

How to exploit social media data to evaluate performing arts: an empirical application at La Scala Opera House

Deborah Agostino (*), Michela Arnaboldi (*), Marco Brambilla (**), Beatrice Cassottana (*),
Riccardo Volonterio (**)

Politecnico di Milano

Keywords: performance measurement, social media, performing arts

(*)

Politecnico di Milano

Dipartimento di Ingegneria Gestionale

Via Lambruschini 4b (20156) Milano

(**)

Politecnico di Milano

Dipartimento di Elettronica, Informazione e Bioingegneria

Via Ponzio 34/5 (20133) Milano

How to exploit social media data to evaluate performing arts: an empirical application at La Scala Opera House

1. INTRODUCTION

Performing arts play a crucial role within a city, contributing to enhance cultural development and city growth (Strom, 1999; Evans, 2005; Eizenberg and Cohen, 2015; Jung et al., 2015). It is recognized that performing arts, and cultural activities in general, render urban areas more attractive, facilitating the physical, social and economic wellbeing of a certain area (Griffiths, 1995; Strom, 1999; Withford, 2009).

In this context, the evaluation of performing arts has long been considered a priority in order to understand if and quantify how arts contribute to urban regeneration and city growth (Pugh and Wood, 2004; Evans, 2005; Wood, 2005; Andersson and Lundberg, 2013). This evaluation is relevant for the performing arts organizations to better understand their viability and competitiveness (Rentschler and Potter, 1996), as well as for policy makers who need to gain insights to better support politics of urban arts development (Strom, 1999; Ashley, 2014).

Several indicators and frameworks have been proposed to date in order to evaluate performing arts (e.g. Baumol and Bowen, 1992; Baden-Fuller et al., 2007; Radbourne et al., 2009; Turbine and Lurin, 2009; Johanson et al., 2014). These frameworks are usually based on the dichotomy between artistic and economic performance. While in the past the attention was catalysed by budgeting systems to monitor the efficient use of resources (Baumol and Bowen, 1992), hence considering economic resources a constraint to artistic performance, over the years multi-objectives performance measurement approaches have been developed to consider simultaneously economic and artistic performance (Bouder-Pailler, 1999; Towse, 2001). Two main problems emerge from extant evaluation models: 1) measures are usually calculated ex-post, without evaluating performance during the event, which could be beneficial to manage potential problems in real time; 2) data to compute these measures are collected from secondary sources (such as press) or from users' surveys when the purpose is to assess customer satisfaction. This can be a problem since those who attend an event do not always have the willingness or the possibility to fill questionnaires.

In this study, we address these two problems about performing arts measurement by exploring the role of social media in supporting the evaluation. Social media have widely developed over these recent years and extant literature recognizes two main social media features that are potentially useful in our context, namely real time data and user generated content (Price, 2006; Richardson, 2006). The

former refers to the possibility to collect real time data from users hence facilitating performance measurement during the event. The latter is related to the exploitation of users data published on social networks rather than surveys, increasing on the one hand the amount of potential respondents and, on the other hand, collecting insights without asking customers to fill interviewer-lead questionnaires, which might cause biased judgements.

Given this context, our objective is to understand if and how social media can contribute to performing arts evaluation. This is associated with three main research questions:

- what are the KPI (Key Performance Indicators) that can be derived from social media, specifically Twitter, in order to evaluate events;
- how do social media data differ from information retrieved from traditional sources (i.e. press review);
- what are the main benefits and drawback that arise from the exploitation of social media data for events evaluation.

These questions were empirically addressed by analysing one event held at the Opera House La Scala in Milan during the season 2014-2015: Aida. For this event, social media data as well as press review data were collected and analysed. Twitter was the selected social media for the analysis, while the press review exploration was considered necessary in order to understand the differences that social media data can provide with respect to traditional sources. The results enhanced the definition of a set of KPIs, social media derived, to support performing arts evaluation, but they also facilitated a more general discussion about benefits and drawback about the usage of social media data to evaluate events in comparison with traditional data sources.

This paper contributes to the practitioner field, by proposing a novel approach that integrates data from social media with traditional measures (Evans, 2000; Hume and Mort, 2008; Radbourne et al., 2009; Johanson et al., 2014) from the press review in order to evaluate events. Moreover, it also identifies a set of problems and benefits managers should consider when exploiting social media data to events evaluation.

The rest of the paper is structured as follows: first extant literature on measurement and evaluation of performing arts is discussed, with a specific focus on the available indicators and frameworks [2]; then social media potentialities are analysed together with available studies about the role of social media in performing arts management [3]. The research methodology is described [4] followed by the discussion of the results in terms of press review and social media analysis together with the proposed set of KPIs [5]. Finally, this study highlights the potentialities of social media applied to event measurement, which can serve both managers and academics [6].

2. EVALUATING PERFORMING ARTS: INDICATORS AND DATA SOURCES

The issue of evaluating performing arts, such as theatre, opera, music and dance, has long received attention, given its importance for performing arts organizations, local administrations and policy makers. Organizations need to understand the viability and competitiveness of arts events (Rentschler and Potter, 1996), while local administrations and policy makers are mainly interested in the expected costs and benefits that an art event will generate for the local community in order to support future urban arts policies (Strom, 1999; Mills and Rosentraub, 2013; Gezici and Er, 2014).

Earlier studies about the evaluation of performing arts date back to the early 1990s (e.g. Baumol and Bowen, 1992), which consider performance measurement a constraint that limits the creativity of the arts field (e.g. Lampel et al., 2000). This view gave rise to a contrast between the artistic dimension, to be enhanced and promoted, and the commercial (revenues) dimension, which simply limited the possible range of actions by setting economic boundaries (Baden-Fuller et al., 2007). Moving from this consideration, initial studies on performing arts evaluation posed a great attention on the economic, and particularly the budget, dimension of performance. The budgeting system was considered a useful approach to evaluate whether resources, mainly personnel, were employed in an efficient way (Baumol and Bowen, 1992; Felton, 1994; Christiansen and Skærbæk, 1997).

