led to their valuelessness [16], and, even if overall NFT prices is driven by cryptocurrencies pricing [21], individuals with inside knowledge of how the NFTs system work can take advantage of it [17, 18]. A relational aspect is also important, where the NFT success is linked with its recognition from experts [19], while the study of a co-ownership network of 16,000 NFTs on the SuperRare market suggests that a highly centralized, small-world structure emerges [20].

In this paper, we analyse a large dataset aiming at providing a first quantitative overview of NFT market in the most prominent domains where NFTs are currently used. To this end, we analyse data concerning 6.1 million trades of 4.7 million digital art pieces tracking primarily the Ethereum and WAX blockchains and covering the period between June 23, 2017 and April 27, 2021. The article is organized as follows. In Section II A, we present an overall analysis of the statistical properties of the NFT market and its evolution over time. In Section II B, we study the network of interactions between NFT traders, and the network of NFT assets, in which two assets are linked if purchased sequentially over time by the same trader. In Section II C, we cluster NFTs based on their visual features. In Section II D and IV F, we present the results of regression and classification models predicting the occurrence and the price of NFT secondary sales.

Before continuing, it is worth stressing that the exact categorisation of the different domains in which NFTs are used is—obviously—outside of the scope of the present paper. For example, "art" objects can be in some cases classified as "collectibles", while some "in-game" objects may present sophisticated aesthetic and cultural properties that may qualify them as "art". Here, we categorise NFTs based on their practical usage and general perception adopting the classification proposed by NonFungible Corporation [22], a specialized company that track NFTs sales, and categorizing the largest remaining collections by manual inspection.

II. RESULTS

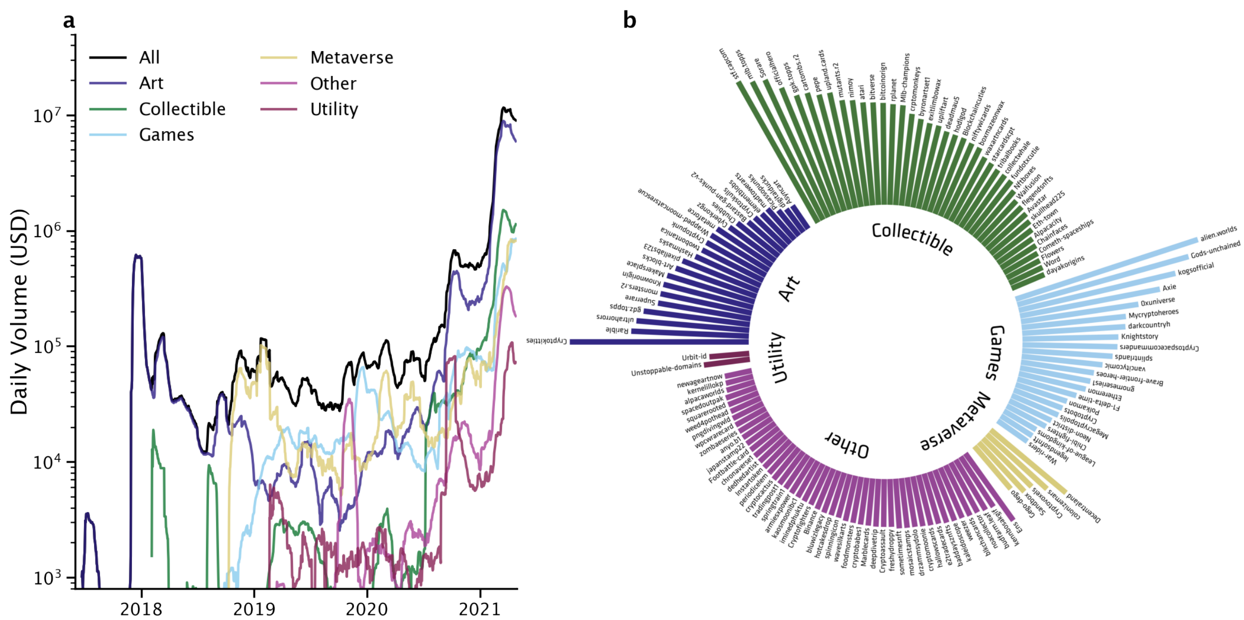


FIG. 1: **The NFT Market. (a)** Top 100 NFTs collections (by number of assets) organized by category. The height of each bar is proportional to the number of assets in each collection. **(b)** Daily volume (in USD) exchanged over time for each category and for all assets (see legend). Results are averaged over a rolling window of 30 days. Days with volume below 1,000 USD are not shown.

A. The NFT Market

Following an initial rapid growth in late 2017, when CryptoKitties gained worldwide popularity, the size of the NFT market has remained substantially stable until mid 2020, with an average of $\sim 60,000$ US dollars traded daily (see Figure 1a). Starting from July 2020, the market has experienced a dramatic growth, with the total volume exchanged daily surpassing ~ 10 million US dollars in March 2021, when was 150 times larger compared to 8 months before.

Items exchanged on the NFT market are organized in *collections*, sets of NFTs that, in most cases, share some common features. Collections can be widely different in nature, from sets of collectible cards, to selections of art masterpieces, to virtual spaces in online games. Roughly speaking, according to the definitions made by NonFungible Corporation [22], most collections can be categorised in six categories: Art, Collectible, Games, Metaverse, Other, and Utility (see also Figure S1). In Figure 1b, we show the top 100 collections by number of unique assets n, organized by Category. The size of collections is well described by a power-law function $P(n) \sim n^{-\alpha}$, with $\alpha = 1.5$ (see Figure 3c). We find that $\sim 75\%$ of collections comprise less than 37 unique assets, and $\sim 1\%$ have more than 10400 unique assets.

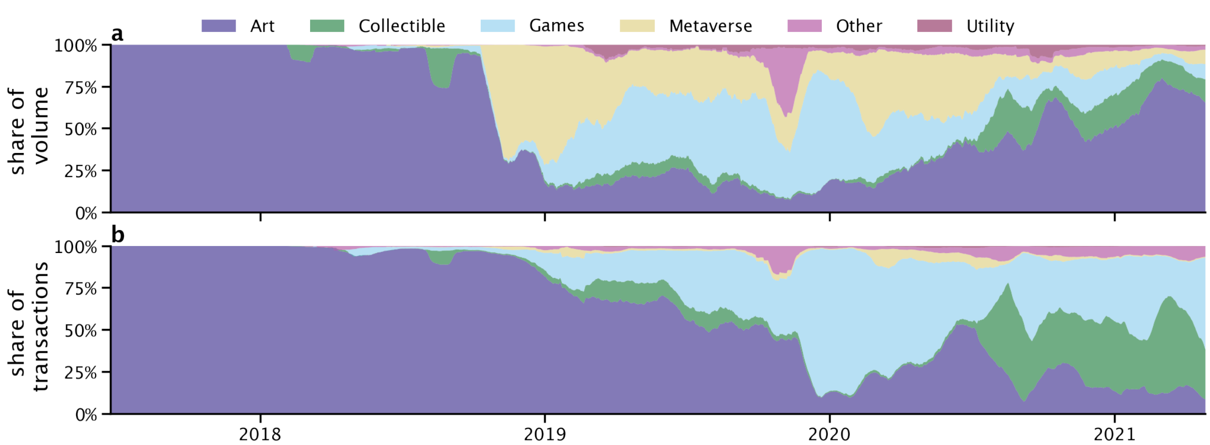


FIG. 2: **Composition of the NFT market. (a)** Share of volume traded by category. **(b)** Share of transactions by category. Results are averaged over a rolling window of 30 days.

A first question is how much different NFTs categories contribute to the whole size of the NFT market. Until the end of 2018, the market was fully dominated by the Art category (see Figure 1), and, in particular, by the Cryptokitties collection. From January 2019, other categories started gaining popularity, both in terms of total volume exchanged (Figure 2a) and number of transactions (Figure 2b).

Overall, in the period between January 2019 and July 2020, $\sim 90\%$ of the total volume exchanged on NFT was shared by the Art, Games and Metaverse categories, contributing 18%, 33% and 39% respectively (see Figure 2a). Starting from mid July, 2020, instead, in terms of total volume spent, the market has been largely dominated by NFTs categorized as Art, which, since then, has contributed $\sim 71\%$ of the total transaction volume, followed by *Collectible* accounting for 12%. Importantly, however, the market composition is quite different when considering the number of transactions (see Figure 2b). Since July 2020, the most exchanged NFTs belong to the categories *Games* and *Collectible*, which account for 44% and 38% of transactions. Instead, only 10% of transactions are related to NFTs categorized as *Art*. Overall, we observe that the share of volume spent in *Art* has been growing since 2020, while the share of transactions has been decreasing (see Figure 2). The discrepancy between volume and transactions reveals that prices of items categorized as *Art* are higher, on average, compared to other categories.

We dig further into these differences by looking at the distribution of NFT prices across categories (see Figure 3a), which we find to be broadly distributed. We observe that the average sale price of NFTs is lower than 10 dollars for 75% of the assets, and larger than 1594 dollars, for 1% or assets. Considering individual categories, NFTs categorized as *Art*, *Metaverse*, and *Utility* reached higher prices compared to other categories, with the top 1% of assets having average sale price higher than 6290, 9485, and 12756 dollars respectively. Note that these categories are different in sizes, so 1% of assets corresponds to 8593, 472, and 78 NFTs in the *Art*, *Metaverse* and *Utility* categories, respectively. The highest prices so far were reached by assets categorized as *Art*, with 4 NFT that were sold for more than 1 million dollars.

