

What can social media data add to the knowledge of arts and humanities? An empirical investigation on Twitter at Teatro Alla Scala

Agostino, D.¹ and Arnaboldi, M.²

1. Introduction

Social media can be considered as a representative example of big data with their high volumes, high velocity and high variety features (George et al., 2014; Gandomi and Haider 2015). Social media are continuously receiving attention in the arts and humanities literature with several studies exploring how social media can be used to enhance audience engagement, informal learning or marketing activities in arts and cultural organizations (Hausmann, 2012; Hausmann and Poellmann, 2013; Padilla-Meléndez and Águila-Obra, 2013; Slatten et al., 2016). These studies have advanced our current understanding about potential applications of social media in this field, but we know a little about the opportunities provided by data extracted from social media to enhance knowledge management in the arts and humanities.

Acknowledging this gap, this chapter is aimed at understanding if and how social media data can contribute to generating new knowledge in the arts and humanities. More specifically, two research questions are here addressed:

- How can social media data be extracted in alignment with intended use?
- How can social media data be analysed?

The first research question tackles the problem about the criteria for downloading data from social media, considering the final use of information. The second research question is focused on the identification of a set of indicators that can be calculated starting from the extracted social media data and allow the generation and communication of new knowledge. These two research questions are addressed by developing a framework for social media data valorisation for the arts and humanities. The framework is constituted by two sections, one specifically addressing the issue of extracting social media data and another focused on analysing social media data. The framework will be here presented theoretically and then empirically applied to the valorisation of the Twitter dataset for Teatro Alla Scala.

¹ Corresponding author
Deborah Agostino, Politecnico di Milano
deborah.agostino@polimi.it

² michela.arnaboldi@polimi.it

This chapter is articulated into five main sections. The next section provides the literature background by introducing the distinctive features of social media data and extant studies about social media in the arts and humanities. The following section presents the framework for social media data valorisation by detailing the phases of data extraction and data analysis. An empirical application of the proposed framework to a one-year Twitter dataset of Teatro Alla Scala is then proposed. This empirical application is considered particularly useful since it tested in practice the proposed theoretical model. Finally, the last section critically discusses the proposed framework posing the attention on the implication of this study for academics and practitioners of knowledge management in the arts and humanities.

2. Literature background: social media data in performing arts

Social media data refer to a set of data extracted from social media platforms, such as Facebook, Twitter, YouTube, Instagram or Google+. Social media data differ from traditional financial and non-financial data along three main dimensions: the frequency at which data are generated, the source of data generation, and the format of data.

The first distinctive feature is related to the *frequency of data generation*. Social media data are real time data (Kietzmann et al., 2011). They are generated continuously during the whole day (24/7). This aspect creates an opportunity to have continuous access to the social media content, but, at the same time, it poses problems of storage capacity given the high volume data generated each moment of the day.

The second distinctive feature concerns the *source of data*. It is said that social media data are generated by users (Richardson 2006; Chun and Luna-Reyes 2012). This aspect provides the great advantage of collecting information without asking directly to the individual user as it happens with customer satisfaction surveys that require a great effort to identify customers/audience and ask them to answer to some questions. Through the analysis of social media data, users are not required to answer questions, but their conversations are monitored unconsciously. At the same time, the major risk connected with this aspect concerns data reliability since the user generated content is not validated nor certified as it happens with traditional financial data.

The third distinctive feature relates the *format of data*. Social media data can be defined as unstructured data (George et al., 2014). They come in a variety of formats, which include texts, videos, links or photos. This aspect renders the analysis of social media data more complex than the analysis of traditional (financial and non-financial) data that arrive in a numeric format only.

These three main features render social media data significantly different from traditional financial and non-financial data. Novel approaches and novel techniques are therefore emerging for valuing and analysing social media data (see Agostino and Sidorova, 2016 for a review). Studies exploring how to value social media data, their benefits and pitfalls are continuously flourishing in information system and accounting literature (e.g. Chen et al., 2012; Gandomi and Haider, 2015). Yet we have a limited evidence on how social media data can add value in the arts and humanities.

The available literature on social media in the arts and humanities mainly tackles the role of social media in enhancing marketing activities (e.g. Hausmann, 2012; Hausmann and Poellmann, 2013), audience engagement (Bakhshi et al., 2010; Freeman, 2010), informal learning (e.g. Russo et al., 2009) or external accountability (e.g. Slatten et al., 2016). These studies highlight the opportunities offered by social media in performing arts organizations and often discuss with a critical lens experiences of social media adoption in these institutions. While these studies have enhanced our understanding about how social media can be used in this type of organizations, far less is known about the role of data generated by these social platforms: can social media data serve performing arts organizations? Can they enrich and provide additional insights to managers of these organizations? How conversations on social media can be valued? These aspects are mainly unexplored in this field.