In later years, both economic and artistic performance were considered relevant within an evaluation system. In this respect, multi-objectives performance measurement approaches were developed and applied to performing arts and opera houses in order to monitor economics as well as artistic performance (Bouder-Pailler, 1999; Towse, 2001). For example, Gilhespy (1999) proposed measures to quantify artistic excellence, innovation, social cohesion, public attendance and financial objectives. Turbine and Lurin (2009) identified six main categories of performance: artistic achievement, audit satisfaction, funder satisfaction, personnel satisfaction, financial management, growth and competition, and image and reputation. In a similar vein, several other studies proposed multiple performance measurement dimensions. Although with different labels, they include both financial data and measures about the quality of artistic performance, and especially the satisfaction of the audience, mainly evaluated through surveys or critical reviews from experts (e.g. Evans, 2000; Hume and Mort, 2008; Radbourne et al., 2009; Johanson et al., 2014). These later group of studies gave rise to a mix of financial and non-financial indicators to support the evaluation of the quality of live performance, while at the same time considering financial aspects (Turbine and Laurin, 2009)

This literature illuminates academics and practitioners on the measurement systems to quantify the feasibility of performing arts. Yet, two main limitations can be identified with reference to these available contributions. The first limitation is related to the real time evaluation of the performing arts events. While the majority of the studies has been concerned with how to quantify performance

after the event and its feasibility in terms of budget before its implementation, there is no evidence on performance evaluation while the event is taking place. This aspect could be useful for organizers in order to introduce corrective actions and increase the possibility of success of the event itself.

The second limitation concerns the difficulty in obtaining data about audience perception on the event. Previous studies showed that customer satisfaction surveys can be adopted to gain evidence on the social perception of the event. However, obtaining answers to questionnaires during or after the event is far from being straightforward and this can compromise the response rate.

In order to address these two main limitations, this study aims to understand if and how social media can contribute to performing arts evaluation.

3. SOCIAL MEDIA AND PERFORMING ARTS

Social media are receiving a continuously growing attention in a variety of fields, including culture and performing arts (e.g. Hausmann and Poellmann, 2013; Colombo, 2014; Ribeiro and Peixoto, 2014). They can be defined as a set of online tools centred on user generated content and social interactions (Kaplan and Haenlein, 2010). Within the umbrella term of social media, several applications are included, such as social networks, blogs, wikis, photo sharing, video sharing platforms or virtual games (Kaplan and Haenlein, 2010). Each of these applications has its own features. For example, social networks (e.g. Facebook, My Space, LinkedIn) allow users to contact a certain number of friends and share texts, photos or files with them. Wikies, like Wikipedia, are networks that favour the development and sharing of knowledge among a group of users. Blogs are the oldest example of social media. They represent personal web pages that allow users to share personal information and opinions that currently find a new development in microblogging, Twitter. Photo sharing or video sharing platforms, such as Instagram, Flickr or YouTube, give users the possibility to interact with each other by posting videos or photos.

Albeit this variety among social media, they are all characterized by two main features: user generated content and real time interactions. The first distinctive characteristic is that users can be active creators of content rather than passively receiving information. Social media applications are based on Web 2.0, also known as the 'Read-Write Web' (Price, 2006; Richardson, 2006), as it enables members of the public to contribute actively and shape the website content. The second distinctive element concerns real-time communication between parties, which shifts the dialogue from offline communication to online instantaneous communication. This allows several users to discuss and share information simultaneously, so that a network of relationships is created more easily.

Potential contributions of social media to performing arts have been recently recognized (e.g. Hausmann and Poellmann, 2013), although the literature is almost silent with respect to how social tools can support performance evaluation. However, some insights can be indirectly gained from the more general literature about the social media role in performing arts organizations. Two issues have emerged as relevant.

The first issue is related to the recognition of the significant influence of social media in engaging the audience in arts events (Bennett, 2012; Hudson, 2015), indirectly suggesting that the quantification of quality performance can be more easily obtained by exploiting social media data rather than filling questionnaires. In this respect, it has been recognized (Hausmann and Poellmann, 2013) that social media can contribute to understand the reputation of an event and its organizer, although no specific measures have been proposed.

The second issue concerns the possibility to interact in real time with the audience, providing rationale to develop a measurement system that focuses on an evaluation while the event is taking place rather than ex-ante or ex-post. For example, real time interactions have emerged being enhanced by exploiting the features of augmented reality and gamification (e.g. de los Rios et al., 2014) or by relying on crowd-sensing applications (e.g. Kanter et al., 2014).

4. METHODOLOGY

In order to understand if and how social media can contribute to performing arts evaluation, we analysed one opera at La Scala Opera House: *Aida*. It was on stage from February 15th until March 15th 2015. For this opera we performed a qualitative and quantitative analysis of both press review and social media.

The selected social media is represented by Twitter given its widespread diffusion at both national and international level (Nielsen, 2012; Cosenza, 2014), which allowed to gain comments to the event potentially on a world scale base. The decision to include the press review analysis is justified by the need to understand how social media data differ from information retrieved from traditional sources. We performed a parallel analysis of the press review and Twitter. As far as the press review analysis is concerned, we retrieved articles from various types of communication media: newspapers, webpages, radio and television. This material has been received directly from the management of La Scala, which systematically collects all the information of interest about its performing events. The only criterion for a source to be included in our analysis is that it has to discuss about the abovementioned opera. The collected sources were either local or national –and a minor part was international- and the articles are predominantly published by generalist periodicals, with only a few

being published by specialised periodicals. Even though no limitations on the time horizon have been imposed, the sources that talk about the opera typically date to a time period close to the show dates. With respect to the Aida, the related sources are all included in a time window that goes from about one month before the first performance of the opera, to the date of the last performance. In conclusion, the analysed dataset of press reviews contains 57 sources about Aida.

