

# An artist ranking system based on social media mining

Amalia F. Foka<sup>1</sup>

Received: 7 January 2017 / Accepted: 14 January 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** Currently users on social media post their opinion and feelings about almost everything. This online behavior has led to numerous applications where social media data are used to measure public opinion in a similar way as a poll or a survey. In this paper, we will present an application of social media mining for the art market. To the best of our knowledge, this will be the first attempt to mine social media to extract quantitative and qualitative data for the art market. Although there are previous works on analyzing and predicting other markets, these methodologies cannot be applied directly to the art market. In our proposed methodology, artists will be treated as brands. That is, we will mine Twitter posts that mention specific artists' names and attempt to rank artists in a similar manner as brand equity and awareness would be measured. The particularities of the art market are considered mainly in the construction of a topic-specific user network where user expertise and influence is evaluated and later used to rank artists. The proposed ranking system is evaluated against two other available systems to identify the advantages it can offer.

**Keywords** Social media mining · Art market · Social media user network · User expertise · User influence · Topic identification

## 1 Introduction

Social media users are currently estimated to approximately 2 billion (Statista 2016). The content of the posts these users make has been used extensively in applications to quantify the digital reputation economy. Users on social media post their personal opinion and feelings about almost everything. On the other hand, almost every business firm from any market sector is also present on the social media mainly for the purposes of marketing and communicating with their customer base.

---

✉ Amalia F. Foka  
afoka@cc.uoi.gr

<sup>1</sup> Department of Fine Arts & of the Sciences of Art, University of Ioannina, 45110 Ioannina, Greece

This online behaviour has led to numerous applications where social media data are used to measure public opinion in a similar way as a poll or a survey. Applications range from prediction of an election winner or product sales (Jin et al. 2010), market analysis and industry competition (He et al. 2013), prediction of the stock market (Bollen et al. 2011), identification of adverse drug reactions (Yang et al. 2012), measurement of brand awareness (Peetz 2015) and brand equity (Culotta and Cutler 2016), to name but a few.

In this paper, we will present an application of social media mining for the art market. To the best of our knowledge, this will be the first attempt to mine social media to extract quantitative and qualitative data for the art market. Although there are previous works on analyzing and predicting other markets, these methodologies cannot be applied directly to the art market. As it is stated in Abbing (2002), the economy of the arts is extraordinary. The foremost difficulty with the art market is the available data. The sales data from the secondary market (i.e. auctions), which some firms collect, are not freely available. However, when we consider the primary market (i.e. sales from galleries, dealers, etc.) there is no data available at all. Another difficulty is that even if there were available price data, we could not make estimates based only on the market value of artworks. Artworks as products incorporate cultural values that are hard to quantify and estimate. Finally, the art world is elitist and certain social groups with higher cultural status or buying power have a greater role in structuring the cultural value and in effect the market value of an artwork. Consequently, mining the general public opinion cannot be exploited to make estimates for the art market.

In our proposed methodology, we will treat artists as brands. That is, we will mine Twitter posts that mention specific artists' names and we will attempt to rank artists based on the inherent cultural value of their works. To achieve this we will work in a similar manner as if we were interested in measuring a brand's awareness and equity but taking into account the particularities of the art market. We will first show how we can build a network of users that are influential. These users will be assessed on their expertise and the extent in which they participate in the social construction of cultural value. User evaluation will then be utilized to weigh the importance of mined tweets to estimate each artist's ranking. Obtained ranking results will be evaluated as trends rather than per se facts for their accuracy since there are no data to compare them with. Estimated trends will be correlated with other available artist rankings to demonstrate the advantages of our methodology.

The remainder of this paper is organized as follows. Section 2 is a brief review of how the art market is organized and establishes the necessity of the proposed application. Section 3 presents a literature review of social media data mining applications for other economies. Sections 4 and 5 detail our methodological approach as well as some key findings. Section 6 presents the final artist ranking results and discusses them in depth. Section 7 concludes with suggestions for future research.

## 2 The art market

The economy of the arts might be considered exceptional for some scholars (Abbing 2002) but for those who are not inside the art world (Danto 1964) it might seem as just odd.

The art world in Becker (1982) is regarded as a network of people: suppliers of materials, distributors of artworks, fellow artists, critics, theorists, and audiences. Works of art are produced within this network of people. In Thornton (2008) the art world is seen as a *“loose network of overlapping subcultures held together by a belief in art”*. The most

prominent actors behind the art scenes encompass roles as artists, dealers, curators, critics, collectors, and auction-house experts. In general, although at different periods of time the notion of who constitutes the art world might vary what remains as a constant is that not everybody can participate in the art world. The latter is mostly comprised of certain social groups, elites, either due to their higher cultural status or their higher purchasing power.

The elitist nature of the art world and/or the cultural superiority of actors in it, is actually what makes the economy of arts exceptional mainly in the sense of how prices are formed. In most economies prices are determined based on the supply and demand. Not surprisingly, this is not the case in the art economy. Certainly, there are also other economies in which prices are not determined based on the supply and demand. For example, in the economy of haute couture the price of a clothing item incorporates a must have value mainly due to the high symbolic value directly relevant to the commercial success and fame of the designer. Haute couture is accessible only to a certain elite group due to its purchasing power. The main distinctive factor of the art economy, when compared to the haute couture economy or other luxury product economies, is that an artwork as a product conveys for its owner cultural values as well as other values. These conveyed cultural values make artworks virtually priceless and can lead to astronomical prices.

The study of the economy of the arts started gaining interest during the 90s when prices reached excessive levels. A significant number of studies have been published since then (Grampp 1989; Velthuis 2007; Thompson 2010; Thornton 2008; Graw 2010; Boll 2011; Degen 2013), which attempt to explore the main factors that contribute to the formation of artwork prices and are summarized below :

<i>aesthetic value</i>	the value of an artwork when appreciated or experienced
<i>social value</i>	the gained social status by the ownership of an artwork
<i>investment value</i>	the safety gained from investing in an artwork since prices never drop.

Wealthy people try to find ways to invest their money and art could be a great investment since the usual practice for living artists is that prices never drop. The statement that prices in the art world never drop might initially sound unreasonable or even illegal, as it would be in most other economies. However, in the art world there are two kinds of markets:

<i>the primary market</i>	where artworks are sold for the first time either directly by the artists or by their representatives (e.g. a gallery) and
<i>the secondary market</i>	where artworks are resold by auction houses.

In the primary market, gallery owners have total control over the price of an artwork. The prices are set according to the gallery's as well the artist's prestige, status and fame. Even in the case where the gallery owner overprices some works of art and they don't sell, prices will not drop. And this is because if prices did drop this would have a negative effect in the art world's perception about the credibility and quality of the gallery as well as the artist. As explained in Abbing (2002), it is the persistent beliefs about art and artists that ultimately govern the art economy. In later sections of this paper, the use of social media mining will be justified as a means for mapping such persistent beliefs about art.

In the case of the secondary market, auction-house experts research previous auction data and retail gallery sales to determine a minimum presale estimate price, a reserve price and the opening bid amount. These data determine whether a certain artwork will be finally

auctioned. If it is not almost certain that the artwork will be sold at a desired price it will not be auctioned. In almost all cases, artwork prices raise excessively after they enter the secondary market. In conclusion, the event of an artwork's price to drop is highly unlikely or even impossible in the way the art economy is structured.

Another significant factor that determines art prices is art's integrated social value. Art has a high symbolic content that is used by art buyers to mark their social status. It is not surprising that works of art can be found primarily among wealthy people and institutions. In a way, this is one of the causes for the recent phenomenon of star artists. The social status gained by the acquisition of an artwork is directly proportional to how recognizable the artist's name is, not necessarily by the general public but at least by the elite of the art world. For this reason, we will later propose to treat artist's names as brands while mining social data with the intent to estimate social values.

The discussion in the art world about how the aesthetic value of an artwork is determined is long and beyond the scope of this paper. In general, today it is concluded that the aesthetic value and market value are interdependent. In Velthuis (2007), it is maintained that in general higher prices signal to consumers a higher quality (i.e. aesthetic value), as it is often the case in other economies as well. Furthermore, today it is accepted that experts with high cultural power are the ones who establish the aesthetic value. Moreover, in Abbing (2002) it is stated that aesthetic value is in practice a social construction that occurs between experts and is related to reputations. Again, we will later show how social data can be utilized to identify the social construction of the aesthetic value based on the notion of reputations.

To conclude our discussion about the art market, we have identified its basic traits and some of the key factors that determine art prices. It is evident, that the actors in the art world would have the necessity to have access to art economy data either for pure research purposes or to decide what to buy. Following, we will present the major available art market information services and later it will be discussed what additional information our proposed methodology can provide.

## 2.1 Art market information services

The global sales of art in 2015 were estimated to \$63.8 billion (Kinsella 2016). Auction sales were estimated to the 47% of the total global art sales and the rest 53% to private sales (galleries, dealers, etc.). Furthermore, the total revenue of global sales of art was generated by 38.1 million transactions. The size of the art market dictates involved actors to entail services that provide data and analysis before making a decision for a transaction. Following, we will describe some of the most popular art market information services.

Artprice (2016) gives access to its subscribers to auction sales data since 1962. They also provide the Artprice indicator™ that estimates the potential current value of an artwork given the amount spent to purchase it at some time in the past. The Art Price Global Index measures the total worldwide turnover of auction sales. Finally, Artprice releases an annual report for the art market, available also to non-subscribers, where statistics for global auction sales are presented as well as a ranking with the top 500 artists by auction revenue.

The European Fine Art Foundation (TEFAF) also publishes an annual report with statistics about the global art market. The TEFAF report does not focus only on auction sales as it is the case with the Artprice annual report. It includes data from auctioneers, dealers, collectors, industry observers and art sales databases. Consequently, it includes data from

the primary market as well. However, we should note that sales data for the primary market are collected with polls to worldwide dealers. There is no available database for the primary art market sales and although respondents to the poll represent only a fraction of the worldwide dealers, the obtained results provide useful insights.

The artnet (2016) is another subscription based service which includes a price database of auction results since 1985. Additionally, they offer an Analytics Reports service where users can create customized reports to get insights on artists' auction sales performance and artwork price rankings.

Other similar services that provide price databases for auction results are the Blouin Art Sales Index (BlouinArtinfo 2016), the ArtNexus (2016) and the FindArtInfo (2016).

What is obvious from the above list of art market information services is that almost all of them are dedicated to auctions. However, as it has already been mentioned, the secondary market represents approximately only half of the total art market. It is obvious that there is a lack of price data for the primary market. This is expected due to its loose structure and it would be impossible to envisage a price database for sales in galleries, art fairs, etc.

Following, we will discuss in more detail Artfacts.Net, an online service that attempts to tackle the problem of the lack of price data in the primary market with the utilization of the economy of attention. That is, they measure exhibition success instead of sales success.

### 2.1.1 *Artfacts.Net*

Artfacts.Net (Artfacts 2016c) is an online subscription service that started in 2001 and offers a database, where someone can find information on exhibitions, artists, works of art, etc. as well as an artist ranking system.

What differentiates Artfacts.Net from the previously mentioned services is that it is not focused on auctions only. The artist ranking system evaluates artists based on their exhibition success, namely on the economy of attention. The museum's or gallery's fame and success is in a way transferred to the artist.

Artfacts.Net has in its database 550,406 artist biographies and 33,954 art institutions worldwide. At this time, it has listed 2737 current or upcoming exhibitions worldwide. The database is updated manually and practically it cannot include everything in the art world globally. This is why users have the option to report missing information on artists, exhibitions, galleries and museums. At present, the Artfacts.Net web page states that the last database update was over three months ago.

Artfacts.Net ranks almost 100,000 artists (Artfacts 2016b) according to a point system. For example, Frank Stella holds the 109th position with 11,530.82 points. Unfortunately, detailed information on how points are awarded is not given. The only information available is that points indicate the amount of attention each artist has received from art institutions.

Marek Claassen, the director of Artfacts.net, illustrates that awarded points attempt to quantify how much embedded an artist is in the international art world (Periferic Biennial 2008). The number of countries, collections and galleries an artist's work is exhibited is counted, along with the number of solo and group shows. Artists are awarded points equal to the exhibition value of the gallery or museum that hosted them. The exhibition value raises according to the number of international artist exhibitions the gallery or museum has hosted.

The Analysis page for Frank Stella (Artfacts 2016a), that is accessible as a trial to the subscription service, reveals more information on awarded points. He received 1,360.77

points for 25 exhibitions in 2016. Again, there are no details on the points for each exhibition but only the average, minimal, medial and maximal points of exhibitions. Subscribers can only get insight on whether an artist tends over the years to participate in more exhibitions or not and whether the average exhibition value increases or decreases. Furthermore, no details can be found for the remaining points up to the total 11,530.82 Frank Stella received in 2016.

Analysis pages also provide statistics and graphs that illustrate an artist's exhibition history by: (1) the different types of art institutions (e.g. festivals, galleries, museums, etc.), (2) the type of exhibitions (i.e. solo or group), (3) ranking over years and finally (4) a peer group analysis. The peer group represents artists who usually exhibit collectively in group shows. Its analysis correlates the ranking evolution for artists in the same group.

In conclusion, Artfacts.Net might have some limitations but currently it is the only service that provides an artist ranking based on data different than auction sales. Subsequently, it is the only service that ultimately ranks artists in the primary market.

### 3 Mining social media for market data

The review of the available art information services makes it apparent that just like for any other market, the actors in the art world require information, data analysis and statistics. Art buyers behave just like any other buyer; they are looking to maximize their gains. The main distinction is that art buyers not only wish to maximise monetary gain but also non-monetary gain from the aesthetic and social values inherent in artworks, that is the cultural value as explained in Sect. 2.

A recent research topic is mining social media data to forecast stock market behavior. Most such approaches mine data for specific firm stocks or indexes and employ sentiment analysis or clustering techniques to forecast the daily stock market behavior or stock return, risk and stock comovement (Bollen et al. 2011; Oliveira et al. 2013; Liu et al. 2015; Smailovic et al. 2013; Arias et al. 2014). The notion behind these approaches is that the general public mood is correlated with stock market behavior.