At the managerial level, the great potentialities associated with social media data have been often claimed:

“A mere Tweet from a trusted source can cause losses or profits of billions of dollars and a chain reaction in the press, social networks, and blogs.” (George et al., 2014, p. 324).

This has prompted a proliferation of studies that seek to develop analytics for big data. This available managerial literature suggests that social media data can favour a better understanding of customers, their opinion and the strength of the relationship created through social media, proposing in some cases ad hoc indicators for social media (e.g. Bonsón, and Ratkai, 2013).

Unlike this recognition, we have to date limited evidence about the potentialities offered by social media data in generating new knowledge in the arts and humanities.

This chapter aims to address this gap by proposing a framework to support knowledge generation from social media data with a particular focus on performing arts institutions. The developed framework tackles two main aspects: the extraction of social media data considering the final use of the information and the analysis of social media data.

3. A reference framework for social media data valorisation in the arts and humanities

This section proposes a reference framework for social media data valorisation in the arts and humanities (see Figure 1).

The background assumption behind the development of the framework is that social media data significantly differ from traditional financial and non-financial data. Social media data are real time, user generated, and come in a variety of formats. These distinctive features requires the development of an ad hoc framework for data valorisation that should be aligned with the final use of the information.

More specifically, the proposed framework suggests the valorisation of social media data by following two main phases of analysis: an initial phase of data extraction and a following phase of data analysis. Both of these phases will be described following.

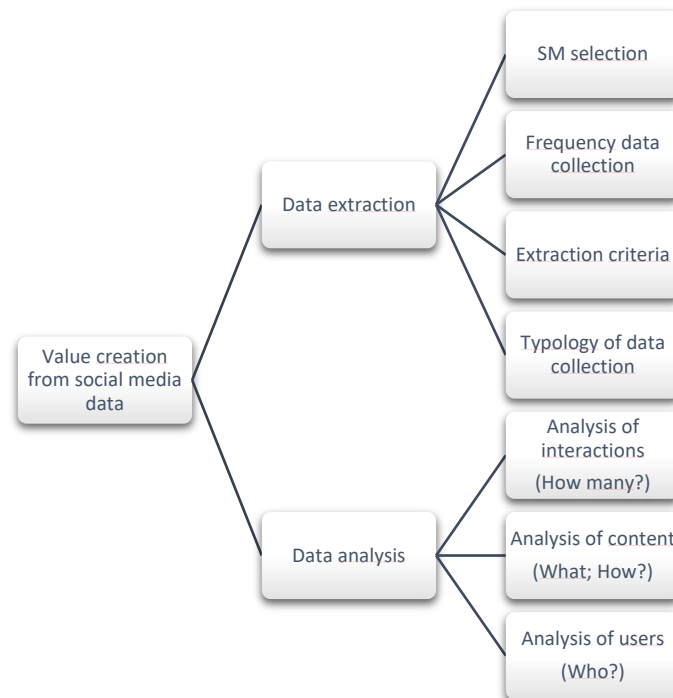


Figure 1: Framework for social media data valorization

3.1 Data extraction

The first phase of the framework requires the extraction of social media data. Extracting data is usually a neglected phase when dealing with traditional financial and non-financial data since data are already available in the information system of the organization. Yet social media data are user generated and they are available in multiple formats. Moreover, data are not available on proprietary systems, but they need to be extracted from the web. This poses the issue about data extraction crucial for the final aim of data valorization. The framework tackles the issue of data extraction not from a

technical information system perspective, but from a managerial perspective posing the attention on the managerial decisions required to generate a meaningful set of social media data.

More specifically, the data extraction phase requires four main decisions (see Table 1):

- Selection of the social media platform object of analysis;
- Frequency of data collection;
- Extraction criteria;
- Typology of data collection.