The software NVivo was adopted to conduct the quantitative and the qualitative analysis. The quantitative analysis consists of the quantification of the following features:

1. *Timing*: number of sources published before, during and after the representation dates,
2. *Mentions*: number of sources that mention the main artists, the director and the orchestra conductor,
3. *Visibility*: number of regional, national and international sources,
4. *Type*: number of periodical, web, radio and television sources,
5. *Topic*: number of sources entirely dedicated to the opera (in opposition to the sources that deal with different topics. Many newspaper pages, for example, advertise all the events happening in Milan, without focusing exclusively on the opera).

Some notes need to be done about the *visibility* of the source and the *topic* of the article. For what concerns *visibility*, we divided regional sources into those belonging to the region Lombardy, the region of Milan, and those that pertain to particular regions outside Lombardy. For example, “La Gazzetta del Mezzogiorno” or “Il Tirreno” are regional and non-Lombardy newspapers. The international sources are those sources that are written in a language other than Italian. The detected languages are English, Spanish, French and German. The *topic* of the article can be detected for periodical articles only. Therefore, in order to count the number of press sources entirely dedicated to the opera analysed, we counted the number of periodical articles containing the name of the opera in the title and we have added the number of radio or TV sources with a length of at least 10 minutes. The qualitative analysis is concerned with automatically identifying the content and the “sentiment” of the press review, through individualizing the recurrent words and studying the context in which these words and other significant keywords appear. In particular, we were able to draw conclusions on the number of sources that express a positive and a negative sentiment about:

6. *Sentiment_Opera*: number of sources that express a positive and a negative sentiment about the opera as a whole,
7. *Sentiment_Music & Artists*: number of sources that express a positive and a negative sentiment about the music and the artists,
8. *Sentiment_Costumes & Scenography*: number of sources that express a positive and a negative sentiment about the costumes and the scenography.

Since the number of sources that express a sentiment about the opera have been individualised through a limited number of keywords (i.e. general public, Peter Stein, scenography, costumes, cast, Zubin Mehta, Carlo Colombara, Anita Rachvelishvili), the above mentioned numbers can underestimate the real number of articles expressing either a positive or a negative sentiment. The search by keywords allowed us to reflect on episodes that have been linked to the opera. Indeed by individualizing those words that are most often associated with the opera, we were able to identify contingencies that have somehow affected the “image” of the analysed opera on the press.

Finally, in order to test the efficacy of the results obtained from Nvivo, two of the involved researchers performed a manual analysis on a sample of press sources and compared their results with those derived from the software NVivo. This manual analysis provided the same result of the automatic one, hence corroborating the validity of the automated procedure adopted.

As far as the Twitter analysis is concerned, we extracted and downloaded all those tweets containing the name of the opera, the name of the conductor and the name of the three main artists. Table 1 shows the detailed keywords adopted (in parenthesis we indicate the role of the artist, even though it has not been adopted as a keyword).

Name of the opera	Aida
Conductor	Zubin Mehta
Cast	Carlo Colombara (the King) Anita Rachvelishvili (Amneris) Kristin Lewis (Aida)

Table 1: list of the Keywords

These tweets were downloaded starting from three days before the first performance until one week after the last performance. During this time-period, in addition to the tweets containing the abovementioned keywords, all the tweets containing the words “La Scala” were downloaded.

Out of the total amount of downloaded tweets, we selected only the ones written in Italian or English in order to facilitate the automatic analysis of the tweets through a machine learning approach. We finally divided the extracted tweets into two main categories: (i) the tweets related to the Opera house, La Scala, but not mentioning the specific operas, and (ii) tweets related to the specific opera of Aida. We finally focused our analysis on those tweets related to the second category, meaning those tweets concerning the opera analysed (Table 2).

Total tweets downloaded	5077
Usable tweets (written in Italian or English)	2296
Relevant tweets (concerning the specific opera)	841

Table 2: details about the amount of Tweets collected and analyzed

Similarly to the analysis of the press review, we conducted both a quantitative and a qualitative analysis. In the case of Twitter the quantitative analysis consists of the quantification of the following features:

1. *Timing*: number of tweets published before, during and after the representation date,
2. *Mentions*: number of tweets that mention the main artists, the director and the orchestra conductor,
3. *Language*: number of tweets written in Italian and in English.

Afterwards, a sentiment analysis was conducted adopting an automatic “tag” based on different categories of classification. In particular, as seen in the case of the press review, we were able to detect the following:

4. *Sentiment_Opera*: number of sources that express a positive and a negative sentiment about the opera as a whole,
5. *Sentiment_Music & Artists*: number of sources that express a positive and a negative sentiment about the music and the artists,
6. *Sentiment_Costumes & Scenography*: number of sources that express a positive and a negative sentiment about the costumes and the scenography.

In addition, the following characteristics about the users were extracted and analysed:

- a. *Level of activity*: average number of tweets posted per month,
- b. *Level of exposure*: number of participated lists,
- c. *Level of influence*: number of followers.

These characteristics helped us ponder the judgements expressed by the users. In this way, we could compute the *average sentiment about the opera* by assigning to each judgement (measured in a scale from -5, for very negative judgements, to +5, for very positive judgements) a weight expressed as the level of relevance of the user.