Another interesting question is how many times individual assets are traded. Here, we refer to the first time an asset is sold as the asset's *primary sale*, and to all other sales as *secondary sales*. All assets considered in this study had a primary sale, but only $\sim 20\%$ of them had a secondary sale. We observe that the distribution of number of sales s per asset has two regimes: for $s \geq 10$ the distribution is well characterized by a power-law function $P(s) \sim s^{-\beta}$, with $\beta = 1.4$ [23] (see Figure 3b). Note that only 0.07% of all assets are sold more than 10 times. Temporal patterns of secondary sales are shown in Figure 4.

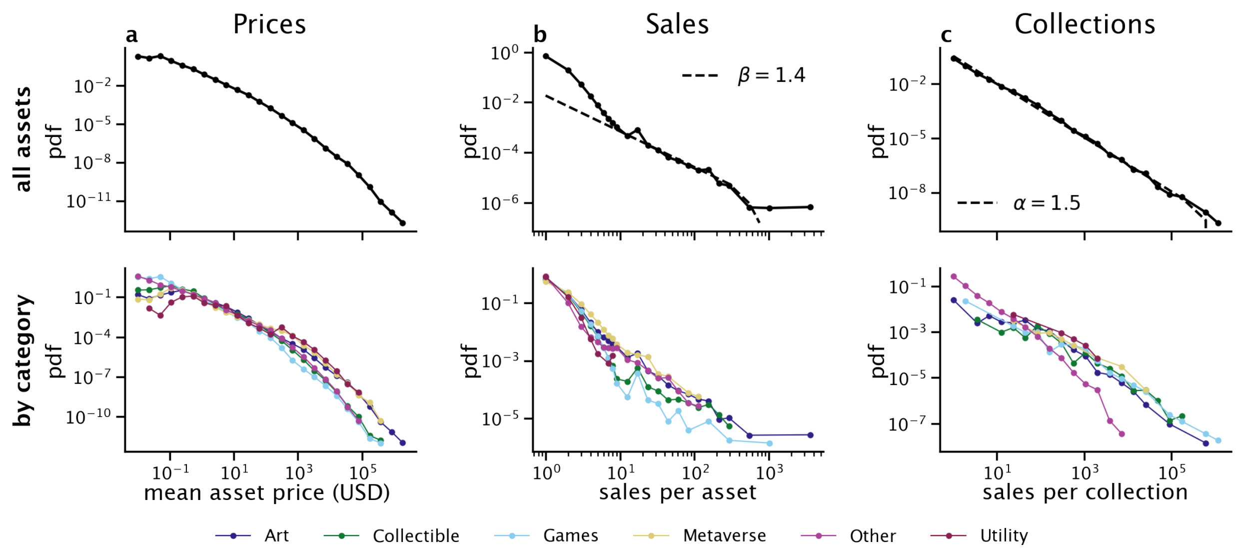


FIG. 3: **Statistical properties of the NFT market. (a)** Distribution of the average price (USD) for all NFTs (top) and by NFT category (bottom). **(b)** Distribution of number of sales per NFT for all NFTs (top) and by category (bottom). The dashed line is a power law fit $P(s) \sim s^{-\beta}$, with $\beta = 1.4$, where s is the number of sales. **(c)** Distribution of number of assets per collection for all NFTs (top) and by category (bottom). The dashed line is a power law fit $P(n) \sim n^{-\alpha}$, with $\alpha = 1.5$, where n is the number of unique assets.

B. The networks of NFT trades

We consider two temporal and directed networks, namely 1) the network of buyer-seller trades and 2) the network of NFTs. The trader network has a directed link drawn from a buyer to a seller when the former purchases an NFT from the latter. It allows to track the behaviour of each trader (either buyer or seller) with respect to the others. The NFT network is constructed by linking NFTs that are purchased in a sequential order by the same buyer: a directed link created from an NFT to another NFT when a buyer purchases first the former and then the latter (see Section IV D for more details). This second network reveals how NFTs are related with one another over time, allowing the study of buyers' temporal purchase patterns.

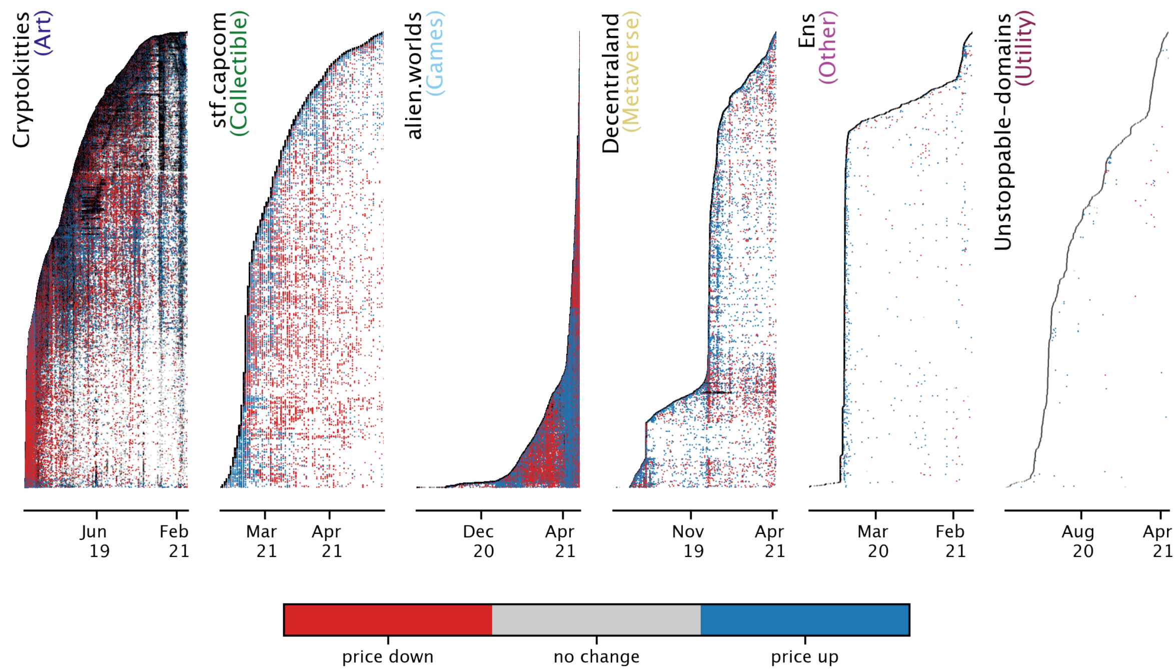


FIG. 4: **Secondary sale prices.** Sales over time for the top collection in terms of number of sales in each NFT category (Criptokitties, Stf.capcorn, Alien, Decentraland, Miscellanea, and Unstoppable). Each horizontal line represent an NFT and each dot represent a sale. Sales are coloured based on the change in price compared to previous sale (see colorbar).

We start by analysing the trader network and examining a key network property, the traders' strength, which represents the number of either purchases or sales each trader has ever made. Figure 5a shows that the decay of the probability distribution function follows a power law with exponent $\lambda_1 = -1.85$. With the top 10% of all traders that perform 85% of all transactions and trade at least once 97% of all NFTs. A similar trend is detected when the number of transactions between a buyer-seller pair is considered. With the top 10% pairs that have the same weight as the remaining 90%, see Figure S2a for more details. Also, traders active a large number of days, trade a lot more than traders active only for few days, with their their strength that increases following a power law with positive exponent $\lambda_2 = 1.28$, as shown in Figure 5b. Traders in our dataset are highly specialized and trade most of the times in only one, or few, collections. Figure 5c shows that, independently from the traders' strength, traders perform at least 73% of their transactions in their top collection, while at least 82% in their top two collections combined.