All the above mentioned approaches for estimation and forecasting cannot be directly applied to the art market since there are no publicly available price data. Furthermore, as it has already been stated in Sect. 2, an artwork as a product possesses cultural values, that affect its price, mostly inherent with the artist's recognition and appreciation by the art world. Thus, it is proposed to mine social media data to estimate the inherent cultural values of an artist's name. The closest analogous to this is measuring brand equity, awareness and perceptions. Hence, the suggestion to treat artists as brands for this study.

A brand's equity value infers intangible aspects like perceived quality, customer loyalty and satisfaction. Recently, more than a few methodologies have been presented that measure a brand's equity value with the utilization of social media data mining. Following, the equity value is often used to estimate a firm's stock market behavior (Yu et al. 2013). Chung (2015) mine Twitter data to determine the adjectives used to describe a brand and its products. Tucker (2015), He et al. (2013) collect product reviews from Twitter and utilize sentiment analysis to classify them as positive or negative to ascertain equity value. Other approaches obtain product reviews from other social media to predict stock market performance based on equity value (Luo et al. 2013). Similar approaches have also been used for measuring a city's brand equity value (Andéhn et al. 2014).

Overall, the foundation of all the above mentioned methodologies is that they attempt to extract the wider possible public opinion for a brand and its products. Their main differences lie in the particular approaches used to collect data, clean data, estimate sentiment and finally in which ways the public opinion will be exploited for price or other estimates.

Yet again, the approaches for brand equity estimation cannot be directly applied to the art market since only the opinion of certain elite groups are important in the social construction of cultural value.

A quite different approach is presented by Culotta and Cutler (2016) to mine brand perceptions from Twitter. The authors wish to correlate brands with specific attributes (e.g. eco-friendliness). Instead of mining tweets about the specified attribute, a user network is built that is relevant to the attribute. Then, the brand's correlation to the attribute is determined as the correlation of its follower's with the user network. Topic specific user networks will be exploited in our proposed methodology as it will be later detailed in Sect. 5.

Subsequent sections of this paper will detail the proposed social media data mining methodology that takes into consideration the particularities of the art market.

## 4 Data collection and cleaning

### 4.1 Data collection

Currently there are numerous social media sites, each one of them with a different set of characteristics and features for their users. For our study, we will mine data from Twitter. Twitter is a microblogging service with over 300 million users (Statista 2016), who post short messages, up to 140 characters. Twitter posts are appropriate for our research due to the large number of publicly available messages and their predominantly textual nature.

Twitter offers developers access to data through three different APIs. The Twitter Search API (Twitter Inc. 2016d) allows to retrieve tweets posted during the last 7 days using search criteria (e.g. keywords, usernames, etc.). Approximately 500 million tweets are generated each day and Twitter imposes certain limitations on the amount of data that can be retrieved (Twitter Inc. 2016a) for the safety of their infrastructure but also because Twitter does not want to offer publicly all user data. A developer is limited to making 180 search requests in a 15 min period. Each search query for a specific keyword returns approximately the most recent 5000 tweets.

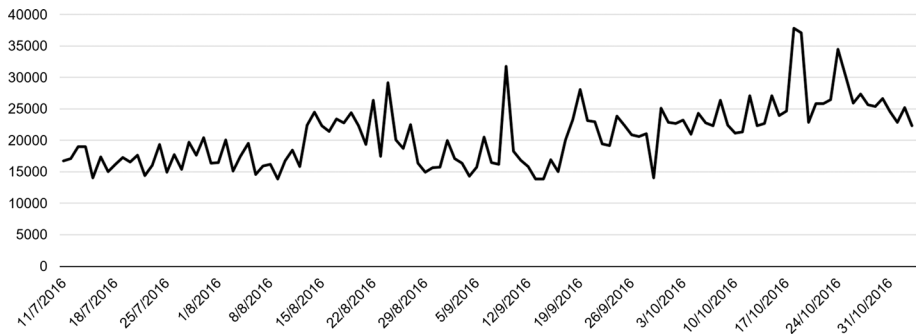
The Twitter Sample Streaming API (Twitter Inc. 2016c) provides tweets as they happen in near real-time. Again, API users can specify the keywords they wish to track. Hence, the streaming API will return in real-time tweets that are relative to the specified keywords. However, the main drawback is that it only provides a small random sample of the total tweets; in most cases 1–2% of the total tweets generated. This percentage is dependent on the actual popularity of the tracked keywords.

Finally, the third way to access Twitter data is by the Twitter Firehose (Twitter Inc. 2016b). The Twitter Firehose is very similar to the Twitter Sample Streaming API only that it guarantees that it will return in near real-time the 100% of generated tweets that match the specified criteria. However, the Twitter Firehose is not free as it is available only through Twitter's enterprise API platform GNIP.<sup>1</sup>

---

<sup>1</sup> <https://gnip.com/>.





**Fig. 1** The total number of tweets retrieved daily for all 548 artist names from July 11th 2016 to November 11th 2016

In our implementation data were collected daily from July 11th 2016 to November 11th 2016 through the Twitter Search API with the utilization of the Tweepy library for Python (Roesslein 2016). Data were collected for the 548 unique artists included in the Artprice Top 500 in 2015 (AMMA & artprice.com 2016) and the Artfacts.Net Top 100 (Artfacts 2016b). The Artprice ranking is based on factual auction revenue data and therefore it refers only to the secondary market. The Artfacts.Net ranking is the only one available that utilizes data other than auction sales, measurements on exhibition success, and therefore includes some evidence on the primary market although its precision cannot be assessed as it was discussed in Sect. 2.1.1. Hence, it was decided to include all 500 artists in the Artprice ranking and the top 100 artists in the Artfacts ranking to subsume an adequate sample of artists that might be successful only in the primary market but not in the secondary. This will allow us to evaluate the results of our approach against both rankings as well as with regard to the primary and secondary market. Additionally, the proposed methodology for the construction of a topic-specific user network will be evaluated as to how well it represents Twitter mentions for these 548 artists.

Every day Twitter was searched for posts that included the name<sup>2</sup> of each one of the 548 artists considered in our study. The total number of tweets retrieved daily is shown in Fig. 1. To overcome some of the limits imposed by the Twitter Search API, a pagination technique was utilized (Twitter Inc. 2016f). Due to the nature of the Twitter Search API there were fetched retweets for which the original tweet was not retrieved. A procedure was implemented where the missing original messages were requested. In this manner, the total number of retrieved messages were further increased by approximately 10%.

In total, over 2.6 million tweets were retrieved from approximately 1 million unique users. It should be noted that the search was limited in tweets written in English. Nonetheless, this does not affect the global nature of our search since users most often write posts in English, even if they are located in a non-English speaking country, when they want to comment something that overreaches the interest of the local public. The 2.6 million retrieved tweets are associated with 150 unique countries. The country field in a Tweet object (Twitter Inc. 2016e) as retrieved from the Twitter API is not always present.

<sup>2</sup> The Twitter Search API returns tweets from all users that include the search string either in plain text or a hashtag.



Nevertheless, we still can get an indication about the validity of our statement that even though we search only for tweets written in English our search reflects the global mentions of an artist's name.

## 4.2 Data cleaning

One of the main difficulties when mining tweets for specific keywords is to verify which tweets are actually relevant. This problem arises due to word polysemy but also particularly in our study due to synonymy.

Since we search for tweets that mention an artist's name the problem of synonymy becomes central. A representative example is with the search for the artist Claudio Bravo which also returned a significant number of tweets about the famous football player Claudio Bravo. Cases like this one inflict a serious integrity problem to the collected data.

Typically in such cases topic modelling techniques are utilized for polysemy disambiguation. In essence, topic modelling techniques analyze plain-text documents for semantic structure. This is achieved by representing a corpus of documents as sparse vectors on which statistical techniques are applied to find word co-occurrence patterns that represent a topic. Popular algorithms for topic inference include the Latent Semantic Analysis (LSA) (Hofmann 1999) and the Latent Dirichlet Allocation (LDA) (Blei et al. 2003).

Although the above mentioned topic modelling techniques have been applied successfully to numerous applications, the same does not hold when applied to short tweets. Since these techniques rely on word frequencies, it is anticipated that it will be hard to infer which words will be more correlated in each short text (Hong and Davison 2010). Furthermore, it is also hard to deduce the context of a short message for polysemy or synonymy disambiguation.

Quite a few approaches are present in the literature that attempt to alleviate the identified problems with topic modelling in short messages (Weng et al. 2010; Cheng et al. 2014; Zhao et al. 2011). In our implementation we employed the approach presented by Weng et al. (2010) where tweets posted by each individual user are aggregated into a single document. This methodology overcomes the deficiencies of short texts as noted previously and is one of the most popular currently in the literature. Although the authors applied tweet aggregation to identify the topics a user is interested into, we will apply it to identify the topic of a tweet.

After tweets were aggregated into documents for each unique user, topic modelling was performed with the LDA implementation of the Gensim library for Python (Rehurek and Sojka 2010). The latest dump of all English Wikipedia articles (Wikipedia 2016) was utilized as a training corpus. Then, similarity queries for each document against the identified Wikipedia topics were performed.

As it was expected this methodology worked well in cases a user had made enough posts. For example, the user Man City News (@mcnewsmcfc), from whom 277 tweets were retrieved, was correctly identified with a football related topic. Hence, all 277 tweets by this user can be disregarded since it is now certain that they do not mention Claudio Bravo the artist. The same holds for a personal Twitter account whose 286 tweets were retrieved when querying about the artist Christo. LDA correctly identified this user to a topic related to video games since this account's posts were talking about Christo the character from the video game *Disgaea 5: Alliance of Vengeance*.

The setback with the tweet aggregation methodology is that it cannot work well for all retrieved tweets. This is due to the fact that almost 70% of the unique users, whose tweets were retrieved, have posted only a single tweet. As a result, the aggregation methodology would not improve LDA results for topic identification. The aforementioned percentage of users with only a single tweet is significantly higher than the stated 15% in the study of Weng et al. (2010).

Intuitively, this could be explained as the result of the common habit to mention very well-known artist names to comment on circumstances unrelated to the art world. For example, when querying about Vincent van Gogh a tweet was fetched with the text *“bruh i swear i just seen vincent van gogh on the train”* by a personal Twitter account who posted only this single tweet regarding van Gogh. However, 11,229 retweets of this post were also retrieved. This number of retweets could affect significantly our later evaluation of how often van Gogh is mentioned if they are not correctly identified as not relevant.

Thus, we will next present a new methodology for evaluating a tweet’s topic relevance based on the originating user’s position in a topic-specific network of users.

## 5 Topic-specific network of users for topic modelling and user influence estimation

As we have already described in Sect. 2, the art market is elitist. Therefore, not all opinions are equally significant in deciding what is important or of high value in the art world. This perception has to be accounted in the construction of the topic-specific network of users.

In the retrieved tweets, there are many occasions where the name of a very well-known artist is used to describe something that is not necessarily pertinent to the purpose of our study. For example, quite a few twitterers comment on plastic surgeons by using Picasso as a reference. One such tweet is: *“Never go to a plastic surgeon with pictures of Picasso on his office wall.”*. This type of tweets undoubtedly do not infer valuable insight on the inherent cultural value of an artist’s name. The only information we can extract from tweets like these is that certain artists are well-known to the general public. The difficulty in estimating an artist’s cultural value is further deepened by the frequent tweets with well-known artists’ quotes or titles of widely recognized artworks. In most cases tweets like these is a mean of expressing a user’s personal feelings.

As we have already stated, in this study we propose to treat artists as brands; for which similar methodologies to ours have been used to measure brand awareness or brand equity. Although, the previously mentioned type of tweets would still be valuable in the case of brands, as wide recognition is typically desired, in the elitist art world wide recognition is not necessarily commended.

Consequently, the topic-specific network of Twitter users will be structured in an elitist manner to reflect the art world. Our methodology for the construction of the elitist topic-specific network was stimulated by the notion of *exemplar* Twitter users used by Culotta and Cutler (2016). The authors attempted to find individuals or organizations that are strongly affiliated with a specific attribute (e.g. eco-friendliness) so that they later can assess the perceived relationship between a brand and this attribute.

## 5.1 Construction of a topic-specific network of users

The elitist topic-specific network of users is built with the following steps summarized below:

### 1. *Input*

- $E$ : the manually defined set of exemplar Twitter accounts

### 2. *Collect followees and followers*

- $F_E$ : Collect the set of Twitter accounts,  $F_E$ , that each exemplar  $E_i$  is following. The term *followees* will be used to denote this group of users.
- $O_E$ : Collect the set of Twitter accounts,  $O_E$ , that follow each exemplar in  $E_i$ . The term *followers* will be used to denote this group of users.

### 3. *User topic expertise and influence evaluation*

- Evaluate the topic expertise,  $T_u$ , for each user  $u \in E \cup F_E \cup O_E$ .
- Evaluate the topic influence,  $I_u$  for each user  $u \in E \cup F_E \cup O_E$ .

## 5.2 Selection of the exemplar twitter accounts

Culotta and Cutler (2016) utilized an automated methodology for determining the exemplar Twitter accounts. Users included in more than two Twitter Lists,<sup>3</sup> both relative to a specific topic, are selected as exemplar users.

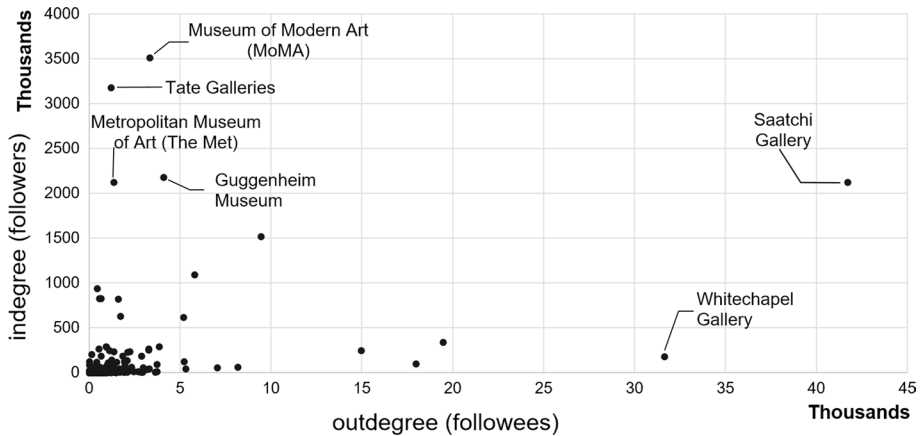
This approach was assessed but it was proven inappropriate for our study. Twitter lists related to the general topic of fine arts do not reflect the desired elitist nature. These lists mostly contain the Twitter accounts of widely recognized museums, galleries and artists. Hence, a manual approach was adopted to select a relatively small set of exemplar Twitter accounts.

