The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0737-8831.htm

Readers' engagement through digital social reading on Twitter: the TwLetteratura case study

Federico Pianzola
University of Milan–Bicocca, Milan, Italy and
Sogang University, Seoul, Republic of Korea, and
Maurizio Toccu and Marco Viviani
Department of Informatics, System and Communication,
University of Milan–Bicocca, Milan, Italy

Received 13 September 2020
Revised 21 February 2021
Accepted 21 February 2021

Abstract

Purpose – The purpose of this article is to explore how participants with different motivations (educational or leisure), familiarity with the medium (newbies and active Twitter users), and participating instructions respond to a highly structured digital social reading (DSR) activity in terms of intensity of engagement and social interaction.

Design/methodology/approach – A case study involving students and teachers of 211 Italian high school classes and 242 other Twitter users, who generated a total of 18,962 tweets commenting on a literary text, was conducted. The authors performed both a quantitative analysis focusing on the number of tweets/retweets generated by participants and a network analysis exploiting the study of interactions between them. The authors also classified the tweets with respect to their originality, by using both automated text reuse detection approaches and manual categorization, to identify quotations, paraphrases and other forms of reader response.

Findings – The decoupling (both in space and time) of text read (in class) and comments (on Twitter) likely led users to mainly share text excerpts rather than original personal reactions to the story. There was almost no interaction outside the classroom, neither with other students nor with generic Twitter users, characterizing this project as a shared experience of "audiencing" a media event. The intensity of social interactions is more related to the breadth of the audience reached by the user-generated content and to a strong retweeting activity. In general, better familiarity with digital (social) media is related to an increase in the level of social interaction.

Originality/value – The authors analyzed one of the largest educational social reading projects ever realized, contributing to the still scarce empirical research about DSR. The authors employed state-of-the-art automated text reuse detection to classify reader response.

Keywords Digital social reading, Twitter, Social media, Text reuse detection, Teaching methods, Teaching literature

Paper type Research paper

1. Introduction

Digital social reading (DSR) is a growing phenomenon regarding the use of social media to discuss reading experiences. Its two most studied forms are online book reviews (Dimitrov *et al.*, 2015; Driscoll and Rehberg Sedo, 2018; Hegel, 2018; Thelwall, 2019; Thelwall and Kousha, 2017) and social annotation of texts (Blyth, 2014; Gao, 2013; Kalir and Garcia, 2019; Marshall and Brush, 2004; Thoms and Poole, 2018). DSR offers great opportunities to several disciplines, such as literary studies, education and library studies, inasmuch as it allows to

© Federico Pianzola, Maurizio Toccu and Marco Viviani. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

Federico Pianzola has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 792849.

Library Hi Tech
Vol. 40 No. 5, 2022
pp. 1305-1321
Emerald Publishing Limited
0737-8831
DOI 10.1108/LHT-12-2020-0317

collect a great quantity of data about readers' response to literature, language learning, and the promotion of reading (Rebora *et al.*, 2021; Rebora and Pianzola, 2018). Tracking the behavior of users when they share online their reading experiences allows to evaluate "how digital media are creating new social valences of reading" (Nakamura, 2013, p. 238; cf. Thomas, 2020; Murray, 2018). In this paper, we analyze the use of Twitter for an educational project that involved Italian high schools' teachers and students, and other external Twitter users, in commenting on a novel over a specified period of time.

The overall goal behind this study was to explore how participants with different motivations (educational or leisure), familiarity with the medium (newbies and active Twitter users), and participating instructions respond to a highly structured DSR activity in terms of intensity of engagement and social interaction. In order to evaluate such differences, we performed a quantitative analysis focusing on the number of tweets/retweets generated by participants and a network analysis exploiting the study of interactions between them. We also classified the tweets by considering their originality, employing both automated text reuse detection approaches and manual categorization to identify quotations, paraphrases and other forms of reader response. Additionally, a related aim was to observe to what extent the organizers' intervention in the social reading activity impacts the participants' response. The identification of positive or negative managing strategies can, in fact, be helpful for the future organization of similar initiatives for the promotion of reading.