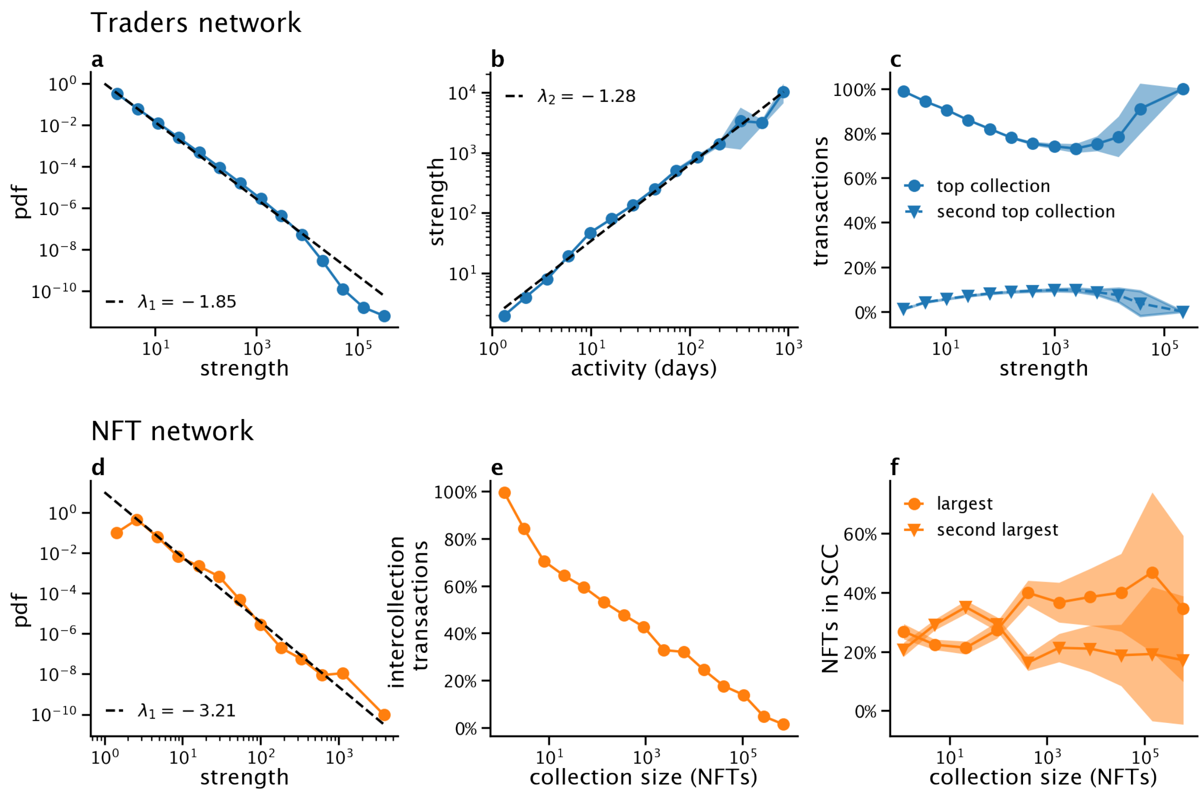
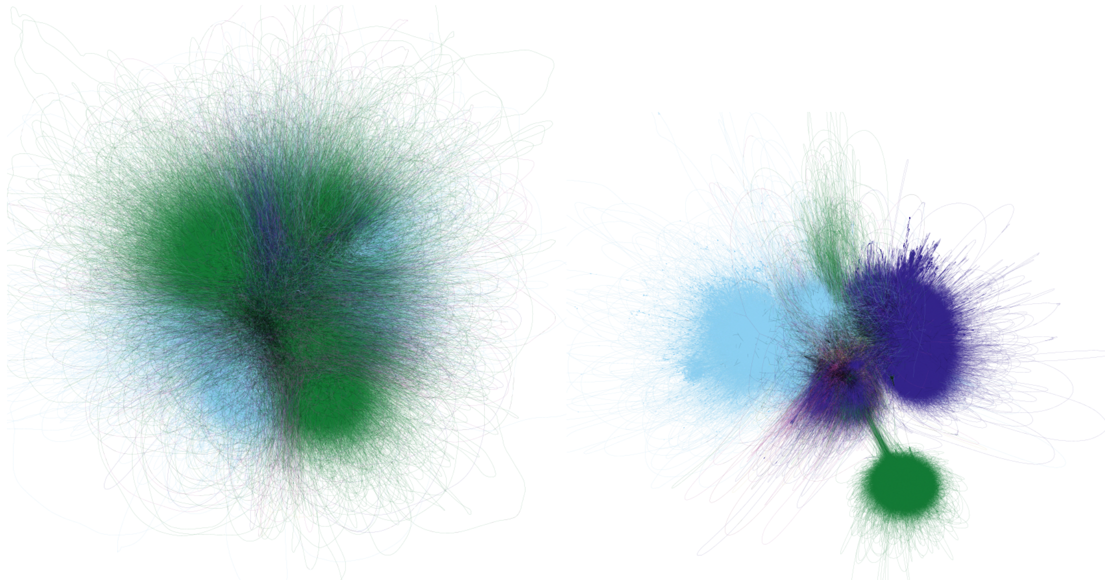


FIG. 5: **Key network properties.** (a) Probability distribution function of the traders' strength. (b) Relationship between the traders' strength and the number of days in which they are active. (c) Percentage of transaction traders make toward their top and second-top NFT collections. (d) Probability distribution function of the NFTs' strength. (e) Percentage of transactions between NFTs in different collections as a function of the size of the collection. (f) Percentage of NFTs belonging to the first and second largest strong connected component (SCC). Solid curves in panels (b)-(c)-(e)-(f) represent average values, while respective bands the 95% confidence interval.

By analysing the traders' behaviour in relation to their strength, we observe that the most specialized traders have either few or tens of thousands transactions. While traders with few transactions can only trade in few collections, traders with many transactions specialize themselves in specific digital art or games. An example is the trader, with Ethereum address "0xfc624f8f58db41bdb95aedee1de3c1cf047105f1", that exchanges tens of thousands CriptoKitties. Similar relationships hold when the behaviour of buyers and seller is separately considered, as shown in Figure S2.

We now turn the analysis to the NFT network, where we can explore how each NFT is connected with the others. Figure 5d illustrates that NFTs strength follows a power law with exponent $\lambda_3 = -3.21$. NFTs with high strength are linked to many others because purchased right before, after, or together with other "NFTs". The next question we ask is: which are these other "NFTs"? Figure 5e shows that NFTs in small collections tend to be bought in sequence with NFTs in other collections. On the contrary, NFTs in large collections, like CryptoKitties or Gods-Unchained, tend to be bought in sequence with NFTs in the same collection. The NFTs network is highly modular and the collections well represent the underlining community structure (modularity $Q = 0.80$ [24]). Despite the NFT network is highly clustered, these communities are not isolated, meaning that their owners would buy few NFTs in other collections, thereby connecting them. Such behaviour allows to form large strong connected components (SCCs), where, inside each of them, by starting from any NFTs it is possible to reach any other NFTs in the SCC. There are two SCCs detected, the largest include tradings in the WAX blockchain and 35% of all NFTs, while the second largest exchanges in the Ethereum blockchain and 20% of all NFTs. Figure 5(e) shows the structure of these SCCs, which include NFTs in collections of all size. Traders' bursty-like behaviour causes the high network modularity, due to sequential purchases in the same collection, and the presence of large SCCs, due to less frequent purchases in other collections.

Key network properties previously discussed, like traders and NFTs strength distributions, hold when each NFTs category is considered separately from the others. Furthermore, traders, independently from the NFTs category considered, are highly specialized, with between 59% (in the *Other* category) and 98% (in the *Utility* category) trades that are performed in the trader's top collection. Relative to the number of total NFTs in each category, the largest SCC contains more than half (the 55.0%) of all NFTs labeled as *Collectible*, but only the 0.06% of all NFTs labeled as *Utility*. On the contrary, the second largest SCC has the 54.8% of all *Art*, but only the 10.6% of *Games*. Figure 6 offers a visual representation of the two SCCs, where we note that each category form partially independent clusters of connections. More specifically, the second largest SCC has few highly defined clusters, corresponding to the NFT collections CryptoKitties, in *Art*, Sorare, in *Collection*, and Gods-Unchained, in *Games*.



Largest SCC | Second largest SCC

FIG. 6: **SCCs in the NFT network.** Vertices are not represented. Link's color, when it connects NFTs in the same category, corresponds to the color of that category, while it is black otherwise. Visualization is done using Graph-tool: [25].

C. Graphical features

NFTs are linked to digital assets of different types, including videos, text, animated gifs, and audio. Currently, the vast majority consists of images. Therefore, we select the NFTs associated with images and animated GIFs, analyzing them with the pre-trained AlexNet convolution neural network and summarizing their graphical features in few meaningful values to investigate their relevance in the sales prediction. The 3d scatter plot in Figure 7 represents the first three principal components obtained through the application of PCA to the 4096-dimensional vectors from AlexNet CNN. The colors in Figure 7 highlighting the categories of the NFTs reveal the ability of the algorithm used for the visual feature extraction to identify common intra-categories characteristics, with images of similar appearance clustering in specific regions of the firsts three principal components space.

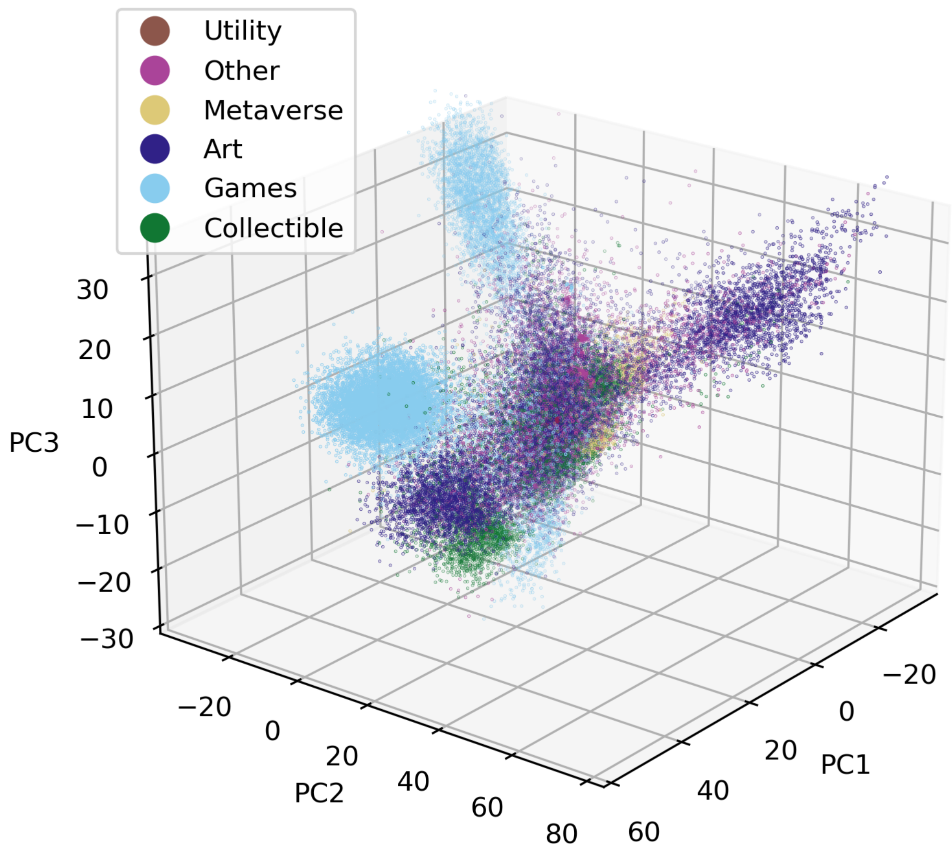


FIG. 7: **PCA of the vectors summarizing the visual features obtained from AlexNet pre-trained CNN.** The visual objects associated with the NFTs in the three-dimensional space identified by the PC1, PC2 and PC3, broken down by NFT categories.