We utilized the following lists of the most influential people and organizations in the art world as they were determined by established magazines and art news websites:

- ArtReview Power 100 (ArtReview 2015, 2016)
- artnet news 100 Most Influential People in the Art World (artnet news 2015, 2016)
- Artsy The 20 Most Influential Young Curators in the United States (Gotthardt 2016)
- Artsy The 20 Most Influential Young Curators in Europe (Bier 2016)
- Artsy The 15 Most Influential Art World Cities of 2015 (Artsy 2015)

The Twitter accounts of the people and organizations included in the above lists were manually entered into the set of exemplar Twitter accounts,  $E$ . Finally, the set of exemplar Twitter accounts resulted with a size of  $|E| = 287$ . The complete list of all exemplar accounts is shown in “Appendix 1”. The exemplar set is comprised of Twitter accounts of museums, galleries, art fairs, art news websites and magazines, auction houses, art foundations, curators, collectors, artists as well as art project spaces. The distinctive factor of the

<sup>3</sup> Any user can create a Twitter List that is a manually curated collection of Twitter accounts that are grouped together under the list’s topic.



**Fig. 2** Scatter plot showing the correlation between followees and followers of the exemplar Twitter accounts

utilized lists when compared to Twitter lists is that they include the most influential entities instead of the most widely recognized. Following, the set of exemplar users will function as the initial seed for populating the user network.

### 5.3 Collection of the exemplars' followees and followers

All Twitter accounts that each exemplar Twitter account,  $E_i$  is following are collected and form the set  $F_E$ , the *followees*. The resulting size of this set is  $|F_E| = 17,645$ . This size represents the unique users that all exemplar Twitter accounts are following.

The set  $O_E$  was also constructed by collecting the *followers* of each exemplar Twitter account. Exemplar accounts with more than 200,000 followers were excluded from this procedure. As it will be explained in the next Section, the expertise of a follower is inversely proportional to the total number of an exemplar's followers. Consequently, followers of such exemplar accounts cannot offer significant expertise or influence and therefore are disregarded.

The unique collected followers of all exemplar Twitter accounts are  $|O_E| = 2,419,865$ . It can be immediately noticed that there is a great difference in the size of  $|F_E|$  and  $|O_E|$ . In general, opinion-makers are expected to have a large number of followers (Leavitt et al. 2009). The number of followers of a user indicates the size of the audience for that user and it is commonly termed as the *indegree* of a user (Cha et al. 2010). Given that exemplar Twitter accounts were manually selected as the most influential people and organizations in the art world it is anticipated that many of them would have a large number of followers. Conversely, the number of other accounts a given user is following is commonly termed as the *outdegree* of a user.

In Fig. 2 the correlation between *followees* and *followers* of exemplar Twitter accounts is shown. It can be observed that there are a few outliers regarding the number of their followers or followees. The Museum of Modern Art (MoMA) in New York and the Tate Galleries are included in the most visited art museums in the world with millions of annual visitors. The great extent to which these accounts surpass the indegree of all other exemplar Twitter account confirms the tendency of Twitter users to follow widely recognized

entities (Liu et al. 2014). Twitter accounts with such followers to followees ratio are termed as *content generators* or *opinion makers*. Accounts with a number of followees proportional to their followers are termed as *conversationalists* (Leavitt et al. 2009). This term signifies that they are more interested in engaging into conversations rather than pushing content to other users. Then, as the number of a Twitter account's followees grows and tends to exceed its number of followers, it demonstrates a behavior of a *listener* rather than that of a content generator. Finally, users with a near equal number of followees and followers are usually *spammers*. Spammers aggressively follow other users to take advantage of the common habit of Twitter users to follow back their followers. The insights behind the number of followers and followees of a Twitter account will next be utilized to evaluate a user's topic expertise and influence.

## 5.4 User topic expertise evaluation

In almost every methodology present in the literature for the estimation of a user's topic expertise or influence, a recursive bottom-to-top approach is utilized (Gayo-Avello 2013). This is because it is attempted to deduce which users own the maximum expertise. Conversely, in our methodology maximum expertise users in the constructed network are known a priori. They are the exemplars which were manually selected for this very purpose. Hence, an inverse top-to-bottom approach is utilized for the topic expertise evaluation since the maximum expertise of exemplars will be transferred to all its followees and followers in the manner that is detailed next.

The topic expertise,  $T_u$ , for all users  $u \in E \cup F_E \cup O_E$  is in the range [0,1]. Initially, the topic expertise of exemplar Twitter accounts is set to the maximum value equal to 1.

Next, the topic expertise of the accounts that exemplars are following, the *followees*  $F_E$ , will be determined by a weighted voting scheme. The notion behind the choice of a weighted voting scheme is that when an exemplar Twitter account follows another account, it is like voting that this other account is of similar expertise with itself. Furthermore, the behavior of an exemplar in following other users will be taken into consideration when their power of vote will be assigned. Hence, exemplars that are more selective in who they choose to follow will be given more voting power. This way the elitist nature of the user network will be preserved. The total votes each followee receives will determine its topic expertise.

Formally, the voting power of each exemplar,  $W_{E_i}$ , is determined by:

$$W_{E_i} = \frac{|O_{E_i}|}{|F_{E_i}|} \quad (1)$$

The use of the ratio of followers,  $|O_{E_i}|$ , to followees,  $|F_{E_i}|$ , is drawn from Leavitt et al. (2009) where it was utilized to characterize a Twitter's account behavior as an opinion maker, a conversationalist, a listener or a spammer.

Table 1 gives some descriptive examples that illustrate how the voting power is distributed among exemplars. The Museum of Contemporary Art Tokyo is the exemplar with the highest voting power and the Museo Rufino Tamayo the one with the lowest. The first is highly selective in following other accounts. The latter has more followees than followers which makes it a listener rather than an opinion maker. It must also be noted that Saatchi Gallery has a voting power very close to the one of Alexandra Munroe and Hirshhorn

**Table 1** Examples of the voting power given to exemplar Twitter accounts according to their number of followers and followees

Screen name	Name	Followers	Followees	Vote weight
@MOT_art_museum	Museum of Contemporary Art Tokyo	$8.5 \times 10^4$	$1.1 \times 10^1$	$4.4 \times 10^{-1}$
@Tate	Tate	$3.2 \times 10^6$	$1.2 \times 10^3$	$1.5 \times 10^{-1}$
@metmuseum	Metropolitan Museum of Art (The Met)	$2.1 \times 10^6$	$1.4 \times 10^3$	$8.9 \times 10^{-2}$
@centrepompidou	Centre Pompidou	$8.3 \times 10^5$	$5.7 \times 10^2$	$8.2 \times 10^{-2}$
@LACMA	Los Angeles County Museum of Art (LACMA)	$8.3 \times 10^5$	$6.5 \times 10^2$	$7.2 \times 10^{-2}$
@MuseumModernArt	Museum of Modern Art (MoMA)	$3.5 \times 10^6$	$3.3 \times 10^3$	$6.0 \times 10^{-2}$
@Guggenheim	Guggenheim Museum	$2.2 \times 10^6$	$4.1 \times 10^3$	$3.0 \times 10^{-2}$
@whitneymuseum	Whitney Museum	$8.2 \times 10^5$	$1.6 \times 10^3$	$2.9 \times 10^{-2}$
@MOCAlosangeles	MOCA Los Angeles	$1.9 \times 10^5$	$1.8 \times 10^3$	$5.7 \times 10^{-3}$
@akmunroe	Alexandra Munroe (Samsung Senior Curator, Asian Art and Senior Adviser, Global Arts @guggenheim)	$1.1 \times 10^4$	$2.1 \times 10^2$	$3.0 \times 10^{-3}$
@saatchi_gallery	Saatchi Gallery	$2.1 \times 10^6$	$4.2 \times 10^4$	$2.9 \times 10^{-3}$
@hirshhorn	Hirshhorn Museum and Sculpture Garden	$6.4 \times 10^4$	$1.3 \times 10^3$	$2.9 \times 10^{-3}$
@_TheWhitechapel	Whitechapel Gallery	$2.4 \times 10^5$	$1.5 \times 10^4$	$9.3 \times 10^{-4}$
@rivolicast	Castello di Rivoli	$2.8 \times 10^4$	$1.7 \times 10^3$	$9.3 \times 10^{-4}$
@DALcontemporary	Dallas Contemporary	$9.4 \times 10^3$	$5.8 \times 10^2$	$9.2 \times 10^{-4}$
@museotamayooax	Museo Rufino Tamayo	$2.1 \times 10^1$	$1.2 \times 10^2$	$9.9 \times 10^{-6}$

**Table 2** Examples of the estimated topic expertise for followees of the exemplars after the voting procedure

Screen_name	Name	Topic expertise	# of votes
@nytimesarts	New York Times Arts	$9.966 \times 10^{-1}$	191
@SFMOMA	San Francisco Museum of Modern Art (SFMOMA)	$9.959 \times 10^{-1}$	165
@smithsonian	Smithsonian	$9.959 \times 10^{-1}$	112
@walkerartcenter	Walker Art Center	$9.959 \times 10^{-1}$	151
@artinfodotcom	BLOUIN ARTINFO	$9.560 \times 10^{-1}$	185
@Phaidon	Phaidon (publisher of the creative arts)	$5.042 \times 10^{-1}$	91
@MaxAndersonUSA	Maxwell L. Anderson (New Cities Foundation)	$5.013 \times 10^{-1}$	54
@TheEconomist	The Economist	$5.012 \times 10^{-1}$	81
@NortonSimon	Norton Simon Museum	$4.996 \times 10^{-1}$	38
@SuzanneMalveaux	Suzanne Malveaux (Proud Mom, CNN National Correspondent)	$9.885 \times 10^{-5}$	2
@YinkaShonibare	Yinka Shonibare (London-based artist)	$8.583 \times 10^{-5}$	2
@WSJdeals	Wall Street Journal Deals	$6.692 \times 10^{-5}$	2

Museum although the number of their followees and followers are of different magnitude. However, Saatchi Gallery is the exemplar with the highest number of followees which makes it highly unselective.

Following, the topic expertise of each exemplar followee,  $T_{F_E^i}$ , is evaluated as the sum of all votes casted by exemplars,  $E' \subseteq E$ , that follow  $F_E^i$ . The total number of followees of all exemplars,  $|F_E|$  is used as a normalization factor, so that topic expertise values are in the range  $[0, 1]$ .

$$T_{F_E^i} = \frac{1}{|F_E|} \sum_{E_i \in E'} W_{E_i} \quad (2)$$

In Table 2, some representative examples of the evaluated topic expertise are shown that validate the proposed methodology. As it can be seen there are quite a few Twitter accounts that their expertise was evaluated close to 1, that is almost the expertise of exemplars.

A similar procedure with the above, a weighted voting scheme, will be exploited for the determination of the topic expertise for the followers of exemplar Twitter accounts,  $O_E$ .

An ingoing following connection to an exemplar account is regarded of less importance than an outgoing connection. Any Twitter account can follow another one without the need of an approval or confirmation. Thus, most Twitter users, that are not opinion makers, typically follow others than being followed (Liu et al. 2014), i.e. they have more followees than followers. They decide to follow a Twitter account in a casual manner without necessarily denoting that they really want to read every post they make or that they have genuinely mutual interests (Cha et al. 2010). For example, 3.53 million users follow the Twitter account of the Museum of Modern Art (MoMA). It is not realistic to anticipate that all 3.53 million users are genuinely interested in modern art or in everything that MoMA posts.

The votes that each exemplar will cast for its followers will signify a vote of confidence in their mutual interest on the topic. Voting power is set to be inversely proportional to the number of followers. This is analogous to the notion that a Twitter account's influence or expertise on the topic is distributed evenly among all of its followers (Gayo-Avello 2013). Formally, the voting weight of each exemplar,  $W'_{E_i}$ , is determined by:

$$W'_{E_i} = \frac{1}{|O_{E_i}|} \quad (3)$$

Following, the topic expertise of each follower,  $T_{O_E^i}$  in the range  $[0, 1]$ , is evaluated as the sum of all weighted votes of exemplars  $E'' \subseteq E$  that are followed by  $O_E^i$  as:

$$T_{O_E^i} = \left| \frac{|F_{O_E^i}|}{|O_{O_E^i}|} - \frac{|O_{O_E^i}|}{|F_{O_E^i}|} \right| \sum_{E_i \in E''} W'_{E_i} \quad (4)$$

The absolute value of the difference of the followers,  $|O_{O_E^i}|$ , to followees,  $|F_{O_E^i}|$ , ratio with its inverse is utilized to identify spammers in addition to rewarding vigorous accounts either they are opinion makers or listeners. Spammers are expected to have an absolute difference close to zero since the number of their followers and followees will be almost equal.