Data extraction decision	Description
Social Media platform selection	One social media platforms vs data multiple social media platforms
Frequency of data collection	Real time vs Periodical
Extraction criteria	Per account vs per keywords
Typology of data collected	Data about interactions generated by a post Data about content of the conversation Data about the users

Table 1: Data extraction decisions

The first decision is related to the *selection of the social media platform* object of analysis. This refers to the decision on whether to download data from a single social media platform or to extract data from multiple sources. This decision should not be driven by the social media platform on which the institution is present since even though an organization is not present on a social media platform, users can however generate content about the organization itself. This resonates with the recognition that : “It’s no longer a choice of whether or not you are on social media. You’ve got to be there. And if you’re there, you have to have governance” (Forbes, 2015). The selection of the social media platform is often influenced by the policies about data access set by social media providers. As of today, for example, data from Twitter can be downloaded for free making easier the possibility to extract data from this platform. Facebook has more restricted policies that allows the data download only for the followers of a given account. This restriction renders more difficult the possibility to download data from accounts not followed.

The second decision concerns the *frequency of social media data collection*. In this respect, two main options are available. On the one hand, data can be downloaded real time. This gives the possibility to analyse social media data while interactions and conversations are taking place and, eventually, immediately intervene. This is suggested when the purpose is a real time customer care. On the other hand, data can be extracted periodically, by setting a period for data collection such as the day, the

week or the month. This is the preferred approach when the purpose is to analyse a given phenomenon without the urgency to take immediate action.

The third decision is related to the definition of the extraction criteria. Defining the extraction criteria means defining whether to download data per account or per keywords. The data download per account is associated with the extraction of all those posts generated by the social media account object of analysis. This means that if a generic user is posting about that account, such post will not appear in the downloaded dataset. The data extraction per account is particularly adopted when the purpose is to monitor how an account is performing on social media (i.e. how many interactions are generated by the posts of a given account? What is the level of virality and dialogue generated by a given account?). A second option requires the data download per keywords. This means that a set of words are identified and all the posts containing that word are extracted irrespectively from the account that has generated that post. This second option is more time consuming since a set of keywords need to be identified and a higher effort in the data crawling from the web is required. The selection of the keywords should be aligned with the final purpose of use of the downloaded information. For example, if the aim is to understand the perception about a given opera, then a set of keywords concerning the title or the actors involved in the operas will be selected. Selecting keywords should be a careful activity that might require a preliminary testing phase in order to avoid “dirty data”, intended as those data containing the identified keywords but not linked to the final aim of the analysis.

The last decision related the typology of data to be extracted. A single post is a rich source of information, which include: text, eventually photos, links, videos, a sender account, a receiver account, a timestamp, a location, eventually tag, mentions, reactions and comments connected to the post itself. This list of available information requires a predefinition of what is to be downloaded per each post. If, on the one hand, all of the extractable information might be considered useful, this is time consuming and costly if data are then not used. Again, there should be an alignment between the purpose of use of the information and the downloaded data. For example, if the purpose is to understand the ability in generating dialogue and engagement with the audience, then reactions, like and replies will be extracted. If the purpose is to understand the topics of discussion and the users profiling, then also text and features of the accounts have to be downloaded.

Once these four decisions are clarified, then data are extracted and an available dataset is ready for analysis.

3.2 Data analysis

The second phase of the framework is centred on the analysis of social media data. The analysis of social media data is often intended in terms of interactions associated to a post on social media. Yet, the analysis of interactions is one of the possible analysis that can be performed on social media data. More precisely, three main typologies of analysis can be performed (see Table 2):

- Analysis of interactions
- Analysis of the content
- Analysis of the users

Dimensions of data analysis	Question addressed	Proposed indicators	Description
Analysis of interactions	How many interactions are generated by a social media post?	Total engagement	Quantifies the total amount of interactions between social media users and the organization by counting the average number of like, share/retweet and comments per post
		Level of dialogue	Quantifies the extent of dialogue between social media users and the organization by counting the average number of comments per post
Analysis of the content	What is the content about?	Word frequency indicator	Sorting words on the basis of their occurrences inside social media messages
	How is the opinion about the content?	Opinion indicator	Sorting social media messages into negative, positive or neutral on the basis of the score of each word inside the message (where, [-5] is a negative opinion, [+5] is a positive opinion and [0] is a neutral opinion)
Analysis of the user	Who is talking about?	Authority Index	Sorting users on the basis of the total number of sent messages
		Hubness Index	Sorting users on the basis of the total number of received messages

Table 2: List of indicators

3.2.1 Analysis of interactions

The analysis of interactions aims at quantifying the strength of the relationship between the institution and its social media users, with particular reference to the ability of the organization to establish dialogue and a two-way relationship with its social media users. It has been widely acknowledged that social media, relying on Web 2.0 features, are intended to facilitate dialogues and interactions (Bonsón and Ratkai, 2013). The analysis of interactions wants therefore to answer to the following questions: How many interactions are generated by a social media post?