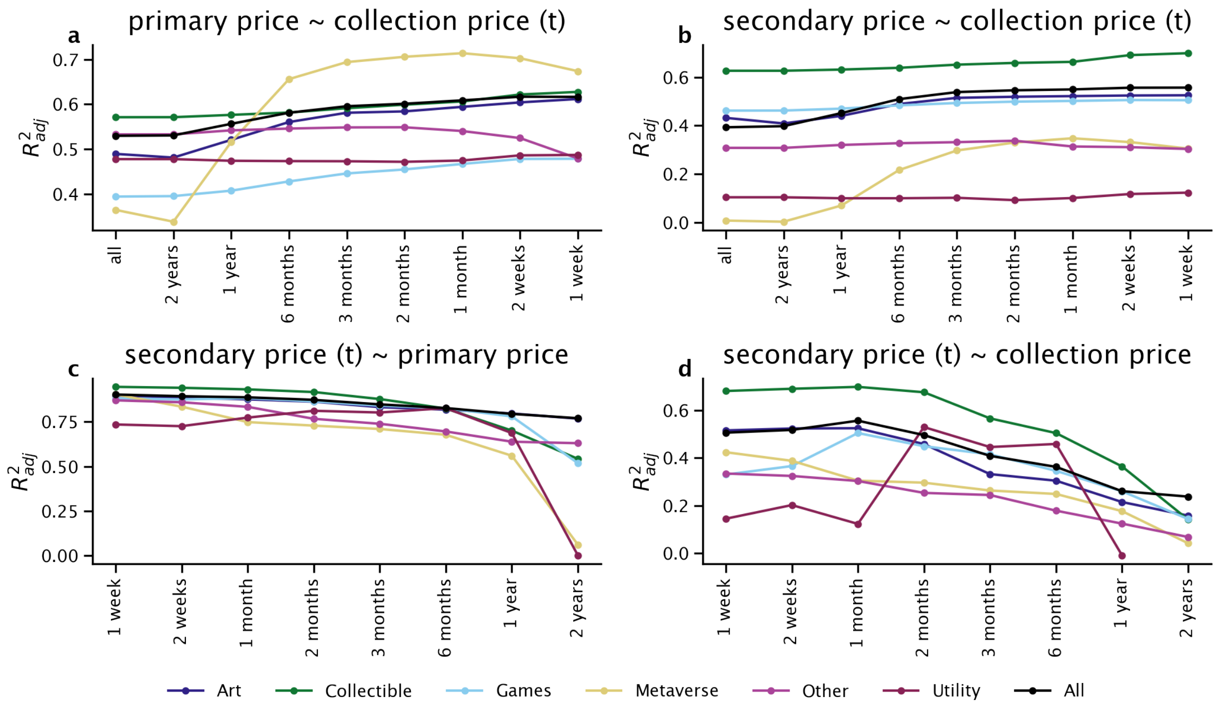


FIG. 8: **Primary and secondary price sale predictions.** Top: R^2_{adj} of a linear regression fit to predict **(a)** the price of primary sales, and **(b)** the median price of secondary sales 1 month after their respective primary sale from the historical median price of sale in the collection calculated over varying time windows (one week to two years) preceding the primary sale. Bottom: R^2_{adj} of a linear regression fit to predict **(c)** the price of secondary sales from the price of their respective primary sales, and **(d)** the price of secondary sales from the median price of sales in the NFT's collection in the previous week; we perform different regressions to predict the median price of secondary sales over varying time windows (one week to two years) after the primary sale. All results are broken down by NFT categories.

D. Predicting sales

To identify the factors associated with an NFT's market value, we fit a linear regression model to estimate the price of primary and secondary sales from different sets of features, calculated considering only the data preceding the day of the NFT's primary sale. The features (whose detailed formulations are provided in Section IV E) include the degree and PageRank centrality of the buyer and seller in the networks of NFT trades ($k_{buyer|seller}$, $PR_{buyer|seller}$), the principal components of visual features of the object linked to the NFT ($vis_{PCA_{1...5}}$), a prior probability of sale within the collection (p_{resale}), and the past median price of primary and secondary sales within the collection ($median\, price$).

Figure 8(a) shows that an NFT's price correlates strongly with the price of NFTs previously sold within the same collection. The median sale price of NFTs in the collection predicts more than half of the variance of price of future primary and secondary sales. The prediction is more accurate when the median of the past sale price is calculated over a recent time window preceding the primary sale, e.g., the prior time window of one week is better than considering the entire time window preceding the NFT's primary sale. Similar results, albeit with generally lower correlations, are found when the secondary sale price is the object of the regression, as shown in Figure 8(b). As one would expect, the price of secondary sales is strongly correlated with the price of primary sale, and the predictive power of the variables declines as one attempts to cast a prediction over longer periods of time: $R^2_{adj} = 0.90$ when predicting the media secondary sale price over the next week, and falls to $R^2_{adj} = 0.77$ when extending the prediction over the next 2 years.

Other features than prior sale history are predictive of future sale and secondary sale prices (Figure 9). Centrality measures of the buyer and seller in the trader network ($R_{adj.^2} \in [0.05, 0.12]$) and visual features of the object linked to the NFT ($R^2_{adj} \in [0, 0.08]$) explain roughly one-fifth to one-fourth of the variance when used in combination ($R^2_{adj} \in [0.18, 0.25]$). When considered in combination with the median price of previous sales, they increase the predictive power by almost 10% for the secondary sale price (R^2_{adj} from 0.55 to 0.6). When fitting separate regressions for each category, it becomes apparent that the predictability of future prices and the predictive power of different sets of features varies depending on the NFT category. The *collectible* category is the easiest to predict, with centrality and visual features yielding $R^2_{adj} \in [0.30, 0.36]$ and $R^2_{adj} \in [0.40, 0.50]$, respectively. These two

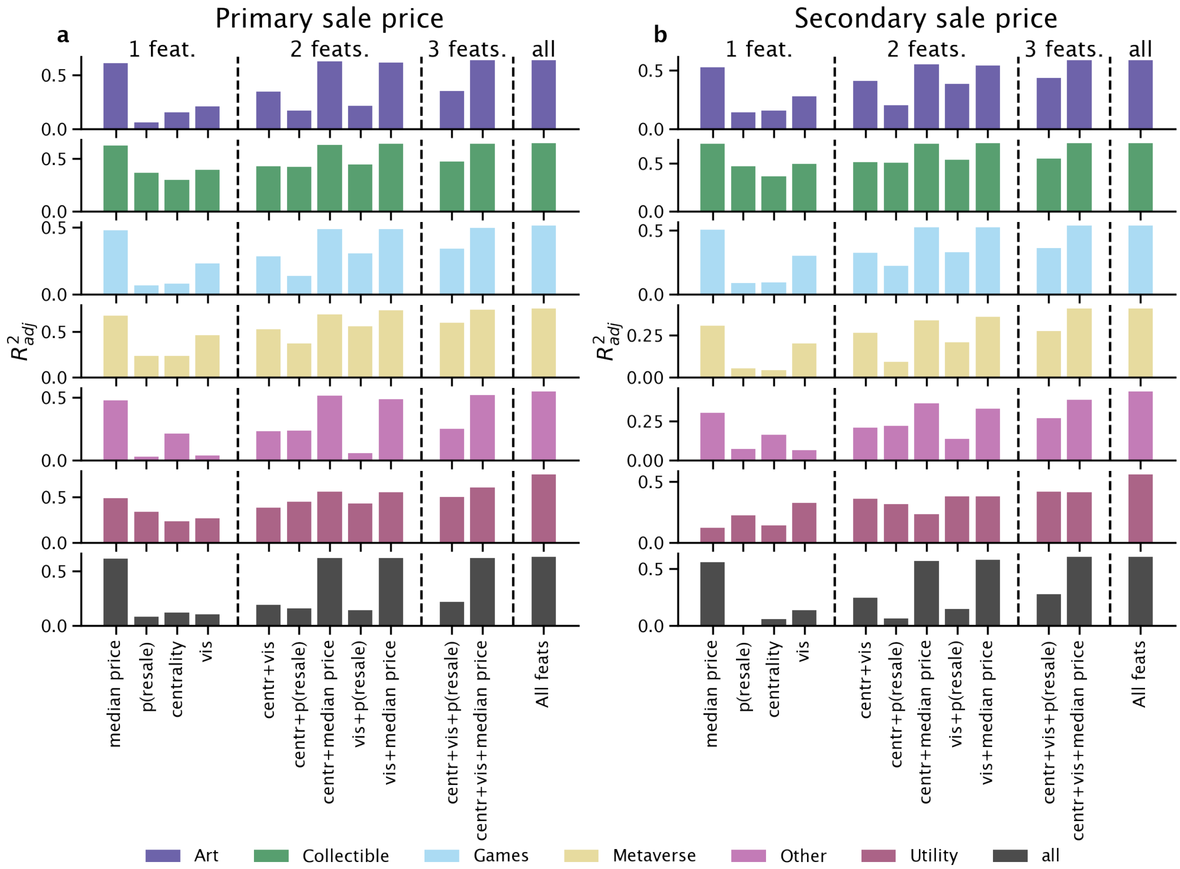


FIG. 9: **Regression results.** R^2_{adj} of a linear regression fit to predict the primary price of sale **(a)** and the median price of secondary sale 1 month after the primary sale **(b)** from different sets of features. Results are broken down by NFT categories.

families of features have the largest compound effect in the *art* category; in the secondary sale price prediction, centrality features boost the predictive power of visual features by more than 50%. Regression coefficients of individual features for the task of secondary sale price prediction one month after the primary sale are presented in Table I.