2. Literature review

DSR can represent a powerful tool for both schools and libraries actively pursuing to promote reading, an example of how the affordances of digital technology can be employed to provide support for linguistic and cultural literacy development. Thoms and Poole (2017) have identified three main kinds of comments generated when using digital annotation tools: (1) linguistic, related to the meaning of words or sentences; (2) literary, related to the interpretation of textual elements; and (3) social, used to display emotional response, provide encouragement, express one's opinion about another's comment or provide a comment not directly related to the text. Blyth (2014) also showed that annotations are used to reflect upon cultural differences, to connect reading to personal experience and to coconstruct meaning, with respect to both foreign words and literary features. Something similar also occurs beyond classroom activities: Pianzola *et al.* (2020) found analogous patterns of interactions in Wattpad users, who use comments in the margins of texts both to have conversations elicited by the story and to leverage peer learning when faced with more complex texts.

One of the main differences between the various DSR practices concerns the time passed between the reading act and the commenting act (Kutzner *et al.*, 2019). In the case of platforms for online reviews and book discussions, like Goodreads, the production of user-generated content is usually something that happens once reading a book or a story is concluded (Thelwall and Kousha, 2017). On the contrary, in the case of platforms designed for annotations in the margins, like Wattpad and Hypothesis, comments are generated during the reading activity. The difference is remarkable because reader response changes and is shaped by the progression of reading. In this regard, TwLetteratura, the organizer of our case study, promotes a hybrid form of DSR. By having a shared calendar and using Twitter hashtags to link comments to specific chapters, they show an interest in encouraging a conversation that follows the progression of the story. However, since text and discussion are hosted on different media, it is not certain if readers write tweets in parallel to reading. More probably, at least some of the comments are generated after having completed the reading of the whole chapter or section to which they refer. So far, there is little research about educational social reading using public social media platforms like Twitter. Vázquez Cano and Esteban (2012) had teachers ask students to collaboratively write on Twitter about topics

shown in class and found that this activity promoted substantial improvement of reading and writing skills. Domizi (2013) and Preston *et al.* (2015) asked students to reflect upon course content and tweet about it, finding that students felt more connected to both content and peers. Moreover, something similar, in terms of the timeframe during which comments are written, is also happening in other pedagogical uses of DSR, where students create comments and annotations on closed platforms as part of classroom assignments, following a calendar, knowing that they will be evaluated and, therefore, pondering them before writing (Blyth, 2014; Thoms and Poole, 2018; Kalir *et al.*, 2020).

DSR on Twitter occurs mainly in the form of recurring media events centered around an account or a hashtag often representing a "trending topic" or an event; for example, with the hashtag #folklorethursday, users share content related to folk tales and local customs. Often, such initiatives follow a regular schedule, an aspect that greatly contributes to establish a sense of community (Thomas, 2020). Normally, tweets are encountered as part of a chronological timeline determined by the date–time of a tweet or by interactions with it, rather than in the order of the original text to which they refer (which is relevant for DSR). This and other affordances of the digital platform – like the "favourite" and "retweet" functions – influence users' participation to DSR, making it more similar to a performance rather than to reading and commenting on a page (Page, 2012; Thomas, 2014).

A recent literature review on Twitter use in educational contexts (Tang and Hew, 2017) focused on how Twitter may impact the various aspects of users' interactions, identifying some open issues to be addressed, namely "to what extent should the instructor be involved in Twitter interactions, what types of interactions occurred and how to sustain the interactions to better facilitate learning" (pp. 109–110). In our case study, we also looked at these aspects in the context of a rigorous comparison-based experiment, thus filling two gaps highlighted by Tang and Hew (2017) in their literature review.

3. Case study

3.1 TwLetteratura

TwLetteratura is a no-profit association promoting a specific DSR practice:

> The community chooses a book, reads it and comments on it – one chapter after the other, following a shared reading calendar – essentially "re-writing" the book through Betwyll (the social reading webapp developed by TwLetteratura) and Twitter. Each participant proposes his or her own personal interpretation in twylls [the name of posts on Betwyll] or tweets (one or multiple, following the inspiration): rewriting includes paraphrasing, variations, comments and/or interpretation; moreover, people can employ media that are different from the original format (like videos, songs, drawings). The employment of different stylistic registers – following the Oulipo model – allows users to experiment with unlimited combinations of deconstruction and reconstruction of the original text, a system of micro-texts produced by the community (comments) related to the original text and between each other (because every twyll or tweet is a source of answers, comments, extensions). (TwLetteratura, 2017)