Accordingly, two indicators have been here proposed: total engagement and level of dialogue.

The indicator of *total engagement* quantifies the overall level of interactions between the performing arts organization and its social media users. It is computed by counting the average value of likes, comments and shares (or retweets) per post, over a given time horizon. Through this indicator, it is possible to have an overall view about the average level of engagement between an organization and its network of social media users. This analysis can support the identification of the types of posts that enhance engagement as well as the preferred social media platform if the analysis is performed on more than one channel.

The indicator about the level of dialogue is specifically focused on the quantification of the level of dialogue generated by a social media post. It is computed as the average number of comments per post, rather than counting also like and shares/retweets. It provides a punctual indication about the ability of a social media post to stimulate dialogue within the social media community.

3.2.2 Analysis of the content

The analysis of the content aims at answering the question on “what people are talking about” on social media and “how people are talking about on social media?” Therefore, it quantifies the content and the opinion of social media conversations. Two indicators are included in this dimension: word frequency indicator and opinion indicator.

The word frequency indicator counts the number of occurrences of a given word in a list of social media posts. By counting the occurrences of words in the list of posts in a given period, it is possible to identify the most frequent topics of discussion. This insight, triangulated with the information derived from indicators about dialogue and engagement, can be useful to understand which topics raise higher interest and interactions.

The opinion indicator quantifies the perception by users in terms of positive, negative or neutral perceptions. This indicator is quantified by assigning a score to each word inside a post within a range [-5; +5], where [-5] consists of a negative opinion and [+5] a positive opinion; then, the weighted

average of the scores of each word in the tweet is computed arriving at the final opinion indicator value. The opinion analysis can be useful to detect real time reputational risks or elements of dissatisfaction for users.

3.2.3 Analysis of users

The third areas of analysis refers to users with the purpose to answer questions such as: how are social media users connected with my institutions how do users interact with each other, which is their influence and importance in my social media network? Indicators in this area are intended to quantify the structure of the network created through social media connections. The computation of these indicators requires the construction of the network of social media users first. Chosen a specific social media platform (e.g. Twitter), the network is built by collecting all the messages from and to the organization's social media account. From the list of messages, the network can be built considering each user (both cited in a message or receiving/sending a message) as a node, and each message as a link between node. Accordingly, two main indicators can be computed to evaluate the structure of network created by social media: level of hubness and level of authorities.

The indicator about the level of hubness allows the identification of the most active social media users included in the network. It is computed as the eigenvector of a transformation AA^T of the adjacent matrix of the network A and, on the bases of the obtained results, it allows to sort users on the basis of the total number of sent messages (Kleinberg, 1998). It ranges from 0 to 1 (i.e. respectively the lowest and highest level of hubness). Users with a high level of hubness are those users that send the highest number of messages in that network; this value should be analysed carefully given that a user with a high level of hubness might also be a "spammer".

The indicator about the level of authorities sorts social media users on the basis of the number of messages they receive inside the network. This is a relevant information given that a social media user with a high level of authority represents someone that receives the highest number of posts in the analysed network and can therefore be considered as an authority inside the network itself. These users can represent key actors to spread key message given their relevance inside the network. This indicator is computed as the eigenvector of the transformation $A^T A$ of the adjacent matrix A associated to the network (Kleinberg, 1998). As the hub index, it ranges from 1 to 0. A high level of authority corresponds to values closed to 1.

Through the proposed framework, after deciding how to extract and how to analyse data, it is then possible to interpret the extracted data providing a novel view on performing arts institution that is based on real time and user generated data.

4. An empirical application of the framework to La Scala Opera House

The proposed framework has been empirically validated and tested on a one year Twitter dataset of Teatro Alla Scala. La Scala was established in 1778 as an independent body, which become a foundation in 1997; its activity on social media started in 2009 and it is now active on 5 social media platforms: Facebook, Twitter, YouTube, Instagram and Pinterest. The empirical application occurred in a joint interaction with the management of the Teatro Alla Scala in order to continuously share insights emerging from the analysis. This joint activity was of crucial importance for the empirical application of the model since it highlighted the centrality of the decision making in the application of the framework and the importance to have knowledge of the context while performing the analysis. The joint work with the management of La Scala lasted nine months, during which several data sources were collected. A first type of data source is represented by interviews with the head of the communication, marketing and social media manager of the opera house in order to understand the strategy behind social media adoption. A second type of data source consists of 6 meetings held at La Scala to discuss insights from the performed analysis and proceed with the project. A final data source is the social media databases of Twitter for the year 2016.