When predicting secondary sale prices, we consider only those NFTs that were sold in a secondary sale. These NFTs are the minority: less than 10% are sold at least once within one week after the primary sale, and only about 22% within one year (Figure 10). Using the same set of features that we selected for the price regression, we trained a binary classifier to assess to what extent it is possible to predict whether an NFT will be sold after its primary sale. We found that this is possible to a certain extent. The prediction was most accurate when training and testing the classifier on *art* NFTs only ($F1 > 0.8$), whereas the prediction

				β coefficients			
Feature	**All**	**Art**	**Collectible**	**Games**	**Metaverse**	**Utility**	**Other**
const.	-0.029	0.030	-0.086	-0.181	0.210	2.054	0.149
k_{buyer}	-0.018	0.022	-0.032	-0.132	-0.078	-0.010$^\bullet$	-0.207
k_{seller}	-0.166	-0.211	0.000	0.026	0.166	0.198$^\bullet$	-0.347
PR_{buyer}	0.129	0.077	0.162	0.317	0.206	-0.241$^\bullet$	0.336
PR_{seller}	0.302	0.367	-0.031	-0.066	0.009$^\bullet$	-0.382	0.459
p_{resale}	0.029	-0.041	0.079	0.023	0.046$^\bullet$	0.465	0.251$^\bullet$
$medianprice$	0.769	0.711	0.970	0.815	0.436	0.478	0.687
vis_{PCA_1}	0.098	0.153	0.049	0.174	0.175	-1.136	0.021
vis_{PCA_2}	-0.120	-0.130	-0.044	-0.064	-0.669	-0.817	-0.181
vis_{PCA_3}	0.019	0.027	0.063	0.203	0.112$^\bullet$	-1.292	-0.037$^\bullet$
vis_{PCA_4}	0.040	0.028	-0.003$^\bullet$	0.130	-0.018$^\bullet$	-0.911	-0.116
vis_{PCA_5}	0.063	0.018	0.276	0.102	0.296	0.071$^\bullet$	0.301
#NFTs	407,549	251,369	69,015	78,848	2,693	314	5,297
#Collections	3307	114	73	48	12	6	3054
R^2_{adj}	**0.6**	**0.589**	**0.709**	**0.535**	**0.408**	**0.562**	**0.44**

TABLE I: Linear regressions to predict the NFTs' median secondary sale price one month after their primary sale from three families of features: centrality on the trader network (k, PR), history of sales in the NFT's collection (namely prior probability of secondary sale p_{resale} and median sale price 1 week before the sale $medianprice$), and visual features (vis_{PCA_i}). Regression models were fit to different categories of NFTs independently. For each category, the number of NFTs and collections it contains is reported. The R^2_{adj} is a measure of goodness of fit, and it quantifies the proportion of the data variance explained by the model. The p-values of all β coefficients are < 0.01 except for those marked with $^\bullet$.

is less reliable for the other categories ($F1 \in [0.14, 0.33]$). The median price of the collection is among the strongest predictors, but not always the strongest. The prior probability of sale in the collection is also a strong signal, and centrality and visual features combined can sometimes outperform other feature combinations (e.g., in the *metaverse* category). Last,

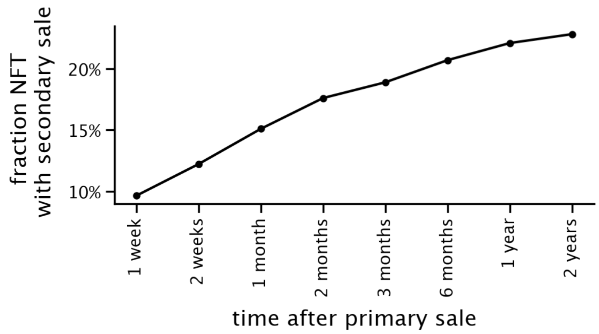


FIG. 10: **Fraction of NFT with secondary sales.** Fraction of NFTs that were sold in at least one secondary sale n days after their primary sale. The vast majority of NFTs got no secondary sale.

the prediction is most accurate when trying to predict the occurrence of a secondary sale over longer periods of time (Figure 12).

III. CONCLUSION

The NFT market is less than four years old and has boomed for just over six months to date. This paper presented the first overview of some key aspects of it by looking at the market history of 6.1 million NFTs across six main categories including art, games and collectibles. In brief, 1) we analyzed the main properties of the market, 2) we built and studied the traders and NFTs networks and found that most traders are highly specialised, 3) we showed that NFT collections tend to be visually homogeneous and 4) we explored the predictability of NFT prices revealing that, while past history is as expected the best predictor, also NFT specific properties, such as the visual features of the associated digital object, help increase predictability.

It is important to highlight the main limitations of our study, which represent also directions for future work. First, we gathered data from a variety of online NFT marketplaces and not directly from the Ethereum blockchain, so that we have likely missed a number of "independent" NFT producers. Second, we adopted a widely accepted categorisation for the NFTs, which could however be further refined and in any case includes a number of arbitrary decisions (as every categorization). Third, since our primary goal was to provide a general overview of the market, we did not extensively explore all the available methods e.g.,

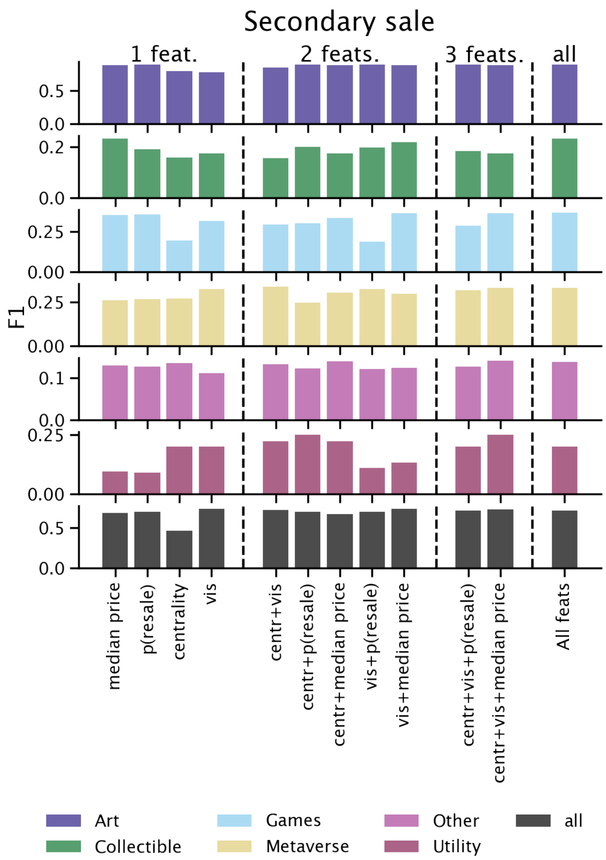


FIG. 11: **Prediction results.** $F1$ score of a binary classification task aimed at predicting whether a NFT will be sold in a secondary sale within 1 year after its primary sale. Results are broken down by different feature sets and NFT categories.

for the clustering of images or price prediction. Fourth, we considered mostly the Ethereum and WAX blockchains, but several other platforms offer smart contracts and therefore likely NFTs. Finally, our price prediction exercise did not include information about the creator of the (digital) object associated to the NFTs. While this is due mainly to the dataset, and in many cases the identity of the creator is not available or does not exist (e.g., for AI generated images), it is clear that in certain contexts, and specifically for art, this can be an important aspect to consider.

Overall, NFTs are a new tool that satisfies some of the needs of creators, users and collectors of a large class of digital and non-digital objects. As such, they are probably here to stay or, at least, they represent a first step towards new tools do deal with property and provenance of such assets. We anticipate that this paper will help accelerate new research

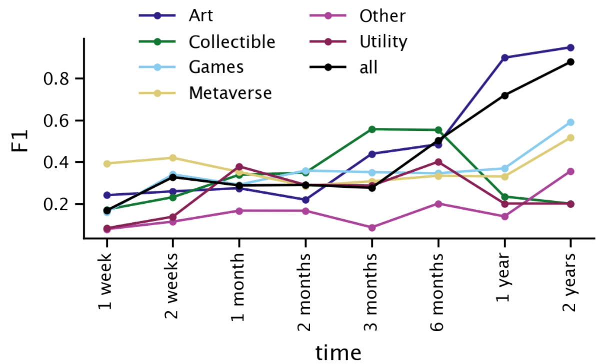


FIG. 12: $F1$ score of a binary classification task aimed at predicting whether a NFT will be sold in a secondary sale within varying time windows after its primary sale. We used all available features for training and testing the models. Results are broken down by different NFT categories.

in a broad array of disciplines, including economics, law, cultural evolution, art history, computational social science and computer science. The results will help practitioners make sense of a rapidly varying landscape and inform the design of more efficient markeplaces as well as the associated regulation.

IV. DATA AND METHODS

A. Data collection

Our dataset includes only transactions representing purchases of NFTs, whose ownership change following that transaction. We exclude from our analysis any transactions representing the minting of NFTs or bids during an auction. We track different cryptocurrencies. Etherum blockchain data for the collections *Superrare*, *Makersplace*, *Knownorigin*, *Cryptopunks*, and *Asyncart* were shared by NonFungible Corporation [22], a company that tracks historical NFT sales data to build NFT valuations. Other Ethereum blockchain data were downloaded from four open-source APIs: CryptoKitties sales [26], Gods-Unchained [27], Decentraland [28], and OpenSea [29]. With OpenSea that allows trading in multiple cryptocurrencies. We also monitored the WAX blockchain, through tracking transactions in the Atomic API [30].