According to an assessment performed by TwLetteratura itself, the people interested in this kind of DSR form a community of around 250–300 members, who regularly take part in their activities, whereas the number of people who participated in at least one game is around 4,000 (Costa, 2016, p. 23). As it can be seen from TwLetteratura's Twitter profile, the audience (followers) is much wider, around 27,000 users when our case study #MattiaTW took place. We decided to analyze TwLetteratura because of the breadth of their active community; this offered us the opportunity to look at a conspicuous number of interactions between students and voluntary participants, thus potentially revealing something about the blending of formal and informal cultural spaces. More specifically, we chose to focus on #MattiaTw

(detailed in the next section) because of its experimental design, which allowed obtaining more rigorous results regarding Twitter use for DSR and learning.

3.2 #MattiaTW: a case study

The case study discussed here is about the novel *The Late Mattia Pascal* by Luigi Pirandello, published in 1904 in the original Italian version. In Italy, this book is considered a 20th century literature classic and is part of the curriculum in all high schools. The #MattiaTW project was announced in May 2016, and it was addressed to high schools, although participation was open to anybody (TwLetteratura, 2016a). Unlike other DSR activities organized by the association, this was a project designed to test the efficacy of the TwLetteratura method. Thus, it mainly involved students and teachers whose participation was part of a classroom activity, rather than a spontaneous practice driven by personal interest. This is a summary of the brief given to teachers:

> The main goal of #MattiaTw is to encourage students to read and develop this competence. The constraint of the brevity (every tweet cannot be longer than 140 characters) stimulates the ability to translate the experience of reading a text into a synthetic form. Such restriction activates the competences necessary for a "deep" reading of a complex text, competences such as critical analysis, deductive reasoning, reflection, intuition. The outcome is a better comprehension of the text.
>
> At least, this is our hypothesis. In order to empirically support it, we are collaborating with Cariplo Foundation [1] and Lombardia's Regional School Office [2] to define an objective framework for the evaluation of the project's results. To the extent that this is possible, we want to objectively measure the efficacy of our method when used to support teaching.
>
> Thus, with respect to previous TwLetteratura activities, #MattiaTw has an additional goal: collaborating with schools to understand whether and to what extent using our method affects the students' reading skills. (TwLetteratura, 2016a; our translation)

The DSR activity has been scheduled in two phases: the first one in November 2016 and the second one in January 2017, after winter recess. The main difference compared to usual TwLetteratura's social readings was that the evaluated classes have been asked to follow a stricter protocol (TwLetteratura, 2016b). Classes were free to choose how to tweet, either using a single class account or multiple individual accounts created by the students. In order to have a rigorous comparison between experimental group and control group, some classes (group B) were required to use the TwLetteratura method for the first phase, and they were free to continue using it or drop it for the second phase. Other classes (group A) were required to adopt the method only in the second phase (Table 4 in Supplementary material). The assignment of participants to distinct groups was randomized. All in all, #MattiaTw met almost all the guidelines for educational use of Twitter suggested by Tang and Hew (2017), namely the organizers provided a clear description of Twitter-related educational activities, made Twitter use mandatory, provided regular support, and provided access to a broader community.

4. Methods

4.1 Dataset description

Participants were instructed to use the hashtag #MattiaTw, followed by the number of the chapter they were currently reading, in their comments posted on Twitter (e.g. #MattiaTw/1). Tweets incorporating the hashtag #MattiaTw have been collected by employing a Python script interrogating the Twitter Streaming Advanced Programming Interfaces (API) [3].

Data collection lasted for a period of three months, from November 2016 to January 2017. In the considered period, a set of 14,072 distinct tweets has been gathered (18,962 including

retweets). Accidental duplicates have been removed before starting the analysis. Due to the possible occasional instability of the Twitter API and of the server running the Python script, the dataset has been compared and reconciliated with a similar but simpler (in terms of metadata) dataset archived by TwLetteratura in the same period. The result of this comparison is the consolidated dataset that has been used for the analyses. Technical details about the considered data are provided in Table 5 in Supplementary material.

4.2 Participants

TwLetteratura provided us with the names of the Twitter accounts of the registered classes and teachers, together with their group information. The number of classes who registered for the DSR was 211, with approximately 3,900 students in total. Group A had 57 classes and group B 54. There was also an additional group of registered classes (OtherReg, 100), who were not asked to follow any specific restriction with respect to the tweeting calendar. In order to include all the students and teachers' accounts that have been used, we manually checked the usernames found in the dataset and, where possible, assigned them to the correct group.