In Table 3 a set of representative examples of the topic expertise evaluation for followers of exemplars is shown. As it can be observed, accounts @ManuelGiraudier and @FIAC have the highest expertise since they mainly followed accounts that are not widely recognized and possess a higher voting power. On the other hand, although account @e\_flux follows almost half the exemplar accounts its expertise is lower. Accounts like @AnaVFCH, @love\_pocan and @pavankpanesar that follow a very small number of the exemplars and



**Table 3** Examples of the estimated topic expertise for followers of the exemplars after the voting procedure

Screen_name	Name / description	Followers	Followees	Topic expertise	# of votes
@ManuelGiraudier	President at Giraudier Group Fine Art	26	872	$9.43 \times 10^{-1}$	41
@FIAC	Contemporary art fair and outdoor projects in Paris	31,627	600	$9.40 \times 10^{-1}$	58
@e_flux	International art network	79,274	1964	$7.68 \times 10^{-1}$	109
@josephklibansky	Dutch Artist	62,521	228	$7.11 \times 10^{-1}$	26
@demartinonunzio	Nunzio De Martino (Artist)	152	1994	$6.96 \times 10^{-1}$	119
@CuratorsINTL	Independent Curators International (ICI)	36,323	1804	$5.34 \times 10^{-1}$	91
@MonopolMagazin	Monopol-Magazin fr Kunst und Leben aus Berlin	20,252	1700	$5.28 \times 10^{-1}$	97
@MagnumPhotos	Photographic cooperative	836,499	462	$5.25 \times 10^{-1}$	11
@AestheticaMag	Aesthetica Magazine	52,423	8717	$3.33 \times 10^{-1}$	136
@AnaVFCH	Personal #blog // #Photography	426	1891	$3.56 \times 10^{-4}$	8
@love_pocan	Photographer ~ Lover of Life	414	1382	$1.49 \times 10^{-6}$	2
@WOAEstore	WOA E-store	66,694	67,035	$7.87 \times 10^{-7}$	36
@pavankpanesar	Pharmacology graduate	62	189	$8.40 \times 10^{-8}$	1

**Table 4** Examples of the retweet and mention influence of exemplar accounts

Screen name	Name	Followers	Retweets	Mentions
@Tate	Tate	$3.18 \times 10^6$	8494	1521
@saatchi_gallery	Saatchi Gallery	$2.12 \times 10^6$	5270	37
@MuseumModernArt	Museum of Modern Art (MoMA)	$3.51 \times 10^6$	4218	372
@metmuseum	Metropolitan Museum of Art (The Met)	$2.12 \times 10^6$	4311	224
@Guggenheim	Guggenheim Museum	$2.18 \times 10^6$	2435	437
@Sothebys	Sotheby's	$7.94 \times 10^4$	2574	234
@aiww	Ai Weiwei	$3.39 \times 10^5$	2212	585
@LACMA	Los Angeles County Museum of Art	$8.26 \times 10^5$	2221	146
@TheBroad	The Broad	$2.19 \times 10^4$	1663	509
@GettyMuseum	J. Paul Getty Museum	$1.09 \times 10^6$	1003	93
@Gagosian	Gagosian Galleries	$2.48 \times 10^5$	422	175
@whitneymuseum	Whitney Museum	$8.19 \times 10^5$	745	216
@_WhiteCube	White Cube Gallery	$9.65 \times 10^4$	518	172

the most widely recognized obtain a very low topic expertise value. Finally, spam accounts like the @WOAEstore also receive a very low topic expertise value.

## 5.5 User topic influence evaluation

Many of the methodologies present in the literature measure influence in a user network with metrics similar to the ones we utilized previously for the topic expertise evaluation (Leavitt et al. 2009; Gayo-Avello 2013). Certainly, a user's expertise and influence are highly correlated. Besides, many studies conclude that measuring influence in a topic specific network is more accurate than globally (Weng et al. 2010). In the context of our study, we are only interested in an elitist topic-specific network of users. That is, a network that will include a rather limited number of users with maximum expertise rather than a network that attempts to be as complete as possible. Moreover, as it was defined in Sect. 2, we are seeking to evaluate the socially constructed aesthetic and social values integrated in an artist's name. In conclusion, the proposed user topic influence evaluation will endeavor to measure the user's contribution to the social construction of aesthetic and social values.

Cha et al. (2010) state that there are three activities that represent different types of influence:

- indegree influence* the number of followers of a user. However, it was found to indicate the user's popularity rather than its ability to engage audience.
- retweet influence* the number of retweets that contain one's name. It indicates the ability to generate content.
- mention influence* the number of mentions that contain one's name. It indicates the ability to engage in conversations.

They also conclude that retweets are driven by the content value of a tweet and mentions are driven by the name value of the user. Thus, the aesthetic value will be associated with retweet influence and the social value with mention influence.

**Table 5** Examples of the retweet and mention influence of followees of exemplar accounts

Screen name	Name/description	Topic expertise	# of votes	Retweets	Mentions
@HISTORY	Official HISTORY Twitter.	$4.155 \times 10^{-2}$	4	40,505	246
@Argy_Papa_	20th + 21st century art + artists	$6.218 \times 10^{-3}$	3	32,450	246
@YouTube	YouTube	$1.114 \times 10^{-1}$	24	1474	9017
@Etsy	Etsy. The marketplace we make together.	$4.092 \times 10^{-2}$	13	1423	3100
@nytimesarts	New York Times Arts	$9.966 \times 10^{-1}$	191	3470	102
@SFMOMA	San Francisco Museum of Modern Art	$9.959 \times 10^{-1}$	165	1131	193
@smithsonian	Smithsonian	$9.959 \times 10^{-1}$	112	509	86
@walkerartcenter	Walker Art Center	$9.959 \times 10^{-1}$	151	133	5
@artinfodotcom	BLOUIN ARTINFO	$9.560 \times 10^{-1}$	185	939	82
@HuffPostArts	HuffPost Culture	$7.410 \times 10^{-1}$	138	2099	333

In Table 4, the exemplar accounts with the highest retweet and mention influence are shown. It can be observed that an account's influence is not directly relevant to its number of followers. Tate and MoMA have approximately the same number of followers but their influence is significantly different. Furthermore, retweet and mention influence are also not directly relevant. For example, Saatchi Gallery has a high number of retweets but on the other hand an unexpectedly low number of mentions. However, it should be stressed out that influence measurements stemmed from the collected set of 2.6 million tweets related to the specific 548 artists considered in our study. Hence, the presented influence metrics are relative to those specific artists that represent only a fragment of the art world.

In Tables 5 and 6, the retweet and mention influence of followees and followers of exemplar accounts is shown. In both cases, there are accounts that were evaluated with a very low topic expertise that exhibit the highest influence. Such accounts are content generators; they have many followers and post many tweets. For example, the History channel's twitter account posts quite often quotes and stories about very well-known artists and many of its 1.5 million followers tend to retweet these posts. On the other hand, numerous users share links to YouTube videos related to artists with Twitter posts where they also tend to mention the name of YouTube's twitter account. Hence, the high retweet influence of the History channel and the high mention influence of YouTube. Though, in our study retweet and mention influence is utilized to determine the aesthetic and social value. Consequently, the total aesthetic and social influence,  $I_u$ , for each user  $u \in E \cup F_E \cup O_E$  will be in proportion to its topic expertise,  $T_u$ , to eliminate the influence of users like YouTube or the History channel. Formally, the influence,  $I_u$ , is determined by:

$$I_u = T_u(R_u + M_u) \quad (5)$$

where  $R_u$  is the retweet count and  $M_u$  is the mention count that are considered of equal importance. The influence values,  $I_u$ , scaled to the range  $[0, 10]$  to threshold their effect will be later utilized to weight each tweet when artist rankings are evaluated.

**Table 6** Examples of the retweet and mention influence of followers of exemplar accounts

Screen name	Name/description	Topic expertise	# of votes	Retweets	Mentions
@GoogleExpertUK	Susan Dolan. Google Expert	$6.582 \times 10^{-3}$	77	10,020	1,715
@___NL	Nans Lombard. Contemporary Waste.	$1.305 \times 10^{-2}$	22	4001	213
@ stefano_pesce	ART COLLECTION & EXHIBITION management	$1.291 \times 10^{-2}$	53	1489	74
@ManuelGiraudier	President at Giraudier Group Fine Art	$9.431 \times 10^{-1}$	41	3470	102
@FIAC	Contemporary art fair and outdoor projects in Paris	$9.399 \times 10^{-1}$	58	1131	193
@e_flux	International art network	$7.678 \times 10^{-1}$	109	509	86
@josephklibansky	Dutch Artist	$7.110 \times 10^{-1}$	26	133	5
@demartinonunzio	Nunzio De Martino (Artist)	$6.961 \times 10^{-1}$	119	939	82

**Table 7** Artist ranking based only on the frequency of all retrieved tweets and after LDA topic identification was performed

Retrieved			LDA with aggregation		
Rank	Artist	Tweets	Rank	Artist	Tweets
1	Vincent van Gogh	476,879	1	Vincent van Gogh	26,391
2	Pablo Picasso	317,845	2	Pablo Picasso	19,374
3	Claude Monet	196,317	3	Andy Warhol	18,321
4	Claudio Bravo	161,168	4	Claude Monet	9693
5	Andy Warhol	121,964	5	Banksy	9249
6	Banksy	110,128	6	Claudio Bravo	7089
7	Salvador Dali	74,014	7	Salvador Dali	6978
8	Edgar Degas	60,171	8	Edgar Degas	6891
9	Yoko Ono	53,758	9	Edvard Munch	6579
10	Edvard Munch	45,615	10	Francis Bacon	6513
11	Henri Matisse	45,488	11	Henri Matisse	5862
12	Francis Bacon	42,115	12	Gustav Klimt	5856
13	Gustav Klimt	34,187	13	Jean-Michel Basquiat	5796
14	Jean-Michel Basquiat	31,068	14	Yoko Ono	5487
15	Wassily Kandinsky	28,023	15	Wassily Kandinsky	5304
16	Paul Klee	24,951	16	Paul Klee	5289
17	Christo	21,142	17	Jackson Pollock	5172
18	Jackson Pollock	20,847	18	Rembrandt van Rijn	4989
19	Rembrandt van Rijn	20,648	19	Mark Rothko	4689
20	Mark Rothko	20,487	20	Man Ray	4428

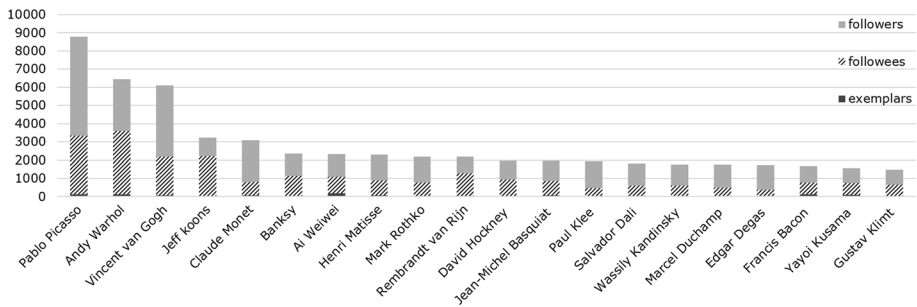
## 6 Results

Previous sections described the proposed methodology for an artist ranking system that included the following steps:

- collect Twitter posts by any user that mention any of the 548 artist names considered in our study,
- clean retrieved tweets with the use of the LDA topic inference algorithm with aggregation,
- construct a topic-specific user network comprised of manually selected exemplar Twitter accounts as well as their followees and followers,
- evaluate users' expertise in the network
- evaluate users' influence in the network

Following, the artist rankings that resulted from our proposed methodology will be presented and discussed to validate the main goal of the proposed system. Namely, to estimate the inherent cultural value of an artist's name in an elitist manner, i.e. the more often an artist is mentioned by users of high expertise and influence the higher the estimated inherent cultural value.

Initially, in Sect. 6.1 rankings based only on the frequency of all retrieved tweets will be shown. These rankings will serve as a reference for comparison with the rankings presented



**Fig. 3** Artist ranking based only on the frequency of retrieved tweets from all users in the topic-specific network

in Sect. 6.2, where only tweets from users in the topic-specific network are utilized for ranking. Furthermore, the effect of gradually increasing the considered user expertise is shown. Finally, in Sect. 6.3 the effect in rankings of weighing the importance of tweets, according to evaluated users' influence is shown.

In all rankings only the top 20 artists will be shown for clarity purposes.

## 6.1 Artist ranking based on LDA with aggregation

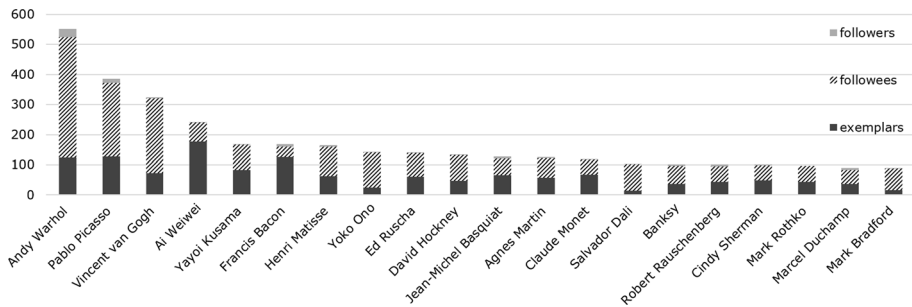
Initially, the artist ranking based only on the frequency of all retrieved tweets is shown. In Table 7, the ranking based on the tweet frequency of all retrieved tweets is presented as well as after the LDA topic inference algorithm with aggregation is applied as described in Sect. 4.2.

When the LDA with aggregation methodology was applied almost 75% of retrieved tweets were rejected as irrelevant. It is not unusual for such a big percentage of data to be rejected after topic modelling is applied (Weng et al. 2010). In Sect. 4.2, the main issues related to word polysemy and synonymy that lead to redundant retrieved data were highlighted. Indicative examples on the reasons for rejecting a great number of tweets after the LDA with aggregation algorithm is applied are:

- synonymy of the artist Claudio Bravo with the footballer,
- synonymy of the artist Christo with the video game character,
- news media tweets reporting the retrieval of stolen van Gogh paintings,
- Yoko Ono tweets related to her work as a musician.

Nevertheless, there are still noisy data present in the ranking after LDA with aggregation is applied. For example, there is still a significant number of tweets related to Claudio Bravo the footballer that could not be identified as sports related. This is mainly due to the significant number of users that have posted only a single tweet; in which case the LDA with aggregation cannot perform well. Thus, Claudio Bravo moves down only two positions in the ranking; Yoko Ono moves down five positions; Christo is no longer in the top 20 and van Gogh still remains in the first position.

In the following section, to eliminate noisy data only tweets from users in a topic-specific network are considered.



**Fig. 4** Artist ranking based only on the frequency of retrieved tweets from users in the topic-specific network with expertise > 0.3

## 6.2 Artist ranking based on the evaluated user expertise

The process for the construction of a topic-specific user network was detailed in Sect. 5. Initially, a set of manually selected exemplar accounts is created to represent the most important users in the art world since our goal is to build an elitist user network as it was described in Sect. 5.2. Then, the user network construction is completed with the collection of the followees and followers of exemplar accounts.