We group NFTs into six categories: *Art* consisting of digital artworks such as images, videos, or GIFs; *Collectible* representing items of interest to collectors; *Games* including digital object used in competitive games; *Metaverse* consisting of pieces of virtual worlds; *Utility* representing items having a specific function; and *Other* including the remaining collections. More details on the NFT categorization are explained in Section IV C. The final, clean dataset includes 935 million USD traded in 6.1 million transactions involving 4.7 million NFTs grouped in 4,624 collections. Our dataset includes transactions in 160 different cryptocurrencies with most of them made in WAX (52% of the total number of transactions), while the volume in USD is mostly ETH (81% of the total volume). Table II shows general statistics of the categories of NFTs considered.

B. Image collection and visual feature extraction

For each NFT in our dataset (except for less than 3,000 exceptions) we managed to collect at least one URL that points to a copy of the NFT's digital object. We focused only on objects with image file formats (e.g. PNG, SVG, JPEG) and GIFs, for a total of about 1.2 million unique graphical objects associated with 4.7 million unique NFTs. Note that a single digital object can be related to the same file; this happens for example for identical playing cards that are minted in multiple copies, each associated with a different NFT. Since our

Category	Buyers	Sellers	NFTs	Volume ($\cdot 10^6$ USD)
Art	161,423	70,623	859,570	655.62
Collectible	62,100	67,173	1,344,449	109.84
Games	151,702	192,772	2,202,432	70.77
Metaverse	12,121	10,283	47,286	68.18
Utility	2,637	1,483	7,752	8.74
Other	34,647	22,308	242,990	21.96
Total	359,561	314,439	4,704,479	935.11

TABLE II: **NFTs categories.** Overall statistics of each NFT category under consideration.

algorithm for visual feature extraction works with static images, we converted the animated GIFs to PNGs by extracting central frame of each GIF. In order to succinctly represent the visual features that characterize an image, we encode it into a latent space using a neural network. Specifically, we pick the PyTorch [31] implementation of AlexNet [32], a deep convolutional neural network architecture designed for image classification. We initialize AlexNet with weights pre-trained on ImageNet [33], a widely-used reference dataset of labeled images. Given an image in input, AlexNet passes it through multiple layers of transformation. The second to last layer (i.e., the layer before the classification layer) is a vector consisting of 4,096 values that constitute a dense representation of the input image into a high-dimensional space. These vectors can be used for a variety of tasks such as similarity ranking, clustering, or classification. To reduce the dimensionality of AlexNet vectors, we extracted their principal components using Principal Component Analysis (PCA) [34], and selected the 5 most relevant ones. PCA projects each point of the high-dimensional space into a space with a desired number of dimensions, while preserving the data variation as much as possible.

C. Data cleaning and categorization

NFTs that share common features are grouped in *collections*, which names are cleaned and even out. The raw names, as downloaded from the selected sources, are stripped by any digits, special characters (e.g., "-"), unusual patters (e.g., "xxxxx"), and capitalized. Cleaned names are then even out by considering a list of words. For instance, the collection "Aavegotchi" renames all collections starting with that string of characters in "Aavegotchi". Some other collections with generic names (e.g. "Stuff") are called "Miscellanea".

Fields considered in our analysis are: buyer address, seller address, time of the transaction, name of the collection, ID of an NFT (here simply called "NFT"), url to the NFT's digital object, type of cryptocurrency and its amount used in the transaction. Transactions with one of the former fields empty (except for the url to the NFT's digital object) are removed from the dataset. From these remaining data, the price in USD is computed considering the exchange rate of the given cryptocurrency at the day of the transaction. Note that, in this work, we use buyer or seller addresses as proxies for real identities, as commonly done in the Ethereum blockchain [35] and with the usernames [20], while in reality an individual may have multiple addresses or usernames. NFTs sharing common features, such as, digital cards of the same online game, belong to the same collection. In turn, collections are assigned to one of the following six categories: "Art", "Collectible", "Games", "Metaverse", "Utility", or "Other". The operative definitions of these categories are inspired from the definitions given by NonFungible Corporation and summarized in Table III. We automatically categorize collections already categorized by NonFungible Corporation, we also automatically classified as "Art" all collections containing that string of characters in their name, manually removing false positives. At least two authors of the present manuscript manually categorized other collections with high trading volume or large number of sales. The manual categorization was done independently by each author, then the final category selected by majority voting, calling additional authors in case of draw between two or more categories. Note that a collection may belong to more than one category and forcing each collection into one category only is a limitation of the present work.

With the exception of the Atomic API, the downloaded datasets are not independent and, for instance, some transactions shared by NonFungible Corporation are available from OpenSea as well. When data are merged together, duplicated transactions are removed by prioritizing (in order) data from NonFungible Corporation, CryptoKitties sales, Gods-Unchained API, Decentraland API, and OpenSea API.

Category	Description
Art	NFTs of digital artworks, such as images, videos, or gifs
Collectible	NFTs of interest to a collector
Games	NFTs used n competitive games
Utility	NFTs for specific purposes (e.g. secure and decentralized name service)
Metaverse	Piece of virtual worlds
Other	NFTs of small collections that are not included in the other categories

TABLE III: Operative definitions of NFTs categories.

D. Generation of the traders and NFTs networks of interaction

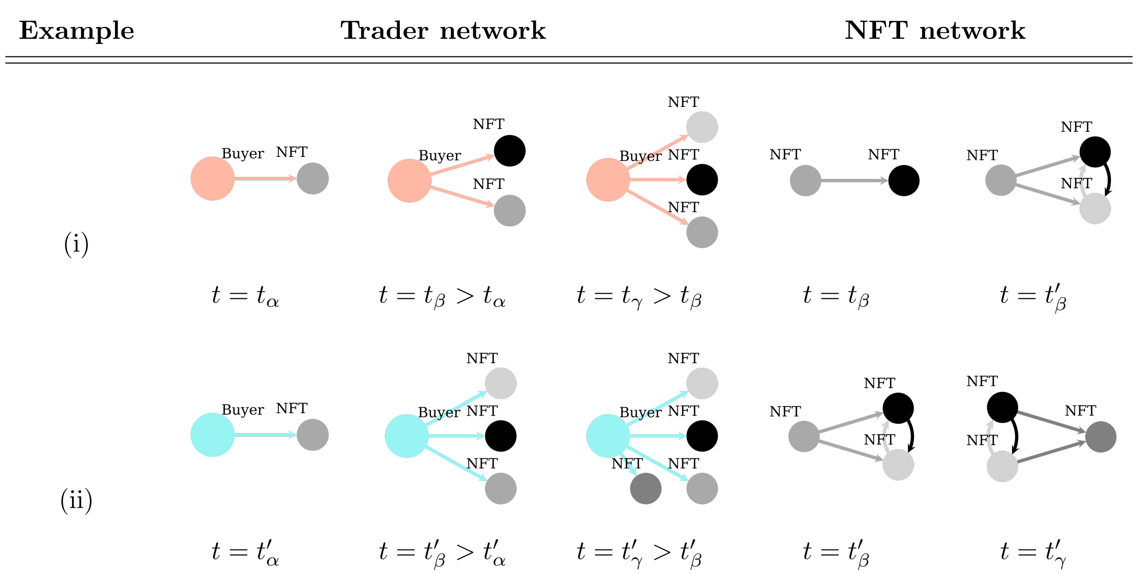


TABLE IV: **Link creation mechanism of the NFT network.** Directed links are generated using the trader network as reference and following three rules. The first two rules take into consideration the same buyer, while the third rule another buyer, both interacting with the same three NFTs. Visualization is done using Graph-tool: [25].

While the trader network was directly obtained from our data collection, the NFT network was created by linking NFTs that are purchased in a sequential order by the same buyer.

Let's consider NFT_i, NFT_j, NFT_k, and NFT_h as identifier of generic NFTs, and t_α, t_β, t_γ, t'_α, t'_β, and t'_γ as identifiers for time instants (with a temporal resolution of seconds). Table IV illustrates two meaningful examples of how the NFT network is created. (i) When a buyer, who purchased NFT_i at time t_α, buy NFT_j at time $t_\beta > t_\alpha$, a directed link from NFT_i to NFT_j is created at time t_β. If the same buyer purchases NFT_k at a later time $t_\gamma > t_\beta$, a directed link from NFT_j to NFT_k is drawn at time t_γ. (ii) When a buyer, who purchased NFT_i at time t'_α buy NFT_j and NFT_k at the same instant $t'_\beta > t'_\alpha$, a directed link from NFT_i to NFT_j and another from NFT_i to NFT_k are drawn. If the same buyer purchases a fourth NFT_h at time t'_α. The NFT network hereby constructed includes 4,657,713 NFTs out of a total of 4,704,479. The NFTs that are left out belongs to buyers who perform only one transaction. The network analysis is done by leveraging selected functions in the networkx Python package.

E. NFT features

We characterize NFTs with a set of 11 features, partitioned in three groups. An NFT's features were calculated only from the data that could be collected until the day before its primary sale, t_s. We used these features in two separate tasks of regression (Section IV F), and classification (Section IV G).

The first group of features includes network centrality scores obtained from the trader network. Specifically, we considered the degree centrality (k), and the PageRank centrality (PR) of the seller and the buyer, for a total of 4 features. The degree centrality of a node is the count of all its incoming and outgoing unique links [36], and its PageRank centrality measures the stationary probability that a random walk on the network ends up in that node [37].