Overall, readers (single Twitter accounts) who participated in #MattiaTw can be divided into six groups:

(1) A (n = 135): instructed to carry out the DSR activity from chapter 11 to chapter 20 (January).

(2) B (n = 187): instructed to carry out the DSR activity from chapter 1 to chapter 10 (November) and free to continue also from chapter 11 to chapter 20.

(3) OtherReg (n = 125): classes registered to the project that did not receive specific instructions.

(4) Nonregistered (NR) users (n = 32): Twitter users who participated spontaneously, writing at least one tweet.

(5) Audience (n = 210): Twitter users who did not write any tweet but retweeted at least one #MattiaTw tweet.

(6) Organizers (n = 12): among the accounts they managed, three were directly related to the project – TwLetteratura, MattiaPascalTw and AdrianoMeisTw. The first is the official account of the association, and the latter two are the double identity of the main character of the novel.

4.3 Analysis

We analyzed the collected data using the *R* programming language (R Core Team, 2013) and the RStudio environment. We computed the number of tweets/retweets over time for each group; the number of tweets/retweets per chapter; the aggregated activity of users within each group, and the mentions that users received. Furthermore, we analyzed the interactions between groups (social network analysis) and performed content analysis (text classification and words frequency).

Since from a preliminary manual check of the dataset we noted that the content of many tweets was constituted by citations or paraphrases of the novel, we employed automated text reuse identification techniques (Smith *et al.*, 2014) to classify as "reuse-based" these kinds of tweets. After testing "passim" (Smith, 2012) and BLAST (Vesanto *et al.*, 2017), we opted for the latter tool because of its better results. The tweets identified as text reuse were 7,267 (52% of the total number of tweets). We then randomly selected a sample of 100 tweets among those automatically identified as not being text reuse and, via a manual investigation, we found that 20% of them were indeed text reuse, but we did not find any false positive. Thus, we can

estimate our classification error rate to be around 10% (see Supplementary material for more details).

To analyze the social reading activity per chapter, we selected a subset of the dataset, keeping only the tweets explicitly mentioning chapters (15,724 tweets and retweets, 83% of the total). Mean values have been calculated by considering only the number of authors active in each phase. For content analysis, we used a data subset including only the tweets identified as not being text reuse. We performed a pre-processing phase by removing stop words, in order to calculate word frequency only for the meaningful terms. For network analysis, we excluded retweets because we wanted to consider direct mentions and replies. Finally, given the very low number of tweets generated by the members of group NR and by the organizers, we decided to focus our analysis on the registered groups only (i.e. A, B, OtherReg). To calculate the importance of authors in the network, we used the degree centrality measure based on the number of incoming links held by each node, i.e. the number of mentions received by an author. Data and commented code are available in a public online repository (Pianzola *et al.*, 2021).

5. Results

5.1 Users' activity

Out of the 18,962 total tweets and retweets, the majority (15,962, 84%) were by registered users, 1,032 (6%) by spontaneous participants, and 1,968 (10%) by the organizers (Table 1). Among the classes, group B had the most intense activity, with 9,112 tweets and retweets (48% of total). Group B was 28% bigger than group A and produced 43% more tweets per phase on average. OtherReg is the only group who tweeted quite a lot in December probably because the shared calendar did not suit the planned classroom activities, and some of the chapters have been read and commented on later than groups A and B did. Another notable fact is that #MattiaTw was mainly an educational activity with almost no participation from nonregistered users who joined spontaneously.

The evolution of the DSR per each chapter (Figure 1) shows that chapters 11 and 15 had marked peaks of tweets, retweets, and engagement for all groups except OtherReg. The increase in engagement (favorites and retweets received by a tweet) is probably due to the organizers' activity: 82 retweets for chapter 11 and 65 retweets for chapter 15, way above their average for phase 2. With respect to tweets, the peak for chapter 11 is mostly due to the promotion done to attract new interest to phase 2, whereas for chapter 15 is probably due to the importance of the chapter for the plot. Being titled "I and my shadow" [*Io e l'ombra mia*], it is about an identity crisis of the protagonist. Something similar happened in chapter 5, which elicited the highest level of engagement in phase 1, after the initial hype of chapter 1. Chapter 5 ("How I was ripened" [*Maturazione*]) is in the middle of the first part of the novel – like chapter 15 is for the second part – and it is about the first identity crisis of the protagonist, when he becomes a grown-up man.