4.1 Data extraction on La Scala Twitter dataset

Once approaching the issue of valuing social media data, four main decisions were implemented in accordance with the proposed framework (Table 3).

Data extraction – main decisions	Choices for Teatro Alla Scala dataset
SM selection	Twitter
Frequency of data collection: real time vs periodical	Periodical
Extraction criteria: per account vs per keywords	per keywords
Typology of data collected	URL Timestamp Content Tweet_id Retweet_count User_name User_language User_description User_id

Table 3: Data extraction decisions for the Twitter dataset of Teatro alla Scala

The first decision about social media selection led to identify Twitter as the preferred source for valuing social media data given the possibility of the social platform to extract data for free.

With reference to the second decision about the frequency of data collection, social media data were extracted on a daily basis since there was no need to monitor conversation real time. The purpose of the analysis was to understand the social media audience of the opera house. This specific intent did not require a continuous access to the social media platform.

The third decision is about the extraction criteria. Data were extracted by keywords given the specific intent to develop a better knowledge of the audience. An analysis of the posts generated by the accounts only would have been too limited. More precisely, all the posts containing the following words were downloaded: #lascala and #teatroallascala.

Finally, the last decision about the type of data to extract led to the identification of the following set of data: URL; timestamp (i.e. timing of the post), content, tweet_id, retweet_count, user_name, user_language, user_description, user_id. The main aim to develop a better knowledge of the audience led the management of La Scala, jointly with the research team, in selecting some data about users and their conversation.

Once clarified these four decisions, the result was a file .csv containing the list of La Scala tweets over one year time horizon.

4.2 Data analysis

On the downloaded data set, the list of proposed indicators were calculated and discussed jointly with the management team.

With reference to the analysis of interactions, both the indicators about the total engagement and level of dialogue were computed (see Table 4). Yet, once performed the analysis, the management of La Scala considered these insights not enough to understand whether than numbers could be considered as a good level of engagement and dialogue. A benchmarking of these two indicators with La Scala posts on Facebook alongside with the dimension of the followers on these two platforms enhanced a more thorough interpretation of the data.

	Twitter	Facebook
Total engagement indicator	43.86 interactions/post	634.74 interactions/post
Level of dialogue indicator	0.54 comments/post	7.62 comments/post
Number of followers/fan of La Scala page	268,000 followers	240,989 fan

Table 4: Analysis of interactions for Teatro alla Scala

With reference to these indicators, two main aspects deserve attention. First, Facebook, if compared with Twitter, is characterized by a high level of engagement and dialogue. On average, it resulted that a post on Facebook generates 634 interactions compared to only 43 interactions that happen with Twitter. In order to understand the validity of this insight, we computed the same indicators for other two opera houses (Metropolitan Opera House in New York and Royal Opera House in London) finding the same evidence: on Facebook the level of engagement is higher than that achieved on Twitter. This insight can support future social media actions.

Second, the dimension of the fan/follower base on Facebook is not as larger as the one on Twitter. This is visible by the number of social media fan/followers that is higher on Twitter rather than on Facebook. This runs quite counterintuitively with respect to the previous insight: even though the Twitter fan base is larger than the one on Facebook, engagement and dialogue are higher on the latter. This insight seems to suggest the importance of the social media Facebook to increase interactions and dialogue.

The second type of social media data analysis concerns the analysis of the content with reference to the word frequency indicator and opinion indicator.

The Word Frequency Indicator gives rise to a list of the most recurrent words on the social media Twitter, graphically represented in Figure 2, where larger babbles correspond to the most frequent words.