Moreover, the characteristics of the users helped us to understand the key influencers of the network, indication that is relevant when designing a marketing campaign. Thanks to the parallel between the press review analysis and Twitter analysis, we were able to compare the two datasets, identifying the main differences between a traditional dataset and a social media source. In particular, we focused on the information added by Twitter analysis deriving a set of KPIs that can be calculated starting from

social media data, supporting a more general discussion about problems and opportunities to exploit social media data to events evaluation.

5. RESULTS

This section discusses the results obtained from the analysis of the press review and Twitter. It is divided into three main paragraphs. The first paragraph discusses the results obtained through the analysis of the press review [5.1], the second discusses the results obtained from the Tweets analysis [5.2] and the third proposed a more general set of KPIs, social media derived, together with a discussion about potentialities and problems about the usage of social media data to evaluate performing arts [5.3].

5.1 Results from the press review analysis

The retrieved data from the press review are showed in Table 3 (in parenthesis the numbers are expressed as percentage points on the total number of retrieved reviews). Starting from a set of 57 reviews we retrieved a set of information about timing, mentions, sentiment, visibility, type and topic, which supported the Aida evaluation exploiting traditional data sources.

The first type of information is related to the *timing* of the press review. The press reviews were concentrated during the performance dates, as one could expect. Instead, it is interesting to observe that the number of reviews published before the first performance is greater than the number of those published after the last performance. We tried to explain this result considering the needs of the different stakeholders that the publishing sources want to satisfy. In particular, we hypothesised that part of the reviews act as an advertisement of the event, and it is therefore more effective to publish them before the first representation. The effectiveness of the advertisement is one of the major interest of La Scala, which, in this case, is the direct stakeholder. A second possible explanation lies in the fact that the publishing sources (i.e. newspaper, web, radio and TV) are interested in having the broadest possible audience, which is the direct stakeholder in this case, and they do so by anticipating, rather than following, the review of the events.

The second type of information is related to the number of *mentions*, which were particularly high with reference to the orchestra conductor and the director, with, respectively, 72% and 74% of the press contributions mentioning them. This result provides a clear indication on the leverages used by La Scala, the opera house in question, in order to have not only a good quality of the opera, but also a great resonance of the event on the traditional media.

The third type of information concerns the *sentiment* about the Aida. The sentiment aroused by the press review is generally positive although a few amount of sources provides judgement about the

opera as a whole (20 reviews out of 57), or the casts or music (5 reviews out of 57), costumes and scenography (13 reviews out of 57). This result further confirms our hypothesis that press reviews are mainly interested in advertising the event rather than providing an evaluation. Furthermore, by analysing the time distribution of the positive and negative sentiments for the Aida, we observed an interesting trend: the positive judgements happen in time before the negative ones. This result can be explained by the high expectation before the event, which led to positive judgement, followed by criticisms after the occurrence of the artistic performance.

Lastly, *visibility*, *type* and *topic* offer an indication to understand the characteristics and the relevance of the press media where the operas appear. In particular, most of the reviews are published on national newspapers such as “Il corriere della Sera” or “La Repubblica” and are not entirely related to the opera. This means that most of the articles are paragraphs at the side of the page with an informative and advertising scope. Similarly, most of the radio and TV broadcasts (which impacted for less than the 25% of the total reviews) dedicates no more than ten minutes to the operas.

(0) Quantity	Number of retrieved reviews	57
(1) Timing	Number of reviews published before the first performance	19 (0.33)
	Number of reviews published during the first and last performance	38 (0.66)
	Number of reviews published after the last performance	0
(2) Mentions	Number of reviews mentioning Carlo Colombara	8 (0.14)
	Number of reviews mentioning Anita Rachvelishvili	16 (0.28)
	Number of reviews mentioning Kristin Lewis	22 (0.38)
	Number of reviews mentioning Peter	41 (0.72)
	Number of reviews mentioning Zubin Mehta	42 (0.74)
(3) Visibility	Number of reviews published in regional-Lombard sources	16 (0.28)
	Number of reviews published in regional-non-Lombard sources	2 (0.03)
	Number of reviews published in national sources	39 (0.68)
	Number of reviews published in international sources	0
(4) Type	Number of reviews published on newspapers	42 (0.74)
	Number of reviews published on the web	1 (0.02)
	Number of reviews published on the radio	7 (0.12)
	Number of reviews published on the TV	7 (0.12)
(5) Topic	Number of reviews entirely dedicated to the opera	27 (0.47)

(6) Sentiment about the opera as a whole	Number of reviews expressing a positive sentiment	17 (0.30)
	Number of reviews expressing a negative sentiment	3 (0.05)
(7) Sentiment about the music and the cast	Number of reviews expressing a positive sentiment	4 (0.07)
	Number of reviews expressing a negative sentiment	1 (0.02)
(8) Sentiment about the costumes and the scenography	Number of reviews expressing a positive sentiment	7 (0.12)
	Number of reviews expressing a negative sentiment	6 (0.10)

Table 3: results from the press review analysis

Finally, starting from a press review, it is possible to calculate the number of recurring words. By performing this analysis we were able to identify contingent facts that have been associated with the opera, providing an additional information about the context of the opera itself.

With respect to Aida, the words “Stein”, “scena” (scene), “regista” (director), “regia” (direction) and “Zeffirelli” have been the most frequently used ones (with an occurrence of respectively 5, 5, 7, 5 and 10 reviews), meaning that the scenography of Aida has been largely discussed because of its originality, and especially in comparison with the more classic interpretation of the director Zeffirelli. To summarize: the press review analysis provides a set of information about timing, mention, visibility, type and sentiment of the opera with the main limitation that the majority of the reviews have advertisement or informative purposes, making it difficult to evaluate the event.