The second group includes the visual features of the object associated with the NFT, namely 5 PCA components extracted from the AlexNet vector of the object ($PCA_{1\ldots5}$). We experimented with a number of components varying from 2 to 10, and results varied only slightly—fewer components caused a feeble decrease in the quality of the regression and prediction results, while additional components did not add any predictive power.

The third and last group includes two features to account for the previous sale history in the NFT's collection. The first is the median price of primary and secondary sales made in the collection of interest during a time window prior to t_s. The latter models the prior probability of secondary sale. We acknowledge that the likelihood that a NFT gets transacted in a secondary sale might depend on the collection it belongs to. For example, NTFs corresponding to collectible items from very popular collections may be more likely to be resold than an NFT serving for a specific purpose, such as determining the ownership of a name server. We defined the probability of secondary sale, p_{resale}, as 0.5 (random probability) when the NFT is the first to be sold in its collection; else, the probability of secondary sale is calculated as:

$$p_{resale} = \frac{0.5}{n+1} + \frac{n}{n+1}\frac{s}{n},$$

where n represents the NFTs with a primary sale up to the day before the first purchase and s the number of these NFTs with at least one secondary sale. When the collection is large, the probability of secondary sales becomes $p\,(n \to +\infty) = s/n$ and corresponds to the ratio between items with secondary sales over all items with one sale.

The frequency distributions of our features have different skews and ranges. To make them comparable and suitable for regression and prediction tasks, we first transform their values to make their distributions closer to a Normal distribution. Specifically, we calculate the logarithm of the network degree and the median sale price (after adding 1, so that zero-values were preserved), and we apply a BoxCox transformation [38] to the PageRank centrality and to p_{resale}; BoxCox uses power functions to create a monotonic transformation that stabilizes variance and makes the data closer to a normal distribution. No transformation was needed for the PCA features. Last, we scale all the variables in the range $[0, 1]$ (i.e., min-max scaling).

F. Sale price regression

We perform linear regressions to estimate an NFT's primary and secondary sale prices. Linear regression is an approach for modeling a linear relationship between a dependent variable (secondary sale price, in our experiments) and a set of independent variables (features describing the NFT at the time it was first sold), and it does so by associating a so-called β-coefficient with each independent variable such as the sum of all independent variables multiplied by their respective β-coefficients approximates the value of the dependent variable with minimal error. Specifically, we used an Ordinary Least Squares regression model to estimate coefficients such that the sum of the squared residuals between the estimation and the actual value is minimized.

We use the NFT features described in Section IV E as independent variables, and either the price of primary sale or the median secondary sale price calculated over a time window starting at t_s as dependent variables. For resale price, the results changed only slightly when using different aggregations other than the median (e.g., mean, maximum). We experimented with different lengths of the time window, ranging from one week after the primary sale up to two years after. To make sure that the secondary sale price of each NFT was calculated over time windows of equal length, we exlcuded from the regression NFTs that were sold for the first time too recently—namely those NFTs whose t_s was within one time window before the most recent timestamp in our dataset. In the regression, we considered only NFTs with at least one secondary sale in the time window considered.

We evaluated the goodness of the linear fit using coefficient of determination R^2, a score in the range $[0, 1]$ that measures the proportion of the variance in the dependent variable that the linear model is able to predict from the independent variables. In particular, we used its 'adjusted' version R^2_{adj}, that discounts the effect of the R^2 spuriously increasing as more independent variables are added to the model.

G. Secondary sale prediction

We performed a binary classification task to predict whether an NFT will be transacted in a secondary sale after its primary sale at time t_s. We adopted a standard supervised learning approach. In supervised learning, instances in a dataset (the NFTs) are described with a number of features (those presented in Section IV E) and marked with a target label (1 if the NFT was transacted in a secondary sale, 0 otherwise). A mathematical model learns a function that maps the features to the target label based on a number of *training* instances from the dataset. The performance of the model is later assessed on a *test* set of unseen instances. In our experiments, we emulate a prediction on future data based on past knowledge. To do so, we sort the NFTs according to their time of primary sale t_s, and we use the first 95% of NFTs for training and the latest 5% for testing. Our dataset is sufficiently large so that the test set, albeit small in relative terms, includes a large selection of tens of thousands of instances. Similar to the regression task, we consider multiple time windows of varying size to determine the target label (i.e., whether the NFT was resold or not), and we exclude from the dataset recent NFTs whose t_s is within one time window before the last timestamp in our dataset.

There are several classes of models that can be used for supervised learning [39]. We pick AdaBoost [40], an ensemble of weak learners (in our case, decision trees) whose output is combined into single score through a weighted sum. Despite its relatively simple design, AdaBoost can achieve good performance compared to more complex model and it effectively limits overfitting the learned function on the training data.

The labels of our dataset are *imbalanced*: the number of negative labels is much higher than the number of positive ones (i.e., 80% of NFTs in our dataset are more not resold). Imbalanced datasets can affect the ability of the model to learn a function that can effectively associate the correct label to both positive and negative instances. To mitigate this problem, we perform random oversampling [41] to balance the classes. Specifically, within the training set, we add multiple copies of positive samples picked at random until the size of the two classes is balanced. Compared to other oversampling techniques [42, 43], random oversampling does not generate synthetic data points, which exhibiting unrealistic features. By applying oversampling, we effectively set the model to assign higher importance to positive samples: misclassifying a positive instance causes a loss in performance that is proportional to the number of its replicas.

To evaluate the performance on the test set, we measure two quantities. The first is the *F1-score*, namely the harmonic mean of the precision (fraction of instances that are classified as positive that are indeed positive) and recall (fraction of positive instances the are correctly classified). The second is the "Area Under the ROC Curve" (AUC); it measures the ability of the model to correctly rank positive and negative samples by confidence score, independent of any fixed decision threshold. AUC is equal to 0.5 for a random classification and it is equal to 1 for a perfect ranking.

ACKNOWLEDGEMENTS

The authors are grateful to Nonfungible Corporation for helpful conversations and data sharing (see text). The research was partly supported by The Alan Turing Institute.

DATA AVAILABILITY

Data are available upon request.

[1] Anny Shaw Ben Luke and Helen Stoilas. WTF are NFTs? Why crypto is dominating the art market. https://www.theartnewspaper.com/podcast/wtf-nfts Accessed May 4, 2021, 2021. The Art Newspaper.

[2] NonFungible Team. Non-fungible tokens quarterly report Q1 2021. https://nonfungible.com/subscribe/nft-report-q1-2021 Accessed May 4, 2021, 2021. NonFungible Corporation.

[3] Tonya M Evans. Cryptokitties, cryptography, and copyright. *AIPLA QJ*, 47:219–247, 2019.

[4] The Writer's Lounge. Choosing the right blockchain for your NFT. https://medium.com/phantasticphantasma/choosing-the-right-blockchain-for-your-nft-d1df2bebae91 Accessed May 4, 2021, 2020. Medium.

[5] Cryptokitties Team. CryptoKitties: Collect and breed furrever friends. https://www.cryptokitties.co/ Accessed May 4, 2021, 2021. Cryptokitties.

[6] Joon Ian Wong. The Ethereum network is getting jammed up because people are rushing to buy cartoon cats on its blockchain. https://qz.com/1145833/cryptokitties-is-causing-ethereum-network-congestion/ Accessed May 4, 2021, 2017. Quartz.

[7] Fitz Tepper. People have spent over $1m buying virtual cats on the Ethereum blockchain. https://techcrunch.com/2017/12/03/people-have-spent-over-1m-buying-virtual-cats-on-the-ethereum-blockchain/ Accessed May 4, 2021, 2017. TechCrunch.

[8] Rebecca Riegelhaupt. Results: Beeple's purely digital NFT-based work of art achieves $69.3 million at Christie's. https://www.christies.com/about-us/press-archive/details?PressReleaseID=9970&lid=1 Accessed May 4, 2021, 2021. Christie's Press Release.

[9] Scott Reyburn. JPG file sells for $69 million, as "NFT mania" gathers pace. https://www.nytimes.com/2021/03/11/arts/design/nft-auction-christies-beeple.html Accessed May 4, 2021, 2021. The New York Times.

[10] Daniel Phillips. The 10 most expensive NFTs ever sold. https://decrypt.co/62898/the-10-most-expensive-nfts-ever-sold Accessed May 20, 2021, 2021. Decrypt.

[11] JohnJo Devlin. The "insane" money in trading collectible cards. https://www.bbc.co.uk/news/business-56413186 Accessed May 20, 2021, 2021. BBC.

[12] Qin Wang, Rujia Li, Qi Wang, and Shiping Chen. Non-fungible token (NFT): Overview, evaluation, opportunities and challenges. *arXiv preprint arXiv:2105.07447*, 2021.

[13] Martin Westerkamp, Friedhelm Victor, and Axel Küpper. Blockchain-based supply chain traceability: Token recipes model manufacturing processes. In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 1595–1602. IEEE, 2018.

[14] Amy Whitaker. Art and blockchain: A primer, history, and taxonomy of blockchain use cases in the arts. *Artivate*, 8(2):21–46, 2019.

[15] Lauren van Haaften-Schick and Amy Whitaker. From the artist's contract to the blockchain ledger: New forms of artists' funding using equity and resale royalties. *Available at SSRN 3842210*, 2021.

[16] Alesja Serada, Tanja Sihvonen, and J Tuomas Harviainen. CryptoKitties and the new ludic economy: How blockchain introduces value, ownership, and scarcity in digital gaming. *Games and Culture*, page 1555412019898305, 2020.

[17] Kentaro Sako, Shin'ichiro Matsuo, and Sachin Meier. Fairness in ERC token markets: A case study of CryptoKitties. *arXiv preprint arXiv:2102.03721*, 2021.