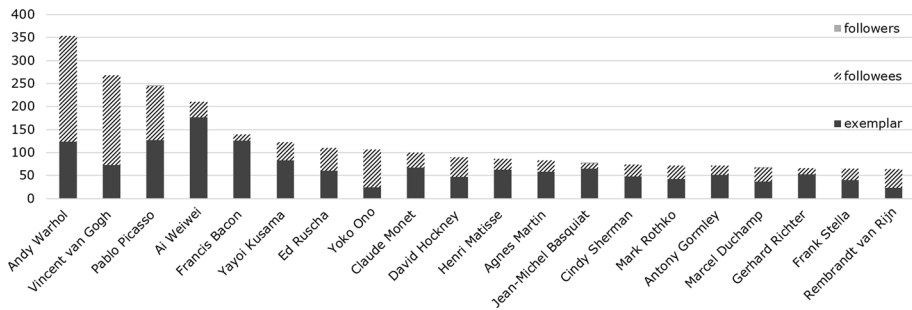
In Fig. 3, the ranking based on the frequency of retrieved tweets only from users in the topic-specific network is shown. It can be observed that Claudio Bravo is no longer present in the top 20 since almost all retrieved tweets were referring to the footballer. Many of the van Gogh related tweets regarding the news about the retrieval of his stolen paintings were not by users in the network and therefore not evaluated in the ranking. Thus, now he holds the 3rd position. Many other differences can be also observed in the current ranking:

- Salvador Dali, Edgar Degas, Francis Bacon and Gustav Klimt move down over seven positions in the ranking,
- Edvard Munch, Yoko Ono, Jackson Pollock and Man Ray are no longer present in the top 20,
- Jeff Koons, Ai Weiwei, David Hockney, Paul Klee, Marcel Duchamp and Yayoi Kusama enter the top 20.

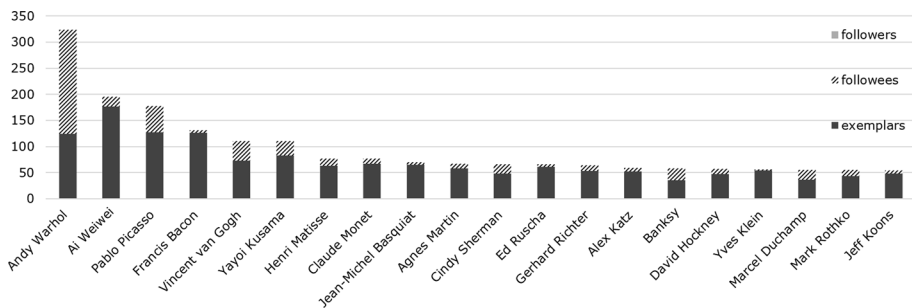
The total number of tweets considered in this ranking was further reduced by 54%. However, only 92 artists are not mentioned at all from the constructed topic-specific user network. These 92 artists were mentioned in only 1512 tweets out of the total 2.6 million originally retrieved. Hence, we can deduce that the topic-specific network produces a reliable sample data set of tweets for our study.

Moreover, the utilization of the topic-specific user network provides a valuable insight on what type of users mostly tweet about each artist. It is apparent in Fig. 3 that the highest volume of tweets is from followers since they are the most populous group of users. It is most interesting though to notice that followers post predominantly about the most widely recognized artists: Picasso, Warhol and van Gogh. This validates our claim while constructing the user network that followers are of lesser importance than followees of exemplars. Therefore, it can be said that the set of followers represents the general public opinion in the art world.





**Fig. 5** Artist ranking based only on the frequency of retrieved tweets from users in the topic-specific network with expertise > 0.5

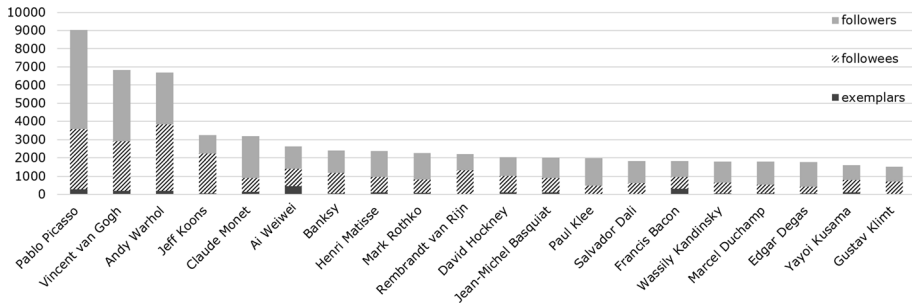


**Fig. 6** Artist ranking based only on the frequency of retrieved tweets from users in the topic-specific network with expertise > 0.7

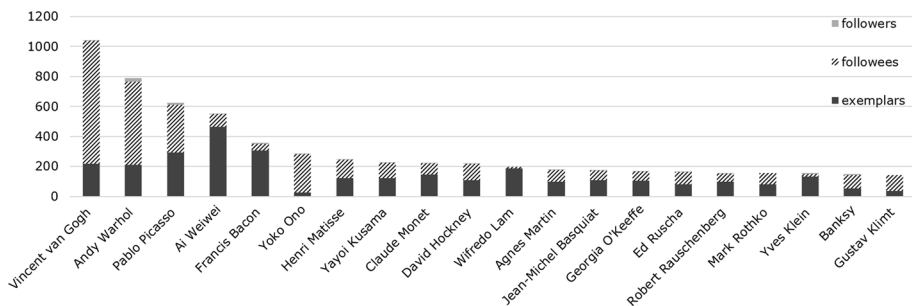
Following, in Figs. 4, 5 and 6 the artist ranking is shown when tweets only from users with an evaluated expertise of over 0.3, 0.5 and 0.7, respectively, are considered. It can be observed that as the user expertise raises, the contribution of followers in the ranking is almost eliminated whereas the exemplars' contribution becomes authoritative. This aspect is highly desired while evaluating artist rankings, in view of the already highlighted fact that the art world is elitist.

Furthermore, it can be noticed that as the considered user expertise raises the artists present in the ranking is significantly different. There are only 7 artists present both in the ranking with tweets from users with an expertise of over 0.7, shown in Fig. 6, and the ranking after LDA with aggregation was applied, shown in Table 7. There are important differences even when the later shown rankings that consider tweets from users in the topic-specific network is considered. This feature further highlights the fact that different type of users talk about different artists.

In conclusion, it can be observed that the higher considered user expertise the less widely recognized artists are present in the ranking who are mentioned by users in a casual manner and not due to an event occurring in the art world with impact. Therefore, artists like Rembrandt, Dali and Kandinsky are not present in Figure's 6 ranking whereas artists like Agnes Martin, Cindy Sherman and Gerhard Richter enter it. Certainly, there are other widely recognized artists still present, but some of them indeed still own greater cultural



**Fig. 7** Artist ranking based on the influence of retrieved tweets from all users in the topic-specific network



**Fig. 8** Artist ranking based on the influence of retrieved tweets from users in the topic-specific network with expertise > 0.3

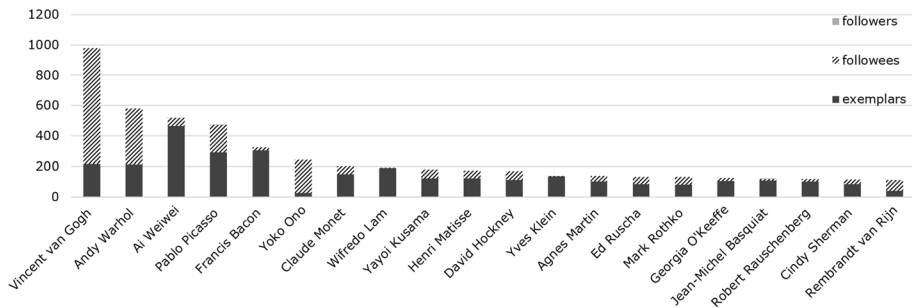
value that is reflected on either auction revenue or exhibition success. For example, Warhol and Picasso occupy the 1st and 3rd position in the final ranking shown in Fig. 6. They also occupy the first two positions in both Artprice and Artfacts rankings. These two artists may probably be the ones that almost everybody recognizes but they also have the greatest auction revenue according to Artprice and the greatest exhibition success according to Artfacts.

In the following section, the evaluated influence for each user in the topic-specific network is also utilized to determine artist rankings. Hence, the effect of the type of users that talk about artists is further enhanced in rankings.

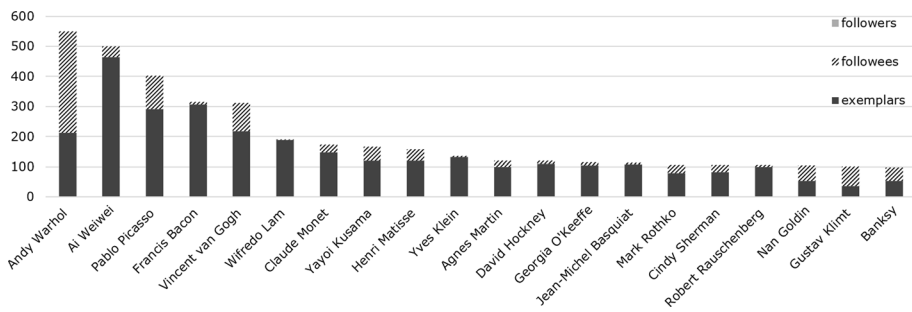
### 6.3 Artist ranking based on the evaluated user expertise and influence

The user influence was evaluated using the retweet and mention count as described in Sect. 5.5. In this study, user influence is exploited to suggest the inherent cultural values (aesthetic and social) in an artist's name. Thus, the influence of a twitterer is used to pose the importance of a tweet in terms of cultural value. Hence, not all tweets contribute equally when counting the frequency of tweets that mention an artist's name. The scaled user influence,  $I_u$ , is used to weight the importance of a tweet.

In Fig. 7, the artist ranking based on the influence of retrieved tweets from all users in the topic-specific network is given. The obtained ranking is almost identical to the one



**Fig. 9** Artist ranking based on the influence of retrieved tweets from users in the topic-specific network with expertise > 0.5



**Fig. 10** Artist ranking based on the influence of retrieved tweets from users in the topic-specific network with expertise > 0.7

shown in Fig. 3 where only the frequency of retrieved tweets from all users in the topic-specific network was considered. In this case, the impact of influence is minimal due to high volume of evaluated tweets by followers.

When we compare the rankings only from users with an evaluated expertise of over 0.3 based on influence, shown in Fig. 8, with the ones based on frequency, shown previously in Fig. 4, the following key observations can be made:

- Vincent van Gogh moves to the 1st place. The van Gogh Museum had a high number of retweets and mentions due to the news about the retrieved stolen paintings. This led into an evaluation of high influence. Furthermore, since the van Gogh Museum is an artist dedicated museum it posts tweets about van Gogh daily.
- Wifredo Lam and Georgia O'Keeffe appear in the ranking. Tate Modern organized periodic exhibitions for both artists that opened while data were collected for our study. Tate Modern is the user with the highest influence and made many posts about these exhibitions which were also retweeted by numerous other users in the network.

When the influence only from users with an expertise of more than 0.5, shown in Fig. 9, no significant changes can be observed in the ranking. However, when an expertise of more than 0.7 is considered, shown in Fig. 10, it can be noticed that:

- Vincent van Gogh moves back to the 5th position. The van Gogh Museum's expertise was evaluated just under 0.7 and therefore its tweets do not contribute to the ranking. Thus, only what other users post is evaluated to rank van Gogh.
- Yoko Ono is no longer present in the top 20. Yoko Ono has a personal Twitter account that exists in the user network as a followee of exemplars. Her account's expertise was also evaluated just under 0.7 and therefore posts originating from her own account are not considered in the evaluation.

Furthermore, when rankings in Figs. 6 and 10 are compared, in which only tweets from high expertise users are considered but in the latter tweets are also weighted according to user influence, it can be observed that Ed Ruscha, Gerhard Richter, Alex Katz, Marcel Duchamp and Jeff Koons have been replaced by Wifredo Lam, Georgia O' Keeffe, Robert Rauschenberg, Nan Goldin and Gustav Klimt. Although, the former artists are still placed in quite high positions in the ranking, positions 20–30, they have been outranked by the latter artists as more influential users tweeted about them. For example, Wifredo Lam has tweets about his concurrent exhibitions in October 2016 at Tate Modern and Philips Auction as well as tweets by Sotheby's about a sale of one his works at 1.1m shortly before the opening of his retrospective at Tate Modern. On the other hand, Ed Ruscha was tweeted by less influential users mainly about a screening of the only two films he ever created at MOCA Los Angeles, his exhibitions at the de Young Museum, the Crown Point Press Gallery and the Gagosian Gallery London as well as about a sale of one his works at \$18,750 by Sotheby's. Another representative example is that of Jeff Koons and Robert Rauschenberg. Jeff Koons was mentioned more often by from less influential users mainly about his participation in The Met's online series The Artists Project, a lawsuit regarding one of his sculpture's, laying off half of the painters working at his studio as well as announcements for auctions of some of his works at Doyle Auctions and Heritage Auction. On the other hand, Robert Rauschenberg was tweeted by more influential users due to his exhibitions at Tate Modern, MoMA New York, San Francisco MoMA as well as the Ullens Center for Contemporary Art (UCCA) in China.

To conclude, the desired features of an elitist ranking system, as stated in the beginning of this paper, are evident in the presented results when a high expertise and influence of users are considered. The effect of general comments and widely recognized artist name mention has been limited. On the other hand, the effect of mentions from authoritative users is augmented where they are mainly initiated from exhibition events. Undoubtedly, certain limitations have also been identified that will be discussed in the following Sections.

## 6.4 Comparison with other rankings

The proposed ranking system will be compared against two well known rankings: the art-price Top 500 artists by auction revenue in 2015 (AMMA & artprice.com 2016) and the Artfacts.Net Top 100 (Artfacts 2016b). However, a direct comparison cannot be performed. The foremost cause is that the data utilized in these two rankings refer to different periods of time. Our data were collected for a duration of 4 months, from July 2016 to November 2016. The artprice ranking refers to auction sales data in 2015 and the artfacts ranking to accumulated data up to the beginning of 2016. Consequently, data collected for a much

**Table 8** The percentage of artists that the proposed system was able to rank for each compared ranking system and each period of art

	Artfacts		Artprice	
	Artists (%)	Ranked	Artists (%)	Ranked (%)
Old masters	0.00	–	10.60	1.89
19th century art	0.00	–	7.20	30.56
Modern art	13.00	92.31%	42.80	18.69
Post-war art	42.00	57.14%	24.20	29.75
Contemporary art	43.00	62.79%	15.20	38.16
Total		63.00%		23.40

longer period of time are required to be able to make a direct comparison that refers to the same time period.