5.2 Results from the Twitter analysis

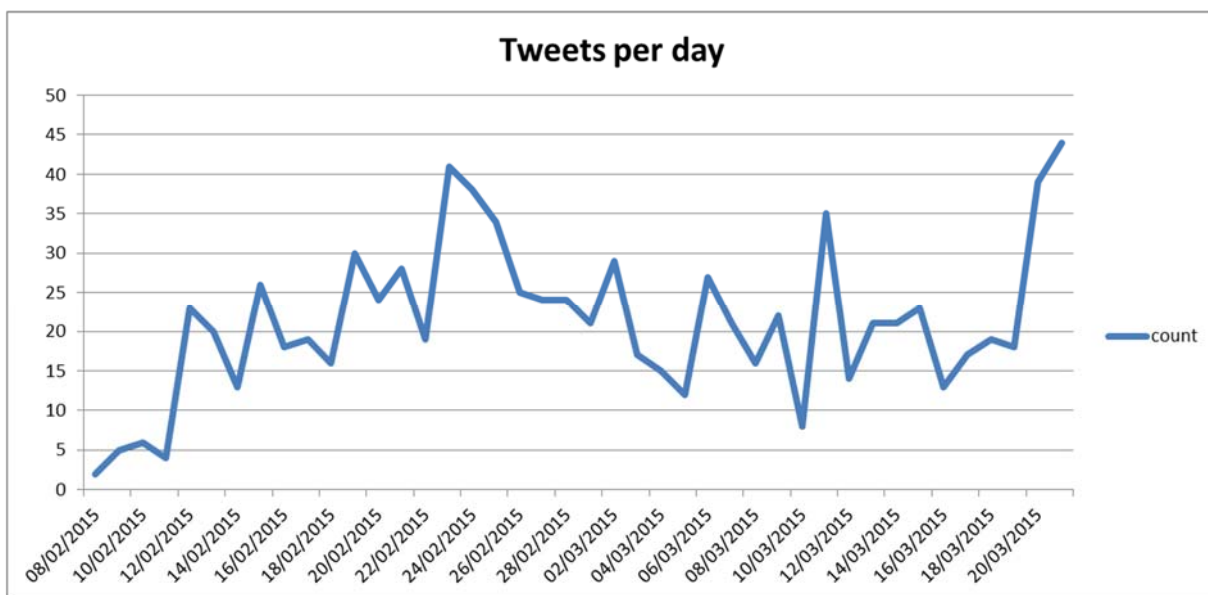
This section discusses the results, showed in Table 4, obtained from the Twitter analysis. Also in this case, we collected and quantify the type of information that can be extracted from social media, with the final aim to understand if i) the type of information differ from traditional sources (i.e. press review) and, when the type of the retrieved information is the same, ii) if the judgement is in accordance or in dissonance .

(0) Quantity	Number of retrieved tweets	891
(1) Timing	Number of tweets published before the first performance	73 (0.08)
	Number of tweets published during the first and last performance	695 (0.78)
	Number of tweets published after the last performance	173 (0.19)
(2) Mentions	Number of tweets mentioning Carlo Colombara	43 (0.05)

	Number of tweets mentioning Anita Rachvelishvili	61 (0.07)
	Number of tweets mentioning Kristin Lewis	48 (0.05)
	Number of tweets mentioning Peter Stein	17 (0.02)
	Number of tweets mentioning Zubin Mehta	199 (0.22)
(3) Language	Number of tweets published in Italian	149 (0.17)
	Number of tweets published in English	742 (0.83)
(4) Sentiment about the opera as a whole	Number of tweets expressing a positive sentiment	324 (0.36)
	Number of tweets expressing a negative sentiment	140 (0.16)
(5) Sentiment about the music and the artists	Number of tweets expressing a positive sentiment	91 (0.10)
	Number of tweets expressing a negative sentiment	27 (0.03)
(6) Sentiment about the costumes and the scenography	Number of tweets expressing a positive sentiment	1 (0.00)
	Number of tweets expressing a negative sentiment	2 (0.00)
(10.b) Average sentiment about the opera	Weighted average of the sentiments on the opera as a whole	1.05

Table 4: results from the Twitter analysis

As for the press review, the first type of retrieved information is related to the *timing* of the tweets, intended as the temporal distribution of the tweets over the time horizon (see Figure 1).



With respect to the Aida, the majority of them were focused between the first and the last performance (occurred respectively at February 15th and at March 15th), with more comments after the last

performance rather than before the first performance. Figure 1 shows this trend with an amount of tweets lower than 25 before the first performance, which achieve the number of 40 during the last representation of February, occurred at February 24th, and then increase again after the last performance in March. This result goes opposite with respect to the finding from the press review, where comments occurred mainly before and during the performance rather than after. This finding about the timing of the tweets provides evidence that the audience used to provide comments about the opera after it is performed rather than talking about the event before its occurrence.

The second type of retrieved information concerns the *mentions*, which were identified by considering the related twitter features. We found that a low percentage of tweets mentioned the cast (the sum of the tweets related to the artists equals the 21% of the overall amount of tweets), while the attention has been mainly catalysed by the conductor Zubin Mehta who obtained alone 199 tweets over a total amount of 891 (22%). A possible explanation of this is represented by his reputation with respect to the popularity of the cast. To confirm this hypothesis, we considered the top mentions related to La Scala during the period of interest. We found the name of Roberto Bolle as the second top mention during the analysed period with an amount of tweets mentioning its name equal to 1305 (the first mention was the name of the Opera House La Scala that achieved 10307 mentions). This result about Roberto Bolle confirmed that even without having any performance on stage during that time period, the popularity of the artist strongly influence the value about the mentions when the popularity is high. However, we need to specify that the low number of mentions for Peter Stein, and, consequently, the low number of tweets expressing a sentiment about the costumes and the scenography (the keywords that we have used to retrieve the comments about the scenography and the costumes include “Peter Stein”), can be due to the methodology that we have used to download the tweets. Indeed the names of the directors have not been used as keywords to download the tweets. The third type of retrieved information is related to the *language* of tweets. We found that the majority of the posts were in English rather than in Italian with 742 English tweets over a total amount of 891 (89%). This is an interesting insight about the audience of the operas, providing information about who talk about an event. This result of a predominance of English tweets can be associated to the fact that the resonance of the event was international, at least for the general audience, even though the location of the event was Italy.