[18] Michael Dowling. Fertile land: Pricing non-fungible tokens. *Finance Research Letters*, page 102096, 2021.

[19] Massimo Franceschet. Art for space. *Journal on Computing and Cultural Heritage (JOCCH)*, 13(3):1–9, 2020.

[20] Albert-László Barabási. The art market often works in secret. Here's a look inside. https://www.nytimes.com/2021/05/07/opinion/nft-art-market.html Accessed May 17, 2021, 2021. The New York Times.

[21] Michael Dowling. Is non-fungible token pricing driven by cryptocurrencies? *Finance Research Letters*, page 102097, 2021.

[22] NonFungible Team. The best place to analyze, track, and discover NFTs. https://nonfungible.com/ Accessed May 4, 2021, 2021. NonFungible Corporation.

[23] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[24] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. Finding community structure in very large networks. *Physical Review E*, 70(6):066111, 2004.

[25] Tiago P. Peixoto. The graph-tool python library. *Figshare*, 2014.

[26] Niel de la Rouviere. A subgraph to index & explore CryptoKitties auctions. https://thegraph.com/explorer/subgraph/nieldlr/cryptokitties-sales Accessed May 4, 2021, 2021. The Graph.

[27] DJ Rosenbaum. Gods-Unchained marketplace. https://github.com/djrosenbaum/unchained-transactions Accessed May 4, 2021, 2021. The Graph.

[28] Decentraland Team. Decentraland marketplace. https://thegraph.com/explorer/subgraph/decentraland/marketplace Accessed May 4, 2021, 2021. The Graph.

[29] OpenSea Team. Api overview. https://docs.opensea.io/reference Accessed May 4, 2021, 2021. OpenSea.

[30] Atomic Market Team. Atomic market api. https://wax.api.atomicassets.io/atomicmarket/docs/swagger/ Accessed May 4, 2021, 2021.

[31] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*, 2019.

[32] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25:1097–1105, 2012.

[33] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. Ieee, 2009.

[34] Ian T Jolliffe. Principal components in regression analysis. In *Principal component analysis*, pages 129–155. Springer, 1986.

[35] Larva Labs Team. All CryptoPunk owners. https://www.larvalabs.com/cryptopunks/leaderboard Accessed May 17, 2021, 2021. The New York Times.

[36] Stanley Wasserman, Katherine Faust, et al. *Social network analysis: Methods and applications.* Cambridge University Press, 1994.

[37] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107–117, 1998.

[38] George EP Box and David R Cox. An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2):211–243, 1964.

[39] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: Data mining, inference, and prediction.* Springer Science & Business Media, 2009.

[40] Yoav Freund, Robert Schapire, and Naoki Abe. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780):1612, 1999.

[41] Charles X Ling and Chenghui Li. Data mining for direct marketing: Problems and solutions. In *Kdd*, volume 98, pages 73–79, 1998.

[42] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.

[43] Haibo He, Yang Bai, Edwardo A Garcia, and Shutao Li. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 1322–1328. IEEE, 2008.

Supplementary Information

S1. ADDITIONAL ANALYSES

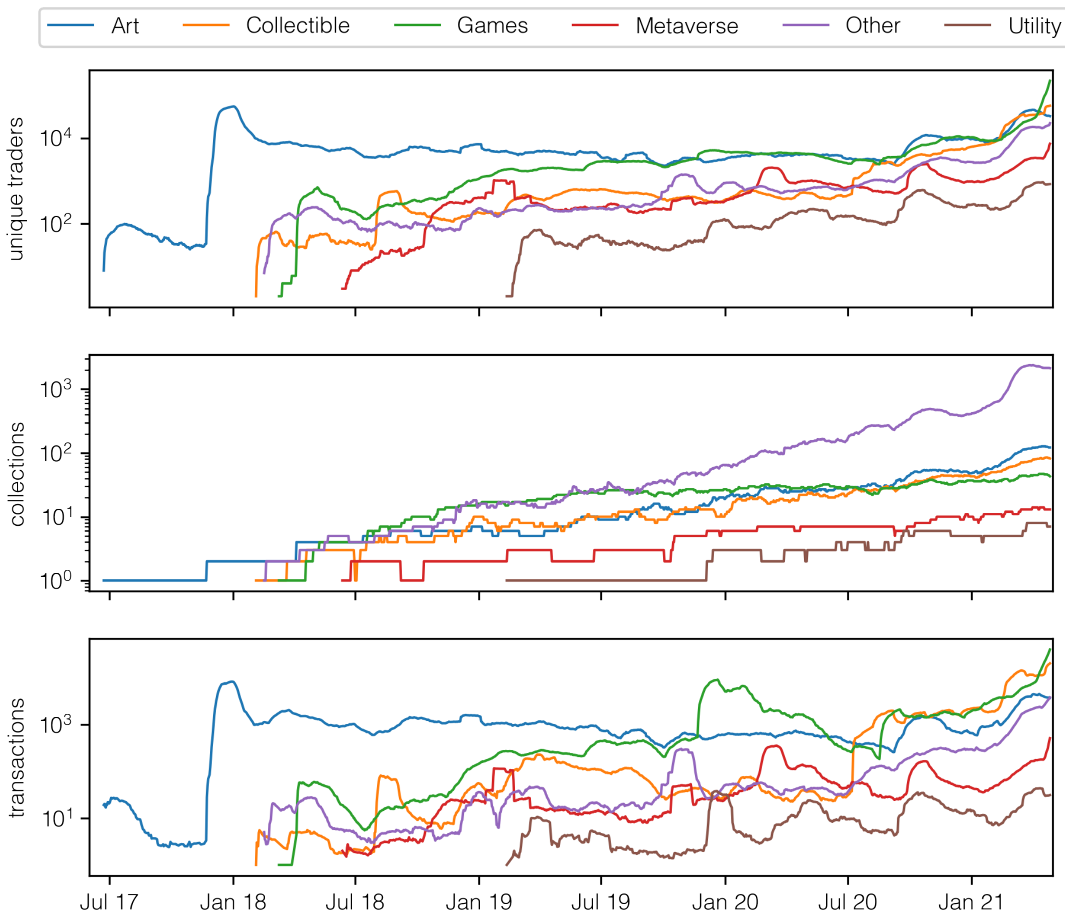


FIG. S1: **Evolution of the NFT Market** The number of unique traders (top), collections (middle) and transactions (bottom) over time for different categories (see legend). Results are computed over a rolling window of 30 days.

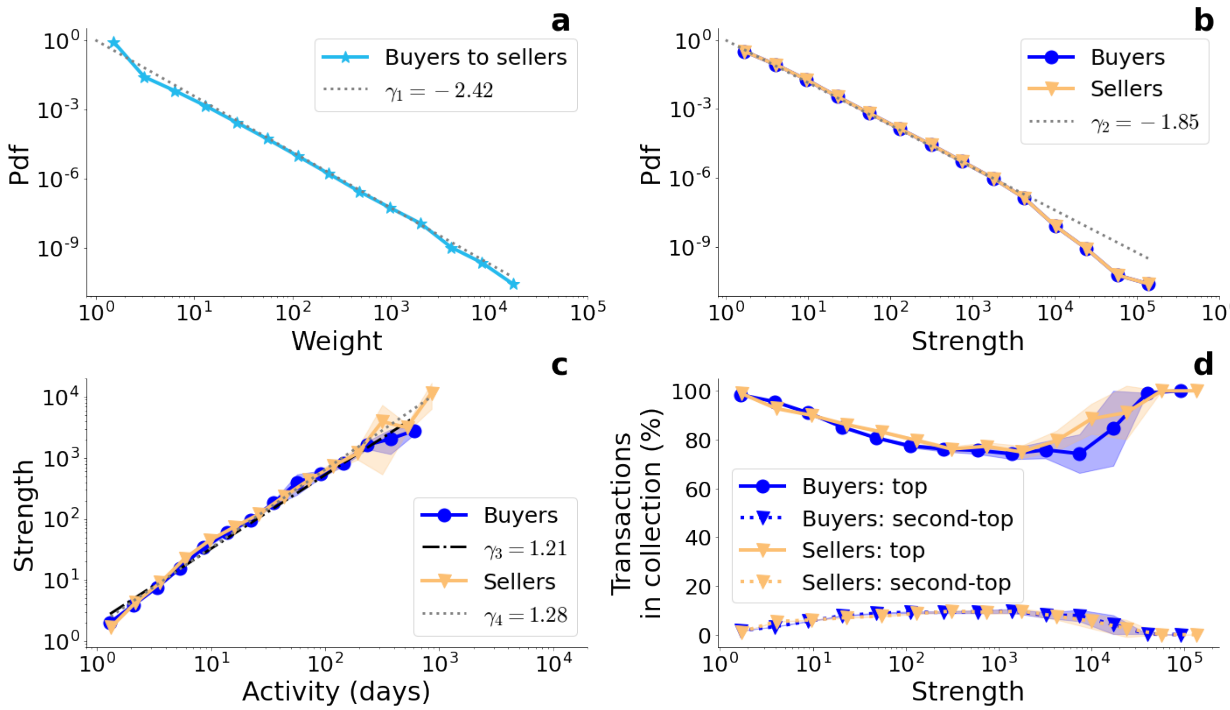


FIG. S2: **Key network properties of buyers and sellers.** (a) Probability distribution function of the number of transactions (weight) from buyers to sellers. (b) Probability distribution function of the buyers and sellers' strength. (c) Relationship between the buyers and sellers' strength and the number of days in which they are active. (d) Percentage of transaction buyers and sellers make toward their top and second-top NFT collections.