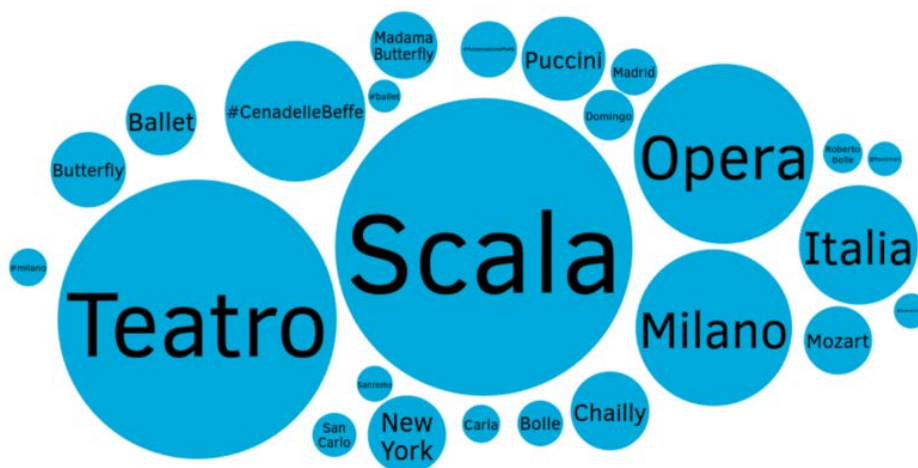


Figure 2: Word Frequency Indicator

Some reflections were shared with the management of the Opera House also in connection with this indicator:

- Although the worldwide presence of La Scala, comments and words on social media are mainly in Italian. This underlines the strong roots of the Opera House in the Italian culture;
- The brand Scala is more powerful than any other opera on stage during the season, even than “La Prima”.
- Some small operas in terms of budget and audience have been found with a high social media resonance (the example was with “La cenadellebeffe”). This was associated with a strong social media campaign and therefore this analysis allows to have a visible return on the initial social media efforts.

The Opinion Indicator allows to quantify the perception by users on social media (see Figure 3). We found mainly neutral opinions (61.5%) followed by positive opinions (31.9%) and a limited number of negative opinions (6.4%).

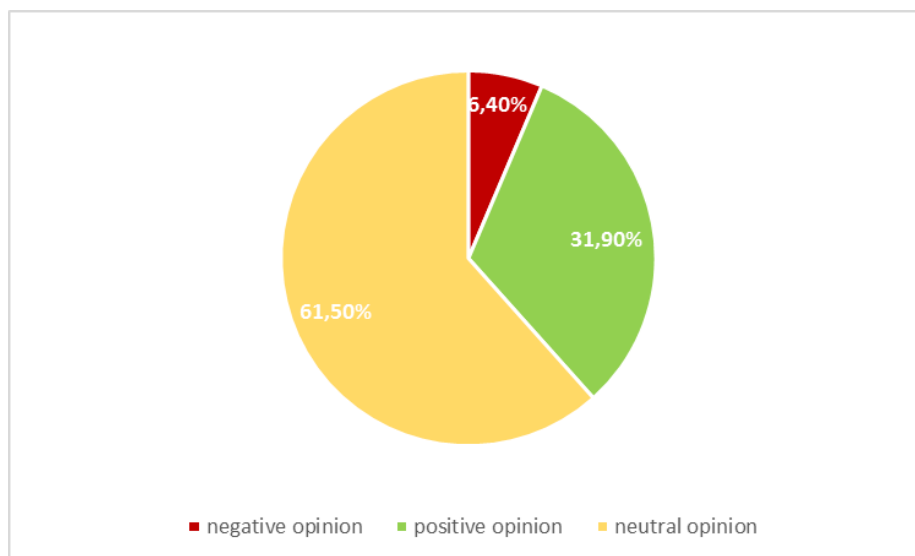


Figure 3: Opinion Indicator

Also this analysis has prompted some reflections:

- Neutral opinions predominate since several social media users use to post their presence at La Scala before attending an opera. This insight also underlines the strong connection between social media users and La Scala audience;
- Positive opinions are related to appreciations of operas and, in general, of performances, posted after the attendance at the event or after watching the performance on television;
- Negative opinions were limited compared with positive and neutral ones. A very few of them were related to the scarce appreciation of the opera, music or artists. We found negative opinions mainly related to an interruption of the online ticketing system that posed several

problems to users that would like to purchase a ticket. This underlines the possibility to use the opinion indicator to detect potential areas of reputational risk.

The third area of analysis is related to the network of users, with the computation of the indicators about level of hubness and authority.

The level of hubness allows to identify the most active Twitter users (i.e. those that send the higher number of posts). By looking individually at their Twitter account, these users have not been classified as “spammers”, but, on the contrary, they were in some cases La Scala employees that commented or promoted some events.

The level of authority supported the identification of Twitter users that receive the highest number of posts inside La Scala network. The first user was inevitably the Opera House itself given that the network has been constructed taking the organization itself at the centre. The following users were artists of La Scala; they are classified as “authorities” meaning that the network listen and take care about their posts, even if they are not the most active users; they can therefore represent strategic actors to increase the virality of La Scala’s posts.