The fourth type of information retrieved concerns the *sentiment* of the tweets, which provides information about the perception of the opera. By calculating the sentiment we obtained mainly positive comments for Aida as a whole (36% of comments were positive against a 16% that were negative), but with a different perception about the music and the cast on the one hand, and the costumes and scenography on the other hand. While the music and cast were particularly appreciated

(10% of the tweets were positive against a 3% that was negative), the latter were almost not considered given that we had only 3 tweets providing comments about this issue. In general, the total sentiment about Aida was positive with a polarity of 1.05. Finally, we also calculated the sentiment per each of the top mentions (Table 5) obtaining a sentiment which is predominantly positive or negative rather than neutral. For example, Zubin Mehta who achieved the higher amount of mentions (199) received 46 positive tweets and 16 negative.

Name	Total tweets containing the name	Tweets with a positive sentiment	Tweets with a negative sentiment	Tweets with a neutral sentiment
Zubin Mehta	199	46 (0.23)	16 (0.08)	137 (0.69)
Peter Stein	17	1 (0.06)	2 (0.12)	14 (0.82)
Kristin Lewis	48	16 (0.33)	5 (0.10)	27 (0.56)
Anita Rachvelishvili	61	20 (0.33)	5 (0.08)	36 (0.56)
Carlo Colombara	43	9 (0.21)	1 (0.02)	33 (0.77)

Table 5: Sentiment per each top mention

5.3 The contribution of Twitter for events evaluation

In order to understand if and how social media can support the evaluation of operas, we analysed analogies and differences between the results derived from the analysis of the press review and Twitter (see Table 6). This approach allowed us to highlight how social media differ from traditional sources for performing arts evaluation, and to propose a set of KPIs for events evaluation based on the exploitation of these social media data.

	Press review	Twitter
Quantity	Small amount (57)	Large dataset (891)
Timing	Mainly before the event	Mainly after the event
Mention	Focused on the director and the conductor	Focused on the cast
Sentiment	Limited and mainly neutral	Diffused and with several negative comments
Language	Mainly Italian	Mainly English

Table 6: comparison between social media data and press review data

By comparing the results from the press review and twitter analysis we found some similarities and differences related to the amount of available data, time horizon, mentions, topic and sentiment.

The discrepancy in terms of the *quantity* of the available data (number of reviews vs number of tweets) shows how Twitter, and social media more in general, can help to have a better idea about the audience perception. Indeed the amount of data gathered through Twitter is significantly higher than the total number of press reviews (891 tweets vs 57 press reviews) that talk about the operas. This difference highlights the first benefits to exploit social media data: the possibility to rely on a large set of comments and judgements that could not be achieved by relying on traditional data sources.

The second difference is related to the *timing* of the available data. The tweets are posted predominantly after the first representation of the opera, while reviews are published predominantly before the last representation. The difference is mainly due to the number of comments dated before the first representation and after the last one. While the press review starts to talk about the opera long before its representation dates and stop to talk about it with the last representation, instead tweets are posted starting from the first representation date until at least one week after the last representation. This discrepancy might be due to the advertising function of the press review commented before and underlines the need of the users to talk about the opera. Moreover, even though social media allowed to collect data in real time, this cannot occur during operas given that the mobile phone is forbidden during representations.

The third difference is related to the number of *mentions*, intended as the frequency of a set of keywords that occurred in both press review and tweets. In general the press review cites the authors, the director and the conductor more than the tweets do. In particular, while the most cited ones on the press review are the conductor and the director, for Twitter the artists of the cast are slightly more popular than the others. This difference about the type and amount of mention can still be associated with the different purpose of the two data sources: the press review, with its mainly advertisement purpose used to cite the director and the conductor of the opera, while tweets, deriving from the general public, used to mention the cast.

Another difference is related to the *sentiment* about the opera. Even though in both the data sources the comments are generally positive, Twitter presents a greater number of negative judgements in comparison to the press further suggesting that the general public used to comment about the opera, also with negative judgements.

Finally, by comparing the *visibility* of the press sources and the *language* of the tweets, it is possible to identify a significant difference in terms of the nationality of the audience of the two kinds of media. On the one hand, the press review is predominantly aimed at a national –Italian- audience, on the other hand, Twitter includes a more international public, since English is the main means of communication.

Starting from the available information that can be extracted from Twitter, we developed a set of KPIs that can be adopted to evaluate performing arts, and events more in general (Table 7).

KPI	Metric
Impact of the opera	Nr of tweet mentioning the opera/total number of tweet about La Scala
Reputation	Sentiment about the Opera as a whole
	Sentiment about the conductor and the cast
	Sentiment about set and costumes
Word of mouth	Average Nr of retweet per tweet
Influencer	Nr of tweet per account

Table 7: KPIs for performing arts evaluation

The measures are relevant, not only to evaluate a specific event, but the publicness of Twitter information allowed to compare these data with those of other events and other Opera Houses other than the individual one. The measures we propose include the following: impact, reputation, word of mouth and influencer.