Table 1. Tweets and retweets per month for each group

	Nov 2016		Dec 2016		Jan 2017		Total	
	tw	rt	tw	rt	tw	rt	tw	rt
A (n = 135)	5	4	0	0	2,454	731	2,459	735
B (n = 187)	4,962	1,027	64	24	2,747	288	7,773	1,339
OtherReg (n = 125)	2,035	261	529	18	707	106	3,271	385
NR (n = 32)	74	87	2	2	38	19	114	108
Audience (n = 210)	–	500	–	23	–	287	–	810
organizers (n = 12)	234	1,049	6	6	215	458	455	1,513
Total	7,310	2,928	601	73	6,161	1,889	14,072	4,890

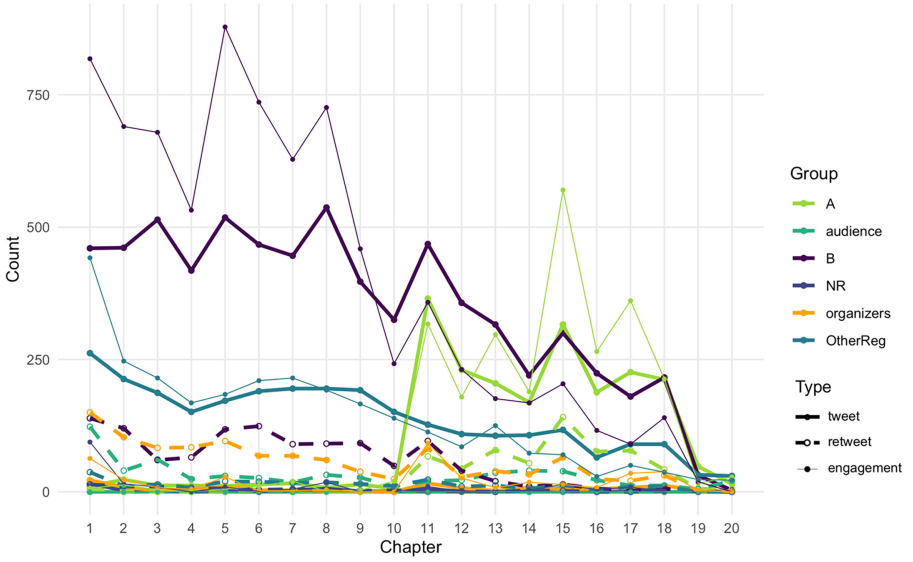

Figure 1. Number of tweets and retweets and engagement per each chapter for each group

More generally, group B's activity was more intense during phase 1 but started declining in the last week (Chapters 9 and 10). Namely, the engagement elicited went from high values to well below an even ratio – one favorite or retweet for every tweet – in phase 2, when group A joined the DSR. Together with OtherReg, group B's activity declined over time and included a very low number of retweets. Group A tweeted constantly during phase 2 but gained an engagement lower than that of group B during phase 1. Its retweeting activity was higher than B and OtherReg but still very low compared to the number of tweets generated. On the contrary, the organizers retweeted much more than creating original tweets, but even their activity declined in phase 2. The audience's attention was caught at the beginning of the project and declined after the first two weeks (chapters 1–4), then staying constant throughout the DSR and declining again in the last chapters.

More details about users' activity can be seen in Table 2. There is an outlier in group A – a class who started commenting from chapter 1, instead of 11 – which is visible in Figure 1 but has been excluded from Table 2 in order to avoid confusion. In phase 2, even though group A joined, the overall numbers of participants dropped (−20%), but this negatively affected only the number of retweets per chapter by groups B (−74% in face of a −54% difference in participants) and NR. Comparing the activity in phase 2, we can see that, on average, users of group B tweeted more than users of group A (29.6 vs 16.3 tweets), despite being 40% smaller in this phase. Moreover, in phase 2 group B tweeted more than group A in total, meaning that the users who continued the DSR were probably the most active ones in phase 1 as well. Indeed, group B's median of tweets for phase 2 was slightly higher (8) than in phase 1 (7) and their average per user increased by 10%, suggesting that a few very engaged users kept a constant intensity, whereas users who commented moderately in phase 1 either dropped out or participated less in phase 2.