Nonetheless, a qualitative comparison against the other two rankings will be attempted. In Table 8, the percentage of artists for which the proposed system collected enough data and was able to rank them is shown for each period of art.<sup>4</sup> Initially, it should be noticed that the artifacts and artprice rankings are very different. The artifacts Top 100 is comprised of artists primarily from the Post-War and Contemporary art period. On the other hand, in the artprice ranking almost half of the artists are from the Modern Art period. As it has been noted in the beginning of this paper, in Sect. 2, the primary and secondary art market are different in many ways and this is reflected in the two rankings. The proposed system collected enough data and was able to rank 63% of the artists included in the artifacts Top 100 but only 23.40% of the ones included in the artprice ranking.

It should also be noted that it is common to both Artprice and Artfacts rankings that contemporary artists, i.e. artists that were born after 1945, do not dominate the top 20 positions. In the Artprice ranking there are only two contemporary artists present in the top 20 positions. This is expected since Artprice is focused on the secondary market. On the other hand, in the Artfacts ranking that utilizes measurements of exhibition success more contemporary artists are present in the top 20 positions, specifically six. Finally, in the final ranking of the proposed approach, shown in Fig. 10, there are five contemporary artists present (Ai Weiwei, Jean-Michel Basquiat, Cindy Sherman, Nan Goldin, Banksy). In conclusion, it is common in the art world that artists from older periods of art dominate the top 20 positions in rankings either due to auction revenue or exhibition success.

There are two key properties that enable the proposed system to rank an artist: the retrieved data and the type of users that belong in the topic-specific user network. Consequently, more research is required to deduce whether a different network of users or different type of data are required for the secondary art market considered in the artprice ranking. An important factor that needs also to be considered is that China accounted for 30% of the annual auction revenue in the global Art Market (AMMA & artprice.com 2016). However, almost all major Chinese auction houses do not have a Twitter account and as a

<sup>4</sup> The periods of art are defined by the year an artist was born, as used in the artprice ranking:

- Old Masters: works by artists born before 1760.
- 19th Century Art: works by artists born between 1760 and 1860.
- Modern Art: works by artists born between 1860 and 1920.
- Post-War Art: works by artists born between 1920 and 1945.
- Contemporary Art: works by artists born after 1945.

**Table 9** Comparison with the artifacts and artprice rankings for the artists present in the top 20 of the proposed system illustrated in Fig. 10 when the evaluated user expertise and influence are considered

	Artist	Artfacts rank	Artprice rank
1	Andy Warhol	1	2
2	Ai Weiwei	23	120
3	Pablo Picasso	2	1
4	Francis Bacon	438	6
5	Vincent van Gogh	238	15
6	Wifredo Lam	858	165
7	Claude Monet	228	3
8	Yayoi Kusama	33	42
9	Henri Matisse	47	25
10	Yves Klein	119	38
11	Agnes Martin	249	106
12	David Hockney	61	114
13	Georgia O'Keeffe	491	111
14	Jean-Michel Basquiat	221	17
15	Mark Rothko	292	8
16	Cindy Sherman	6	214
17	Robert Rauschenberg	12	58
18	Nan Goldin	54	–
19	Gustav Klimt	598	54
20	Banksy	4113	352

result the users interested in these auctions were not represented in the constructed topic-specific network.

On the other hand, most of the artists present in the artifacts Top 100 were ranked by the proposed system. More importantly, the discussion of the results in Sect. 6.3 provides evidence that the proposed system ranks artists based on their exhibition success as it is the case with the artifacts ranking system. In Table 9, the artifacts and artprice ranking is shown for the artists present in the top 20 of the proposed system illustrated in Fig. 10 when the evaluated user expertise and influence are considered. The common characteristic of all rankings is that Andy Warhol and Pablo Picasso are in the top positions. In general, for other artists there is a great difference in the position they occupy in each ranking. This is expected since the artprice ranking is based on auction turnover, the artifacts on exhibition success and the proposed system measures exhibition success by means of reputation success. Moreover, each ranking utilizes data from a different period of time.

Nevertheless, the case of the artist Banksy should be pointed out. Banksy is a street artist and although he has international recognition and a significant auction turnover that places him in the artprice ranking, he appears in an extremely low position in the artifacts ranking, the 4113th position. As the case of Banksy is discussed in Periferic Biennial (2008), his very low position in the artifacts ranking is due to the different type of art he makes that is not well represented in the registered exhibition spaces and data in the artifacts database. It is promising that the proposed system was able to rank more appropriately Banksy since influential users talk about his work although he does not hold exhibitions in established museums or galleries.

Another example is the case of Nan Goldin who is ranked in a high position in the artifacts ranking but is not present in the artprice top 500. Again in this case the proposed

system ranked Nan Goldin appropriately although her annual auction turnover is not as high but she is appreciated and recognized in the primary market.

As a result, the presented preliminary results are very promising since they exhibit the ability of the proposed system to collect data and measure the reputation and success of artists from a diverse collection of sources that originate from different periods of art as well as different art practices. More data have to be collected though, for a long time period of at least a year, to be able to make a more direct comparison with the other two rankings.

In conclusion, the proposed system can offer two main advantages over the other considered ranking systems. The first one is that it can provide rankings that reflect the present time. The two other systems update their rankings quarterly or even yearly due to the demanding process of manually collecting data. The proposed system is able to continuously integrate newly collected data into the rankings. Additionally, it is also able to present rankings based on accumulated data over time or only on data collected during a specified period of time that is of interest. Certainly, if rankings for instance based on monthly obtained data are observed there will be high fluctuations. This will occur since rankings will reflect only where the short-term attention of the art world is. However, as obtained data will be accumulated over larger periods of time, e.g. a whole year or even a 5-year period, these fluctuations will be limited, as it is the case with the Artfacts ranking where data are accumulated from long periods of time (in some cases since 1996). In this case, rankings will reflect the long-term attention of the art world.

The other significant advantage the proposed system can exhibit is that it can provide detailed insights on how rankings were evaluated. For the calculation of rankings, a tweet's importance is scaled according to the user's influence. Therefore, tweets that contribute significantly in an artist's ranking can be identified. These tweets will infer insights on the events and the users that largely impacted each artist's ranking. Thus, the potential user of the system can further evaluate the presented rankings based on his/her specific interests.

## 7 Conclusions and future work

In this paper, a social data mining methodology was proposed to rank artists. The proposed ranking system takes into account the singularities of the art market. The main difficulties and differentiations of the art economy when compared to other economies are summarized below:

- there are no publicly available price data,
- the market value of an artwork is dependent on its inherent cultural value,
- the cultural value is socially constructed by elite social groups with higher cultural status or higher purchasing power,
- the features and structure of the primary and secondary art market are different in many ways.

The main goal of the proposed system is to mine social data to obtain trend estimates for the art market. Based on the singularities of the art market, the main goal becomes to estimate the inherent cultural value of an artists name, since market value and cultural value are interdependent. Thus, the obtained rankings represent artists with the highest estimated cultural value which could be reflected in the market value of their artworks. As such, obtained artist rankings can also be viewed as art market trends.



Estimates of cultural value, that is socially constructed by elite social groups, are provided by measuring how often an artist is mentioned by users of high expertise and influence. The proposed elitist topic-specific user network determines the expertise and influence of users and was shown that it can reflect the elitist nature of the art world. The evaluated user expertise incorporates into rankings that the attention an artist receives from the art world is important rather than wide recognition from the general public. Furthermore, when the evaluated user influence is also incorporated into rankings, reputation within the art world is esteemed. The contribution of the general public opinion for widely recognized artists was minimized whereas the contribution of certain influential users was authoritative for the final artist ranking determination.

A comparison with two other ranking systems was performed to highlight the advantages of the proposed methodology as well as topics for further research and improvement. The proposed methodology has provided evidence that it can collect data, estimate cultural value and rank a great variety of artists that originate from different periods of art as well as different art practices. However, it was deduced that it cannot represent well all the characteristics that govern the secondary market and especially the Chinese art market. Conversely, promising results were indicated regarding the measurement of exhibition success that largely governs the primary art market. The accuracy of the exhibition success measurement cannot be evaluated since there is no “ground truth” available. Other ways for evaluation, possibly with the participation of art market experts, will have to be considered in the future.

Two main advantages of the proposed system were identified. The first one is that it does not rely on manual collection of data to estimate rankings, as it is the case with other systems currently available. Hence, rankings can be updated continuously to signify current trends in the art market. The second core advantage is that potential users of the system can explore the events that had the greatest impact on ranking results for further assessment based on their particular interests.

To the best of our knowledge, this is the first attempt to mine social media to extract quantitative and qualitative data for the art market. This study being preliminary in a novel application undoubtedly requires further research in a number of identified topics. Firstly, more data have to be collected for a longer period of time and for a larger number of artists to further validate the obtained results. A reverse procedure should also be investigated where all exemplars’ tweets are retrieved and artist mentions are detected, thus evaluation of rankings is not limited to artists included in a predetermined list.

The selection of which users are regarded as exemplars governs greatly the structure of the user network and consequently the obtained rankings. Therefore, alternative ways, automated or semi-automated, for the initial setup of the set of exemplar accounts have to be explored. The objective of exemplar accounts is to represent as closely as possible experts in the art world. However, not all art world experts are Twitter users and those who are exhibit an online behavior different than the one in the real world. Although some may be highly influential in the art world their online presence is limited and thus not well represented in mined data. Another issue identified is the potential bias that can affect rankings from accounts that predominately post about themselves (e.g. artist’s personal accounts, artist dedicated museums).

Consequently, alternate approaches for user expertise and influence measurement should be considered to investigate their performance on the previously identified issues. Firstly, other metrics than retweet and mention count should be incorporated for influence evaluation as well as ranking evaluation. An example would be how many different users posted for an artist. Another approach would be to continuously reassess the expertise and

influence of users, either in the user network or not, while collecting new tweets. This approach would result in an adaptive user network structure. Thus, being able to include new expert and influential users as well as determining their main characteristics. This approach could also contribute to the identification of the distinctive characteristics of the secondary market that are not currently well represented in the proposed methodology.

Finally, further research for the development of a formal model for the construction of elitist user networks based on topic expertise is essential for the application of the proposed approach to other markets.

## Appendix 1: List of exemplar accounts

The exemplar accounts were selected in the manner described in Sect. 5.2. Following, the complete list of the manually selected exemplar accounts is presented.

1. @theartmarket, The Art Market. News and developments in the global art market, covering commercial galleries, museums, auctions and art fairs.
2. @MOT\_art\_museum, Museum of Contemporary Art, Tokyo
3. @ArtistisPresent, Marina Abramovic, Artist
4. @documenta\_13, documenta\_13, Exhibition
5. @Tate, Tate Art galleries in UK: Tate Britain, Tate Modern, Tate Liverpool & Tate St Ives.
6. @AiANews, Art in America. The World's Premier Art Magazine since 1913.
7. @JeffKoons, Jeff Koons, Artist
8. @metmuseum, The Met (The Metropolitan Museum of Art), New York
9. @centrepompidou, Centre Pompidou, Paris
10. @LACMA, LACMA (Los Angeles County Museum of Art), Los Angeles
11. @MuseumModernArt, MoMA (Museum of Modern Art), New York
12. @YukoHasegawa, Yuko Hasegawa, Curator
13. @Guggenheim, Guggenheim Museum, New York
14. @whitneymuseum, Whitney Museum, New York
15. @MoMAPS1, MoMA PS1, Nonprofit Art Center, New York
16. @Takashipom\_En, Takashi Murakami, Artist
17. @Qatar\_Museums, Qatar Museums, Qatar
18. @HitoSteyerl, Hito Steyerl, Artist
19. @brooklynmuseum, Brooklyn Museum, New York
20. @almayassahamad, Al Mayassa Al Thani, Chairperson of @Qatar\_Museums, @Doha-Film, @ReachOutToAsia, @QLCentre and Vice Chairperson of @AlShaqab
21. @newmuseum, New Museum, New York
22. @ArtBasel, Art Basel, Switzerland
23. @ArtObserved, Art Observed, Global contemporary and fine art events and news observed from New York City.
24. @Stedelijk, Stedelijk Museum, Amsterdam
25. @artmarket, Art Market Monitor, Updates on artists, auctions and sales.
26. @artstagesg, Art Stage Singapore, International Art Fair, Singapore
27. @ArtTactic, ArtTactic, An insider's view to the art market, London–New York
28. @nicholasserota, Sir Nicholas Serota, Current director of the Tate
29. @Gagosian, Gagosian Gallery, Worldwide