The measure about the *impact* provides information about the relevance of the event (the opera in our empirical analysis) for the organizer. It is calculated as the ratio between the amount of tweets related to the detailed event (i.e. Aida) with respect to the total tweets related to the organizer (i.e. La Scala). The higher is this value, the higher is the importance of the event for the twitter audience since it means that the majority of people who twitted about La Scala were also interested in that specific event.

The second indicator concerns the *reputation* of the event, which is evaluated exploiting the sentiment of the event itself. It can be calculated by adopting three different metrics: the first calculates the sentiment of the event as a whole, the second instead calculated the sentiment with respect to specific features of the event (in our empirical analysis represented by music and cast, and scene and costumes). Through this measure evidence is provided about the perception of the event by twitter users, providing suggestions on the most favourite or worst aspects of the events. A comment is required for what concern how the sentiment is calculated: depending on how tweets are classified, results can vary. We adopted an automatic tag, but the exploitation of different approaches could led to different results.

The third indicator is related to the *word of mouth*, calculated as the average number of retweet per tweet, which provides insights about the potential diffusion of the tweet in the social community. Indeed the higher is the amount of retweet per tweet, the higher is the amount of people that can

access the comment. This indicator is particularly useful to monitor the trend of tweets with a negative sentiment since their widespread diffusion can be higher.

Finally, the fourth indicator concerns the *influencer*, which represents those twitter users that mainly twitted about the opera. It is calculated by considering those accounts with the highest number of tweets about the specific event. This value provide relevant information about who can influence the perception of the event itself: thinking at our empirical analysis, the top influencer for both the operas is represented by La Scala suggesting that it intention to promote and advertise the operas generated a higher amount of tweets than comments by users. If the intention of the evaluation is to receive a judgement about the perception of the event, those tweets posted by the organizers need to be removed.

6. CONCLUSION

This paper aimed at understanding if and how social media can contribute to performing arts evaluation in terms of: identification of a set of KPI (Key Performance Indicators) that can be derived from social media in order to evaluate events; analysis of the added value of social media data from traditional sources (i.e. press review); and benefits and drawback associated with social media data. We relied on the Opera Aida on stage at La Scala during the season 2014-2015 and performed both a press review and social media analysis in order to understand the main differences between these two different data sources and therefore better capture benefits and drawback of social media data. Our results highlight the contribution of social media data to perform events evaluation.

First, the amount of comments that can be retrieved from social media is much higher that those from the press review analysis, providing a larger data base to support the judgement of an event and therefore having a better understanding about the reputation of the event itself.

Second, the timing of the tweets occur mainly after the performance facilitating the collection of comments about the perception of the opera. This is connected with the different users and purposes of social media with respect to the press review. On the one hand, the “user” is different: while press reviews provide a description of the event by a journalist, the social media contains information by the general public who might had attended the event or not. At the same time, also the purpose of the comment is different: while press review is mainly intended at inform the public about the event, and in some cases advertise the event, social media mainly provides discussion around a certain event. These two difference further corroborate the relevance to exploit social media data to understand the perception of an event by the public.

Third, social media allowed the analysis of the sentiment about each tweet, and therefore about the opera and the analysis of the influencer within the network. These features can be exploited by calculating a set of KPI, in terms of impact, sentiment, word of mouth and influencers, that relies on twitter data in order to complement traditional press review analysis.

These results suggest that social media data do differ from traditional data sources in terms of quantity, timing, mention, sentiment and language and they can be exploited to integrate traditional data sources analysis in order to understand and evaluate performing arts.

However, the computation of social media analysis led to the identification of a set of problems, which suggests some warnings about the exploitation of social media data need that need to be considered. First, the identification of the key words: depending on the keywords selected, the amount and type of downloaded tweets can vary. Second, the setting of the time horizon: it influences the analysis since the discussion about an event can start or finish long before or long after the event itself, therefore influencing the available data set. Third, inclusion/exclusion criteria: depending on which type of tweets wanted to be included in the analysis, the evaluation of the event can change. For example, our analysis included Italian and English tweets only; the inclusion of other languages could have led to different judgements. Fourth, the approach to sentiment analysis: several different approaches are available to perform sentiment analysis. We relied on a machine learning approach that, starting from a manual tagging of the tweets, automatically proceeded to the classification of tweets. The adoption of a different algorithm again could have provide a different sentiment.

Finally, we want to synthesise the benefits of our analysis for practitioners and academics. At the practitioner level, these results can provide a practical guidance to Opera House managers to understand the advantages to exploit social media data to evaluate their events, integrating the available social media information with the traditional press review analysis. Given the general public that use to post on twitter, the content and sentiment of the tweets can provide a set of information larger and more reliable that the traditional press review with its mainly informative purposes. Second, we also suggest managers a set of KPIs that can be applied to evaluate events by relying on twitter data. At the academic level, we enlarged the available literature on performing arts evaluation by discussing how social media can provide benefits, but also their limitation, to the evaluation of the event.

Limitations of our paper need to be addressed. We relied on a single event only, the Aida. More robust conclusions could be achieved by performing this analysis to other events. This was our initial intention by setting the social media analysis emerged as a difficult path. The next steps will include a triangulation of these results with other operas at La Scala and with interviews with La Scala managers in order to get their comments about these findings.