Level of hubness		Level of authority	
Twitter name	Value	Twitter name	Value
private individual	1,0000	company	1,0000
employee	0,8828	artist (dancer)	0,1080
private individual	0,6902	artist (dancer)	0,0768
private individual	0,5836	company	0,0486
private individual	0,4029	company	0,0435
private individual	0,3969	artist (dancer)	0,0208
private individual	0,3553	artist (choreographer)	0,0197
company	0,3443	artist (dancer)	0,0174
artist (dancer)	0,3072	artist (soloist)	0,0168
company	0,2828	artist (dancer)	0,0105
private individual	0,2627	artist (dancer)	0,0102
association	0,2302	artist (dancer)	0,0100
company	0,2249	company	0,0094
private individual	0,2220	artist (dancer)	0,0093
private individual	0,2123	artist (conductor)	0,0090

Table 5: Indicators about the network of users

5. Discussion and conclusion

This study aimed at understanding if and how social media data can contribute to generate new knowledge in the arts and humanities. Two research questions are addressed: how can social media

data be extracted and then analysed in alignment with the intended use of data. The empirical experiment allows to draw some more general conclusions of interest for both practitioners and academics.

The first area of results is related to the framework proposed, in which a missing variable need to be inserted: the decision makers. Both the social media extraction and analysis were carried out in close cooperation with the decision makers, sharing since the beginning the practical research question they wanted to face. All along the study the research team and the managers shared decisions enhancing the quality of data in relation to the organizational interest. The close cooperation required several meetings, but it improved the finale output. For example after the first data collection, the marketing manager highlighted the need to consider a larger data set of key words, in order to trace some trends at the international level. This was not considered at the beginning, as he thought about a more limited geographical boundary, linked to the possibility of people to come to the theatre. The analysis of the data clarified to the manager the diversity of thinking in the digital layer. A similar pattern was visible in the data analysis, where the managers asked the research team to focus on some issues, especially for monitoring the network of users.

Our findings highlight that social media are translated in new *actionable* knowledge with a joint path between analysts and decision makers; through this interactive pattern there is a reciprocal process of learning and the construction of a new decision making space, that in the digital layer has different space and time.

A second area of findings is more specifically related to Performing Arts and Humanities and their “control” in the digital age. Through the path of data collection and analysis, the results of our study provides a reference scheme for performance management articulated in new units of analysis, indicators, and action time. Regarding the unit of analysis, social media data revealed the possibility of managing two units of analysis previously neglected: individuals and networks. In both cases the empirical application allowed to develop a theoretical enhancement but also an instrument for managers. The characterization of network of users is exemplary of that. This new unit of analysis was defined and measured with an articulated set of indicators stemmed from the theory but filtered to fit the needs of the theatre, arriving to a restricted and coherent set of indicators. The indicators about users allows to trace: the structure of the network and its evolution in comparison with major international theatres; different typologies of influencers, considering the performing art sector dynamics; the relationship between users.

Through the proposed indicators, it is possible to identify the type of posts that increase the level of engagement, potential risks or strategic social media actors. However, the emerging insights, are intended, not only to serve social media managers, but also other managers inside performing arts

institutions. For example, the communication area is enriched with a set of information related to the social media audience, their preferences and their relevance. The marketing area can use insights derived from the proposed indicators to better shape promotional campaigns on the basis of the general preferences and characteristics of social media users. Finally, the word frequency indicator and the opinion indicator can support the overall management of an opera house detecting potential areas of risks, such as the problem with the online ticketing system happened at La Scala.

Of course, also pitfalls and drawback need to be underlined: some indicators (i.e. those about users) require the availability of social media posts that are not always available for free. At the time of our analysis, social media posts could be downloaded for free for Twitter, while it was not possible for Facebook (they can be downloaded using ad hoc platforms of social media analytics). A second drawback is related to the need of some statistical skills in order to compute indicators: the indicators about users, for example, required the construction of the network of social media users that was performed through an ad hoc software (i.e. we used the software R). When practically applied in an organization, this might imply the involvement of external technical experts or to develop internal competences in order to perform the analysis. A third element of attention concerns the importance to know the history, background and context of the organization object of analysis; some indicators might be useless if they are not accompanied by an interpretation that can be provided only with reference to managerial and previous experiences. For example, results obtained from the word frequency indicators could appear useless if analysed “per se”. On the contrary, when interpreted within a broader performing arts strategy, that can provide insights to evaluate past actions and drive future decisions.