30. @GettyMuseum, J. Paul Getty Museum, Los Angeles
31. @lindsaypollock, Lindsay Pollock, NYC arts writer
32. @Artforum, Artforum, Defining the world of contemporary art since 1962, New York
33. @KimbellArt, Kimbell Art Museum, Texas
34. @MAMoficial, Museu de Arte Moderna de Sao Paulo, Brazil
35. @artnet, artnet, The Art World Online, New York
36. @klausbiesenbach, Klaus Biesenbach, Director, MoMA PS1 & Chief Curator at Large, The Museum of Modern Art
37. @TheasterGates, Theaster Gates, Artist, Director of Arts + Public Life, UChicago Arts, Rebuild Foundation, Black Monks of Mississippi
38. @la\_Biennale, Biennale di Venezia, Italy
39. @FriezeArtFair, Frieze Fairs, presents three of the leading international art fairs in London and New York.
40. @FTLifeArts, Financial Times's Life & Arts, London
41. @TheArtNewspaper, The Art Newspaper, Compelling stories and in-depth analysis about the art world, New York–London
42. @frieze\_magazine, frieze, the leading magazine of contemporary art and culture, London / Berlin / New York
43. @ARTnewsmag, ARTnews, The oldest and most widely circulated art magazine in the world.
44. @MOCALosangeles, The Museum of Contemporary Art, Los Angeles
45. @E\_Perrotin, Emmanuel Perrotin, French art dealer specializing in contemporary art.
46. @\_WhiteCube, White Cube Galleries, London and Hong Kong
47. @OkwuiEnwezor, Okwui Enwezor, Curator and Art critic
48. @nodecenter, Node Center for Curatorial Studies, Berlin
49. @ArtReview\_, ArtReview, the leading international, independent, critical voice in contemporary art, London
50. @artsy, Artsy, New York
51. @Museo\_MAXXI, MAXXI (Museo Naziolane Delle Arti Del XXI Secolo), Rome
52. @MuseoGuggenheim, Guggenheim Bilbao, Spain
53. @thearmoryshow, The Armory Show, International Art Fair, New York
54. @Sothebys, Sotheby's, Auction House,
55. @PaceGallery, Pace Gallery, Contemporary art gallery with locations in New York, London, Hong Kong, Beijing, Seoul, Paris, and Palo Alto.
56. @SerpentineUK, Serpentine Galleries, London
57. @luisastrina, Galeria Luisa Strina, Brazil
58. @tuymanstour, Luc Tuymans, Artist
59. @ICALondon, The Institute of Contemporary Arts, London
60. @Lisson\_Gallery, Lisson Gallery, London–New York
61. @DallasMuseumArt, Dallas Museum of Art, Texas
62. @galapagosDET, Galapagos Art Space, Detroit, Michigan
63. @FondazionePrada, Fondazione Prada, Institution dedicated to art and culture, Milan and Venice, Italy
64. @gamt torino, Galleria Civica d'Arte Moderna e Contemporanea (GAM), Torino, Italy
65. @akmunroe, Alexandra Munroe, Samsung Senior Curator, Asian Art and Senior Adviser, Global Arts at Guggenheim
66. @saatchi\_gallery, Saatchi Gallery, London
67. @hirshhorn, Hirshhorn Museum and Sculpture Garden, Washington
68. @rhizome, Rhizome, Commissioning, presenting, and preserving digital art

69. @theamoncarter, Amon Carter Museum, American Art Museum, Texas
70. @berlinbiennale, Berlin Biennale
71. @SharjahArt, Sharjah Art Foundation, United Arab Emirates
72. @HauserWirth, Hauser & Wirth Gallery, Zrich, London, Somerset, New York and Los Angeles
73. @crystalbridges, Crystal Bridges Museum of American Art, Bentonville, Arkansas
74. @abyrosen, Aby Rosen, Principal of RFR Holding, LLC, Lever House Art Collection, New York
75. @FundacionJumex, Fundacin Jumex, Mexico
76. @HUObrist, Hans Ulrich Obrist, Curator
77. @FondTrussardi, Fondazione Trussardi, No profit institution for the promotion of Contemporary Art and Culture, Milan, Italy
78. @pamm, Prez Art Museum Miami, Florida
79. @gerhardrichter, Gerhard Richter, Artist
80. @studiomuseum, Studio Museum Harlem, New York
81. @themodernfw, The Modern Art Museum of Fort Worth, Texas
82. @Artcurial, Artcurial, Auction House
83. @thesubstation, The Substation, Contemporary arts centre, Singapore
84. @RubellMiami, The Rubell Family Collection (RFC), New York
85. @Contemporaryist, Contemporary Istanbul, Art Fair, Turkey
86. @XavierHufkens, Xavier Hufkens, Art gallery, Belgium
87. @garage\_mca, Garage Museum of Contemporary Art, Moscow
88. @CAPCmusee, CAPC muse d'art contemporain de Bordeaux, France
89. @MarianGoodman, Marian Goodman Gallery, New York, Paris, London
90. @biennalesydney, Biennale of Sydney
91. @davidzwirner, David Zwirner, Contemporary art gallery, New York–London
92. @deSarthe, de Sarthe Gallery, Hong Kong
93. @ChristiesInc, Christie's, Auction House
94. @mcchicago, Museum of Contemporary Art, Chicago
95. @thegetty, J. Paul Getty Trust, Nonprofit dedicated to the visual arts, Los Angeles
96. @DalArtsDistrict, Dallas Arts District, Texas
97. @SadieColesHQ, Sadie Coles HQ, Contemporary art gallery, London
98. @GladstoneNYC, Gladstone Gallery, New York
99. @bonhams1793, Bonhams, Auction House
100. @nashersculpture, Nasher Sculpture Center, Dallas, Texas
101. @SeanKellyNY, Sean Kelly Gallery, New York
102. @galerieperrotin, Galerie Perrotin, Paris, France
103. @GagosianLondon, Gagosian Gallery, London
104. @jeudepaume, Galerie nationale du Jeu de Paume, Paris, France
105. @RyanTrecartin, Ryan Trecartin, Artist
106. @PerformaNYC, Performa Arts, New York
107. @gagosiany, Gagosian Gallery, New York
108. @vanabbemuseum, Van Abbemuseum, Eindhoven, Netherlands
109. @simondepury, Simon de Pury, Auctioneer, Artist & Photographer
110. @NatoThompson, Nato Thompson, Artistic Director at Creative Time, New York
111. @TheBroad, The Broad, Contemporary art museum, Los Angeles, California
112. @kurimanzutto, kurimanzutto, Art gallery, Mexico
113. @aiww, Ai Weiwei, Artist
114. @\_TheWhitechapel, Whitechapel Gallery, London

115. @rivolicast, Castello di Rivoli, Museum of Contemporary Art, Torino, Italy
116. @DALcontemporary, Dallas Contemporary Museum, Texas
117. @DrZamalek, Graham Harman, Philosopher
118. @GContinua, Galleria Continua, San Gimignano, Beijing, Les Moulins, Habana
119. @MetroPictures, Metro Pictures Gallery, New York
120. @artbrussels, Art Brussels, Contemporary art fair, Belgium
121. @BlumandPoe, Blum & Poe, Art gallery, Los Angeles, California
122. @CFA\_Berlin, CFA (Contemporary Fine Arts Gallery), Berlin
123. @GagosianAsia, Gagosian Gallery, Hong Kong
124. @fondsr, Fondazione Sandretto Re Rebaudengo, Italy
125. @UCCA798, Ullens Center for Contemporary Art, Beijing, China
126. @LehmannMaupin, Lehmann Maupin, Contemporary art gallery, New York and Hong Kong
127. @arthk, ART HK, Hong Kong International Art Fair
128. @KochiBiennale, Kochi Biennale Foundation, India
129. @marcglimcher, Marc Glimcher, President of Pace Gallery
130. @natgalleriesg, National Gallery Singapore
131. @GagosianParis, Gagosian Gallery, Paris
132. @phillipsauction, Phillips, Auction House
133. @sculpturecenter, SculptureCenter, Contemporary art museum, New York
134. @trevorpaglen, Trevor Paglen, Artist
135. @AlainSeban, Alain Seban, President of the Pompidou Centre in Paris until february 2015
136. @Kasmingallery, Paul Kasmin Gallery, New York
137. @carterac, Carter Cleveland, CEO at Artsy
138. @KunsthalleBasel, Kunsthalle Basel, Switzerland
139. @AlmineRech, Almine Rech Gallery, Paris, Brussels, London, New York
140. @GagosianLA, Gagosian Gallery, Los Angeles
141. @Tate\_StIves, Tate St Ives
142. @lcorneil, Lauren Cornell, Curator and Associate Director, Technology Initiatives at the New Museum in New York City.
143. @Crow\_Collection, Trammell & Margaret Crow Collection of Asian Art, Dallas, Texas
144. @GagosianGeneva, Gagosian Gallery, Geneva
145. @textezurkunst, Texte zur Kunst, Berlin
146. @ThaddaeusRopac, Galerie Ropac, Paris, Salzburg, London
147. @airdeparis, Air de Paris, Art gallery in Paris, France
148. @HeritageAuction, Heritage Auctions, Auction House
149. @sp\_arte, SP-Arte, Art Fair, Brazil
150. @SAMUSO\_in\_Seoul, SAMUSO, Space for Contemporary Art, Seoul, South Korea
151. @ManifestaDotOrg, Manifesta Biennial, Amsterdam, The Netherlands
152. @tateliverpool, Tate Liverpool
153. @shanghart, ShanghART Gallery, China
154. @ICAinBOSTON, The ICA/Boston, Massachusetts, USA
155. @AsiaArtArchive, Asia Art Archive, Non-profit organisation, Hong Kong
156. @koeniggalerie, KNIG GALERIE, Berlin, Germany
157. @KunsthalleZH, Kunsthalle Zrich, Switzerland
158. @MuholiZanele, Zanele Muholi, Artist
159. @TMIGLA, The Modern Institute, Contemporary art gallery, Glasgow, Scotland

160. @JackShainman, Jack Shainman, Contemporary Art Gallery, New York
161. @CPPCisneros, Coleccin Cisneros, New York and Caracas
162. @robertelmes, Robert Elmes, Founder, Galapagos Art Space
163. @MoMA\_Warsaw, Museum of Modern Art, Warsaw, Poland
164. @HendrikFolkerts, Hendrik Folkerts, Curator
165. @CuratorOnTheRun, Eva Respini, Curator
166. @evapresenhuber, Presenhuber Galerie, Zurich
167. @bukowskis, Bukowskis, Auction House
168. @SimonLeeGallery, Simon Lee Gallery, London
169. @CharlesEsche, Charles Esche, Museum Director and Curator
170. @BerlinArtWeek, Berlin Art Week
171. @SaatchiArt, Saatchi Art, Los Angeles
172. @artbma, Baltimore Museum of Art, Maryland
173. @HausderKunst, Haus der Kunst, Munich, Germany
174. @GagosianRome, Gagosian Gallery, Rome
175. @maspmuseu, Museu de Arte de Sao Paulo, Brazil
176. @Slotover, Matthew Slotover, Co-founder, Frieze
177. @GagosianAthens, Gagosian Gallery, Athens
178. @Cultured\_Mag, Cultured Magazine
179. @ran\_dom, Domenico Quaranta, Contemporary art critic and curator
180. @Zabludowicz\_Col, Zabludowicz Collection, London
181. @philipitinari, Philip Tinari, Director, Ullens Center for Contemporary Art
182. @PSandretto, Patrizia Sandretto, President Fondazione Sandretto Re Rebaudengo
183. @ALAContemporary, ALAC (Art Los Angeles Contemporary)
184. @julietaaranda, Julieta Aranda, Artist
185. @gagosiansf, Gagosian Gallery, San Francisco
186. @artcentralhk, Art Central, Hong Kong
187. @RoseArtMuseum, The Rose Art Museum, Massachusetts
188. @documenta\_\_14, documenta 14, Exhibition
189. @SupporticoLopez, Supportico Lopez, Berlin, Germany
190. @The\_ADAA, ADAA (The Art Dealers Association of America), New York
191. @PearlLamFineArt, Pearl Lam Galleries, Hong Kong, Shanghai, Singapore
192. @vitamincreative, Vitamin Creative, Artist Space, China
193. @nadimsamman, Nadim Samman, Curator
194. @wittedewith\_\_, Witte de With Center for Contemporary Art, Rotterdam, Netherlands
195. @126gallery, 126 Gallery, Ireland
196. @154artfair, 1:54 Contemporary African Art Fair
197. @BroadMSU, Eli and Edythe Broad Art Museum, Michigan
198. @ashkal\_alwan, Ashkal Alwan, Nonprofit, Beirut, Lebanon
199. @anish Kapoor\_art, Anish Kapoor, Artist
200. @KukjeGallery, Kukje Gallery, Seoul, South Korea
201. @SpruethMagers, Sprueth Magers, Berlin, London & Los Angeles
202. @VictoriaSiddall, Victoria Siddall, Director, Frieze Fairs
203. @MendesWood, Mendes Wood DM, Sao Paulo, Brazil
204. @alexlogsdail, Alex Logsdail, Director Lisson Gallery
205. @eva\_Limerick, EVA International Biennial, Ireland
206. @Mnuchin\_Gallery, Mnuchin Gallery, New York
207. @pablodelabarra, Pablo Len de la Barra, Guggenheim UBS MAP Curator
208. @ArtIntl, Art International, Istanbul, Turkey

209. @RubaKatrib, Ruba Katrib, Curator
210. @clearinggallery, c-l-e-a-r-i-n-g Gallery, New York, Brussels
211. @TBA21, TBA21 (Thyssen-Bornemisza Art Contemporary), Vienna
212. @GavinBrownsEnt, Gavin Brown's Enterprise, New York
213. @BrantFoundation, The Brant Foundation Art Study Center, Greenwich, Connecticut
214. @daataeditions, Daata Editions
215. @Guggenheim\_HKI, Guggenheim Helsinki
216. @aCCeSsions\_bard, aCCeSsions
217. @Lopez\_Muse, Lopez Museum Library, Pasig, Philippines
218. @Hanne\_M, Hanne Mugaas, Director, Kunsthall Stavanger
219. @maryceruti, Mary Ceruti, Executive Director and Chief Curator of SculptureCenter
220. @MKechBiennale, Marrakech Biennale
221. @mdcgallery, Massimo De Carlo Gallery, Milan - London
222. @KomuNow, riyas komu, Artist
223. @tillmans\_art, Wolfgang Tillmans, Artist
224. @\_artandpractice, ART+PRACTICE, Nonprofit, Los Angeles
225. @karawalker\_art, Kara Walker, Artist
226. @GAM\_Milano, Galleria d'Arte Moderna, Milan
227. @WoolOfficial, Christopher Wool, Artist
228. @DLevyGallery, Dominique Lvy Gallery, New York–London
229. @ArcadiaMissa, Arcadia Missa, London
230. @wolfsonian, The WolfsonianFIU, Miami Beach, Florida
231. @DhakaArtSummit, The Samdani Art Foundation, Dhaka, Bangladesh
232. @MatthewMarks, Matthew Marks Gallery, New York
233. @dorotheum, Dorotheum, Auctin House
234. @Tajan\_Auction, TAJAN, Auction House
235. @defneayas, Defne Ayas, Director of Witte de With Center for Contemporary Art
236. @SundaramTagore, Sundaram Tagore Galleries, New York, Hong Kong, Singapore
237. @koyo\_kouoh, Koyo Kouoh, Independent exhibition maker and cultural producer
238. @sammlunggoetz, Sammlung Goetz, Munich, German
239. @Piasa\_auction, PIASA, Auction House
240. @AmyCappellazzo, Amy Cappellazzo, Chairman of the Fine Art division of Sotheby's
241. @KrishBose, Bose Krishnamachari, Artist & Curator
242. @TheLabSF, The Lab, San Francisco
243. @Kunsthall\_Stvgr, Kunsthall Stavanger, Norway
244. @benvickers\_, Ben Vickers, Curator
245. @TulcaFestival, TULCA Festival, Ireland
246. @mialocks, Mia Locks, Curator
247. @rjkhckly, Rujeko Hockley, Curator
248. @SavvyArtSpace, SAVVY Contemporary, Berlin, German
249. @mfsirmans, Franklin Sirmans, Curator
250. @sgfutureperfect, Future Perfect, Singapore
251. @leoxuprojects, Leo Xu Projects, Shanghai
252. @MaccaroneLLC, Maccarone Gallery, New York
253. @PhilippePirotte, Philippe Pirotte, Art historian, critic, and curator
254. @KmTemporaer, km temporaer, Zurich
255. @\_SUPERDAKOTA\_, SUPER DAKOTA, Contemporary art gallery, Brussels
256. @matthewabess, Matthew Abess, Curator
257. @jjonutyte, Juste Jonutyte, Curator