References

- Andersson, T., and Lundberg, E. (2013), “Commensurability and sustainability. Triple impact assessment of a tourism event”, *Tourism Management*, 37, pp. 99–109.
- Ashley, A.J. (2014), “Negotiating risk in property-based arts economic development: Exploring the innovative but untimely development partnership between the Seattle Art Museum and Washington Mutual”, *Cities*, 37, pp. 92-103
- Bennett, L. (2012), “Patterns of listening through social media: online fan engagement with the live music experience”, *Social Semiotics*, 22(5), pp. 545-547.
- Baumol, W.J. and Bowen, W.G. (1992). On the Performing Arts: The Anatomy of Their Economic Problems, in Blaug, M.(ed.)*The Economics of the Arts*, Aldershot, Gregg Revivals,pp.218–226
- Bouder-Pailler, D. (1999), “A model for measuring the goals of theatre attendance”, *International Journal of Arts Management*, 1(2),4–15.
- Christiansen, J.K. and Skærbæk, P. (1997), “Implementing budgetary control in the performing arts: games in the organizational theatre”, *Management Accounting Research*, 8, pp. 405-438.
- Colombo, A. (2014), “Social Media As a Key Element for Cultural Global Events”, paper presented at XVIII ISA World Conference of Sociology, Yokohama, 13-19 July.
- Cosenza, V. (2014) World map of social networks, available at <http://vincos.it/world-map-of-social-networks/>
- De los Rios, S., Cabrera-Umpiérrez, M.F., Arredondo, M.T., Páramo, M., Baranski, B., Meis, J., Gerhard, M. Prados, B., Pérez, L., and del Mar Villafranca, M. (2014), “Using Augmented Reality and Social Media in Mobile Applications to Engage People on Cultural Sites”, In Stephanidis, C. and Antona, M. (Eds), *Universal Access in Human-Computer Interaction, Universal access to Information and Knowledge*, pp. 662-672.
- Eizenberg, E., and Cohen, N. (2015), “Reconstructing urban image through cultural flagship events: The case of Bat-Yam”, *Cities*, 42, pp. 54-62
- Evans, G. (2000), “Measure for measure: evaluating performance and the arts organisation”, *Studies in Cultures, Organizations and Societies*, 6(2), pp. 243-266
- Evans, G. (2005), “Measure for Measure: Evaluating the Evidence of Culture’s Contribution to Regeneration”, *Urban Studies*, 42(5/6), pp. 959-983
- Felton, M.V., (1994), “Evidence of the Existence of the Cost Disease in the Performing Arts”, *Journal of Cultural Economics*, 18, 93–112.
- Jung, T.H., Lee, J., Yap, M.H.T., and Ineson, E.M. (2015), “The role of stakeholder collaboration in culture-led urban regeneration: A case study of the Gwangju project, Korea”, *Cities*, 44, pp. 29-39

- Johanson, K., Glow, H., and Kershaw, A. (2014), “New modes of arts participation and the limits of cultural indicators for local government”, *Poetics*, 43, pp. 43-59.
- Gezici, F., and Er, S. (2014), “What has been left after hosting the Formula 1 Grand Prix in Istanbul?”, *Cities*, 41, pp. 44-53
- Gilhespy, I. (1999), “Measuring the Performance of Cultural Organizations: A Model.” *International Journal of Arts Management*, 2(1), pp. 38–52.
- Griffiths, R. (1995), “Cultural strategies and new modes of urban intervention”, *Cities*, 12(4), pp. 253-265
- Hausmann, A., and Poellmann, L. (2013), “Using social media for arts marketing: theoretical analysis and empirical insights for performing arts organizations”, *International Review of Public and Non Profit Marketing*, 10(2), pp. 143-161.
- Hume, M. and Mort, G.S. (2008) “Satisfaction in performing arts: the role of value?”, *European Journal of Marketing*, 42(3/4), pp.311 - 326
- Kanter, T., Rahmani, R., Walters, J., and Sauter, W. (2014), “Distributed Immersive Participation as Crowd-Sensing in Culture Events”, *Journal of Virtual Worlds Research*, 7(2), pp. 1-18
- Kaplan, A. and Haenlein, M. (2010), “Users of the world, unite! The challenges and opportunities of social media”, *Business Horizons*, Vol. 53 No. 1, pp. 59—68.
- Mills, B.M., and Rosentraub, M.S. (2013), “Hosting mega-events: A guide to the evaluation of development effects in integrated metropolitan regions”, *Tourism Management*, 34, pp.238-246
- Nielsen (2012), State of the media: the social media report 2012, available at <http://www.nielsen.com/us/en/insights/reports/2012/state-of-the-media-the-social-media-report-2012.html>
- Price, K. (2006) ‘Web 2.0 and education: What it means for us all’ In *2006 Australian Computers in Education Conference*, 2–4 October, Cairns, Australia.
- Radbourne, J., Johanson, K., Glow, H., and White, T. (2009), “The Audience Experience: Measuring Quality in the Performing Arts”, *International Journal of Arts Management*, 11(3), pp. 16-29
- Rentschler, R. and Potter, B. (1996), “Accountability versus artistic development”, *Accounting Auditing & Accountability Journal*, 9(5), 100 – 11.
- Ribeiro, J.P., and Peixoto, J.P. (2014), “Searching for Events through Social Media: An Exploratory Study”, *International Journal of Marketing, Communication and New Media*, 2(3), pp. 91-116
- Richardson, W. (2006) *Blogs, Wikis, Podcasts, and other powerful tools for classrooms*. Thousand Oaks, CA: Sage.

Strom, E. (1999), "Let's put on a show! Performing arts and Urban Revitalization in Newark, New Jersey", *Journal of Urban Affairs*, 21(4), pp. 423-435.

Towse, R. (2001), "Quis custodiet? Or managing the management: the case of Royal Opera House, Covent Garden", *International Journal of Arts Management* 3(3), 38-50.

Turbine, J. and Laurin, C. (2009), "Performance Measurement in the arts sector: the case of performing arts", *International Journal of Arts Management*,

Whitford, M. (2009), "A framework for the development of event public policy: Facilitating regional development", *Tourism Management*, 30(5), pp. 674-682.

Wood, E. H. (2005), "Measuring the economic and social impacts of local authority events", *International Journal of Public Sector Management*, 18(1), pp. 37-53