To summarise, our study contributes at the managerial and academic level to performing arts but also more general to managerial studies. The developed framework enhances the current management literature with a set of decisions for data extraction and indicators to value social media data, answering the recent calls to exploit the value connected with social media data (George et al., 2014). The set of indicators also serves the performing arts literature, where measurement system has often been perceived with a negative connotation as limiting the creativity of organizations (Felton, 1994). This study underlines benefits that can be derived by using indicators in these organizations and contributes in this way to an emergent stream that recognised that measurement and creativity that characterised performing arts organizations are linked and not opposed.

References

- Agostino, D., & Sidorova, Y. (2016). "A performance measurement system to quantify the contribution of social media: New requirements for metrics and methods" *Measuring Business Excellence*, 20(2), 1–21.
- Bakhshi, H., Garcia, J.M. and Throsby, D. (2010) "Beyond Live: Digital Innovation in the Performing Arts," National Endowment for Science, Technology and the Arts (NESTA), London, Research briefing (February 2010), at https://www.nesta.org.uk/sites/default/files/beyond_live.pdf
- Batagelj, V. and Zaversnik, M. (2002) An O(m) Algorithm for Cores Decomposition of Networks, arXiv:cs/031004
- Bonsón, E. and Ratkai, M. (2013) "A set of metrics to assess stakeholder engagement and social legitimacy on a corporate Facebook page", *Online Information Review*, Vol. 37, No. 5, pp. 787-803.
- Chen, H., Chiang, R.H.L., and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, 36(4): 1165-1188
- Chun, S. A., and L. F. Luna-Reyes (2012) "Social Media in Government." *Government Information Quarterly*, 29 (4): 441–445.
- Felton, M.V., (1994) "Evidence of the Existence of the Cost Disease in the Performing Arts", *Journal of Cultural Economics*, Vol. 18, pp. 93–112.
- Forbes (2015), "4 Components Of Robust Social Media Governance Program In Financial Services", 15 November 2015, available at <https://www.forbes.com/sites/joannabelbey/2015/11/15/4-components-of-robust-social-media-governance-program-in-financial-services/#7190ef122156> (retrieved April 6th 2018)
- Freeman, J. (2010) "Web-based collaboration, live musical performance and open-form scores", *International Journal of Performance Arts and Digital Media*, Vol. 6, No. 2, pp. 149-170.
- Gandomi, A. and Haider, M. (2015) "Beyond the hype: big data concepts, methods and analytics", *International Journal of Information Management*, Vol. 35, No. 2, pp.137-144.
- George, G., Haas, M.R. and Pentland, A. (2014) "Big Data and Management", *Academy of Management Journal*, Vol. 57, No. 2, pp. 321 – 26.

- Hausmann, A. (2012) “Creating ‘buzz’: opportunities and limitations of social media for arts institutions and their viral marketing”, *International Journal of Nonprofit and Voluntary Sector Marketing*, Vol. 17, No. 3, pp. 173-182.
- 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.
- Kietzmann, J. H., K. Hermkens, I. P. McCarthy, and B. S. Silvestre (2011), “Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media.” *Business Horizons* 54 (3): 241–251
- Kleinberg, J. (1998) “Authoritative sources in a hyperlinked environment” Proc. 9th ACM-SIAM Symposium on Discrete Algorithms. Extended version in *Journal of the ACM* 46(1999). Also appears as IBM Research Report RJ 10076, May 1997.
- Padilla-Meléndez, A. and Águila-Obra, A.R. (2013), “Web and social media usage by museums: Online value creation”, *International Journal of Information Management*, 33(5): 892-898.
- Rapoport, R. (1970) “Three Dilemmas of Action Research”, *Human Relations*, Vol. 23, No. 6, pp. 499-513
- Richardson, W. 2006. *Blogs, Wikis, Podcasts, and Other Powerful Tools for Classrooms*. Thousand Oaks, CA: Sage
- Russo, A., Watkins, J. and Groundwater-Smith, S. (2009) “The impact of social media on informal learning in museums”, *Educational Media International*, Vol. 46, No. 2, pp. 153-166
- Slatten, L.A.D., Hollier, B.N.G., Stevens, D.P., Austin, W. and Carson, P.P. (2016), “Web-Based Accountability in the Nonprofit Sector: A Closer Look at Arts, Culture, and Humanities Organizations”, *The Journal of Arts Management Law and Society*, Vol. 46, No. 5, pp. 213-230.
- Seidman S. B. (1983), “Network structure and minimum degree”, *Social Networks*, Vol. 5, pp. 269–287.