258. @IsabelleGraw, Isabelle Graw, Art historian and critic
259. @RavenelArt, Ravenel International Art Group, Auction House, Taiwan
260. @MoscowBiennale, Moscow Biennale
261. @LudwigJustine, Justine Ludwig, Curator
262. @erinzulie, Erin Christovale, Curator
263. @wormholelab, Jose Carlos Diaz, Curator
264. @sp\_muenster, Skulptur Projekte Mnster, Germany
265. @tanya\_leighton, Tanya Leighton Gallery, Berlin
266. @Biennale\_Young, Moscow International Biennale for Young Art
267. @Lempertzauktion, Lempertz Auctioneers, Auction House
268. @tadtweets, Tad Smith, President & CEO, Sotheby's
269. @Sam4Art, Samuel Leuenberger, Curator
270. @Yuzmuseum, Yuz Museum Shanghai
271. @art\_sixteen, Art Sixteen, Berlin
272. @NeueGalerieBER, Neue Galerie Berlin
273. @WANNENES, WANNENES, Auction House
274. @isagenzken, Isa Genzken, Artist
275. @BudiTek, buditek, founder: yuzmuseum shanghai
276. @NahmadProjects, Nahmad Projects, London
277. @FarsettiArte, Farsetti Arte, Auction House
278. @kentridge\_art, William Kentridge, Artist
279. @wbartdesign, West Bund Art & Design, Shanghai
280. @museotamayooax, Museo Rufino Tamayo, Mexico
281. @artpricedotcom, Artprice.com
282. @majahoffmann, Maja Hoffmann, Collector
283. @DesteFoundation, DESTE Foundation for Contemporary Art, Athens, Greece
284. @JAltmannParis, Thaddaeus Ropac, Art dealer
285. @documenta\_14, documenta 14, Exhibition
286. @K11ArtFoundtn, K11 Art Foundation, Hong Kong
287. @simcosclub, Stefan Simchowitz, Collector

## References

- Abbing, H. (2002). *Why are artists poor? The exceptional economy of the arts*. Amsterdam: Amsterdam University Press.
- AMMA & artpricecom. (2016). The art market in 2015. [http://imgpublic.artprice.com/pdf/rama2016\\_en.pdf](http://imgpublic.artprice.com/pdf/rama2016_en.pdf). November 1, 2016.
- Andéhn, M., Kazeminia, A., Lucarelli, A., & Sevin, E. (2014). User-generated place brand equity on twitter: The dynamics of brand associations in social media. *Place Branding and Public Diplomacy*, 10(2), 132–144. <https://doi.org/10.1057/pb.2014.8>.
- Arias, M., Arratia, A., & Xuriguera, R. (2014). Forecasting with twitter data. *ACM Transactions on Intelligent Systems and Technology*, 5(1), 8:1–8:24. <https://doi.org/10.1145/2542182.2542190>.
- Artfacts. (2016a). Artfacts analysis page for frank stella accessible without a subscription. [http://www.artfacts.net/index.php/pageType/career\\_analyser/artist/23/lang/1](http://www.artfacts.net/index.php/pageType/career_analyser/artist/23/lang/1). October 31, 2016.
- Artfacts. (2016b). Artfacts artist ranking—top 100. <http://www.artfacts.net/en/artists/top100.html>. October 31, 2016.
- Artfacts. (2016c). Artfacts homepage. <http://www.artfacts.net/>. October 31, 2016.
- Artnet. (2016). Artnet homepage. <http://www.artnet.com>. October 31, 2016.
- Artnet news. (2015). 100 most influential people in the art world. <https://news.artnet.com/art-world/100-art-world-influencers-part-one-373718>, <https://news.artnet.com/art-world/100-most-influential-art-peopl>

- e-part-two-374075, <https://news.artnet.com/art-world/most-influential-art-people-part-three-374084>. November 6, 2016.
- Artnet news. (2016). 100 most influential people in the art world. <https://news.artnet.com/art-world/artnet-titans-2016-part-one-654851>, <https://news.artnet.com/art-world/artnet-titans-2016-part-two-654854>, <https://news.artnet.com/art-world/artnet-titans-2016-part-three-654857>. November 6, 2016.
- ArtNexus. (2016). ArtNexus auction results. <http://artnexus.com/Auctions.aspx>. November 3, 2016.
- Artprice. (2016). artprice homepage. <http://www.artprice.com>. October 31, 2016.
- ArtReview. (2015). 2015 power 100. [https://artreview.com/power\\_100/2015/](https://artreview.com/power_100/2015/). November 6, 2016.
- ArtReview. (2016). 2016 power 100. [https://artreview.com/power\\_100/](https://artreview.com/power_100/). November 6, 2016.
- Artsy. (2015). Artsy editorial: The 15 most influential art world cities of 2015. <https://www.artsy.net/article/artsy-editorial-contemporary-art-s-most-influential-cities>. November 6, 2016.
- Becker, H. S. (1982). *Art worlds*. Berkeley: University of California Press.
- Bier, A. (2016). Artsy editorial: The 20 most influential young curators in Europe. <https://www.artsy.net/article/artsy-editorial-the-20-most-influential-young-curators-in-europe>. November 6, 2016.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993–1022.
- BlouinArtinfo. (2016). Blouin art sales index. <http://artsalesindex.artinfo.com/>. November 3, 2016.
- Boll, D. (2011). *Art for sale*. Berlin: Hatje Cantz.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1–8. <https://doi.org/10.1016/j.jocs.2010.12.007>.
- Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K. P., (2010). Measuring user influence in twitter: The million follower fallacy. In *ICWSM 10: Proceedings of international AAAI conference on weblogs and social*.
- Cheng, X., Yan, X., Lan, Y., & Guo, J. (2014). Btm: Topic modeling over short texts. *IEEE Transactions on Knowledge and Data Engineering*, 26(12), 2928–2941. <https://doi.org/10.1109/TKDE.2014.2313872>.
- Chung, A. D. (2015). Brand personality research on twitter. Master thesis, The University of Texas at Austin.
- Culotta, A., & Cutler, J. (2016). Mining brand perceptions from twitter social networks. *Marketing Science*, 35(3), 343–362.
- Danto, A. (1964). The artworld. *The Journal of Philosophy*, 61(19), 571–584. <http://www.jstor.org/stable/2022937>.
- FindArtInfo. (2016). Findartinfo homepage. <http://www.findartinfo.com>. November 3, 2016.
- Gayo-Avello, D. (2013). Nepotistic relationships in twitter and their impact on rank prestige algorithms. *Information Processing & Management*, 49(6), 1250–1280. <https://doi.org/10.1016/j.ipm.2013.06.003>.
- Gotthardt, A. (2016). Artsy editorial: The 20 most influential young curators in the united states. <https://www.artsy.net/article/artsy-editorial-the-20-most-influential-young-curators-in-the-united-states>. November 6, 2016.
- Grampp, W. D. (1989). *Pricing the priceless: Art, artists and economics*. New York: Basic Books.
- Graw, I. (2010). *High price: Art between the market and celebrity culture*. Berlin: Sternberg Press.
- He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, 33(3), 464–472. <https://doi.org/10.1016/j.ijinfomgt.2013.01.001>.
- Hofmann, T. (1999). Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international acm sigir conference on research and development in information retrieval, SIGIR '99* (pp. 50–57). New York, NY: ACM. <https://doi.org/10.1145/312624.312649>.
- Hong, L., Davison, B. D. (2010). Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics, SOMA '10* (pp. 80–88). New York, NY: ACM. <https://doi.org/10.1145/1964858.1964870>.
- Jin, X., Gallagher, A., Cao, L., Luo, J., Han, J. (2010). The wisdom of social multimedia: Using flickr for prediction and forecast. In *Proceedings of the 18th ACM international conference on multimedia, MM '10* (pp. 1235–1244). New York, NY: ACM. <https://doi.org/10.1145/1873951.1874196>.
- Kinsella, E. (2016). What does tefaf 2016 art market report tell us about the global art trade? <https://news.artnet.com/market/tefaf-2016-art-market-report-443615>. November 3, 2016.
- Leavitt, A., Burchard, E., Fisher, D., Gilbert, S. (2009). The influentials: New approaches for analyzing influence on twitter, technical report, web ecology project. <http://www.webecologyproject.org/2009/09/analyzing-influence-on-twitter/>.
- Liu, L., Wu, J., Li, P., & Li, Q. (2015). A social-media-based approach to predicting stock comovement. *Expert Systems with Applications*, 42(8), 3893–3901. <https://doi.org/10.1016/j.eswa.2014.12.049>.
- Liu, Y., Kliman-Silver, C., Mislove, A. (2014). The tweets they are a-changin: Evolution of twitter users and behavior. In *International AAAI conference on web and social media*.

- Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research*, 24(1), 146–163. <https://doi.org/10.1287/isre.1120.0462>.
- Degen, N. (Ed.). (2013). *The market, whitechapel documents of contemporary art*. Cambridge: The MIT Press.
- Oliveira, N., Cortez, P., Areal, N. (2013). On the predictability of stock market behavior using stocktwits sentiment and posting volume. In L. Correia, L. P. Reis & J. Cascalho (Eds.), *Progress in artificial intelligence: 16th Portuguese conference on artificial intelligence, EPIA 2013, Angra do Heroísmo, Azores, Portugal, Proceedings, 9–12 September 2013* (pp. 355–365). Berlin Heidelberg: Springer.
- Peetz, M. H. (2015). Time-aware online reputation analysis. Phd thesis, University of Amsterdam.
- Periferic Biennial. (2008). Interview with Marek Claassen, the director of artifacts.net by Zsuzsa Laszlo. <https://perifericbiennial.wordpress.com/2008/10/20/interview-with-with-marek-claassen-the-director-of-artifactsnet/>. November 3, 2016.
- Rehurek, R., Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks* (pp. 45–50).
- Roesslein, J. (2016). Tweepy version 3.5.0. <http://www.tweepy.org/>. November 5, 2016.
- Smailovic, J., Grčar, M., Lavrač, N., Žnidaršič, M. (2013). Predictive sentiment analysis of tweets: A stock market application. In A. Holzinger, & G. Pasi (Eds.), *Human–computer interaction and knowledge discovery in complex, unstructured, big data: Third international workshop, HCI-KDD 2013, Held at SouthCHI 2013, Maribor, Slovenia, Proceedings, July 1–3 2013* (pp. 77–88). Berlin, Heidelberg: Springer.
- Statista. (2016). Most famous social network sites worldwide as of september 2016, ranked by number of active users (in millions). <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>. November 3, 2016.
- Thompson, D. (2010). *The \$12 million stuffed shark: The curious economics of contemporary art*. New York: St. Martin's Griffin.
- Thornton, S. (2008). *Seven days in the art world*. New York: W. W. Norton & Company.
- Tucker, C. S. (2015). Quantifying product favorability and extracting notable product features using large scale social media data. *Journal of Computing and Information Science in Engineering*, 15, 031003.
- Twitter Inc. (2016a). Api rate limits. <https://dev.twitter.com/rest/public/rate-limiting>. November 5, 2016.
- Twitter Inc. (2016b). Get statuses/firehose. <https://dev.twitter.com/streaming/reference/get/statuses/firehose>. November 5, 2016.
- Twitter Inc. (2016c). Get statuses/sample. <https://dev.twitter.com/rest/public/search>. November 5, 2016.
- Twitter Inc. (2016d). The search api. <https://dev.twitter.com/rest/public/search>. November 5, 2016.
- Twitter Inc. (2016e). Tweets. <https://dev.twitter.com/overview/api/tweets>. November 5, 2016.
- Twitter Inc. (2016f) Working with timelines. <https://dev.twitter.com/rest/public/timelines>. November 5, 2016.
- Velthuis, O. (2007). *Talking prices: Symbolic meanings of prices on the market for contemporary art* (4th ed.). Princeton: Princeton University Press.
- Weng, J., Lim, E. P., Jiang, J., He, Q. (2010). Twitterrank: Finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on web search and data mining, WSDM '10* (pp. 261–270). New York, NY: ACM. <https://doi.org/10.1145/1718487.1718520>.
- Wikipedia. (2016). English wikipedia dump on 01/10/2016. <https://dumps.wikimedia.org/enwiki/20161001/>. October 1, 2016.
- Yang, C. C., Yang, H., Jiang, L., Zhang, M. (2012). Social media mining for drug safety signal detection. In *Proceedings of the 2012 international workshop on smart health and wellbeing, SHB '12* (pp. 33–40). New York, NY: ACM. <https://doi.org/10.1145/2389707.2389714>.
- Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. 1. Social Media research and applications 2. theory and applications of social networks. *Decision Support Systems*, 55(4), 919–926. <https://doi.org/10.1016/j.dss.2012.12.028>.
- Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E. P., Yan, H., Li, X. (2011). Comparing twitter and traditional media using topic models. In *Proceedings of the 33rd European conference on advances in information retrieval, ECIR'11* (pp 338–349). Berlin, Heidelberg: Springer.