In phase 1, the organizers retweeted an average of 77.4 tweets per chapter. However, the contraction in phase 2 is quite drastic (−48%), leading to 40.4 retweets per chapter, even though the total number of tweets per chapter generated by all users remained almost constant (−3%). The fact that for all groups the median of tweets and retweets per user was always much smaller than the average per user means that the majority of authors generated

Table 2.
Average number of tweets and retweets per each user (avg/usr), per chapter (avg_ch) and per chapter per each user (avg_ch/usr). Medians per user are reported to better understand the distribution of data. Data for chapters 19 and 20 have been excluded, since they are two afterwords to the novel and have been commented on only by a few users. Variation from phase 1 to phase 2 is indicated in brackets as a percentage

		A –		B $n = 169$		OtherReg $n = 85$		NR $n = 13$		Audience $n = 105$		Organizers $n = 10$		Total $n = 382$	
		tw	rt	Tw	rt	tw	rt	tw	rt	tw	rt	tw	rt	tw	rt
Phase 1 (10 ch.)	Total	–	–	4,543	948	1,908	170	35	65	–	391	54	774	6,540	2,348
	avg/usr	–	–	26.9	5.6	22.4	2	2.7	5	–	3.7	5.4	77.4	17.1	6.1
	median	–	–	7	0	7	0	2	0	–	1	1.5	7.5	3	1
	avg_ch	–	–	454.3	94.8	190.8	17	3.5	6.5	–	39.1	5.4	77.4	654	234.8
	avg_ch/usr	–	–	2.7	0.6	2.2	0.2	0.3	0.5	–	0.4	0.5	7.7	1.7	0.6
		$n = 117$		$n = 70$ (−54%)		$n = 46$ (−46%)		$n = 7$ (−30%)		$n = 53$ (−49%)		$n = 7$ (−30%)		$n = 307$ (−20%)	
Phase 2 (8 ch.)	Total	1,911	581	2,281 (−50%)	194 (−79%)	811 (−57%)	86 (−49%)	30 (−14%)	9 (−86%)	–	203 (−48%)	63 (+17%)	323 (−58%)	5,096 (−22%)	1,396 (−40%)
	avg/usr	16.3	5	29.6 (+10%)	2.5 (−55%)	17.6 (−21%)	1.9 (−6%)	4.3 (+59%)	1.3 (−74%)	–	3.8 (+3%)	9 (+67%)	46.1 (−40%)	16.6 (−3%)	4.5 (−26%)
	median	5	0	8	0	3	0	1	1	–	2	4	11	3	0
	avg_ch	238.9	72.6	285.1 (−37%)	24.2 (−74%)	101.4 (−47%)	10.7 (−37%)	3.7 (+7%)	1.1 (−83%)	–	25.4 (−35%)	7.9 (+46%)	40.4 (−48%)	637 (−3%)	174.5 (−26%)
	avg_ch/usr	2	0.6	3.7 (+38%)	0.3 (−44%)	2.2 (−2%)	0.2 (+17%)	0.5 (+98%)	0.2 (−68%)	–	0.5 (+29%)	1.1 (+108%)	5.8 (−25%)	2.1 (+21%)	0.6 (−7%)

very little tweets and retweets, but there were a few very active authors. This can be better seen with a treemap chart (Figure 2).

For all groups, there is a remarkable gap between the top authors and the rest of the participants. The top retweeters in the three registered groups were teachers using their personal account to mainly retweet comments by their classrooms. In the audience and group NR users often retweeted the same authors repeatedly, meaning that probably the amplification of the original tweet was due more to a personal connection rather than to interest in the content. Regarding tweets, the top four authors of group B are all more active than the first authors of groups A and OtherReg. The organizers retweeted using their main account but tweeted using the fictional characters' accounts. The effect of this choice can be seen through network analysis.

5.2 Network analysis

Among the 704 authors, there are 469 who, in total, mentioned 331 other authors in 1,262 tweets. The three organizers' accounts are the most central in the network (Figure 3). A few users of group B mentioned TwLetteratura very often in their tweets. On the other hand, group A almost never mentioned them, preferring to engage in conversation with the accounts of the two fictional characters managed by the organizers. There are no visible differences between the number of users from each group who mention the organizers. The network is very loosely connected, meaning that there were few social interactions, as also suggested by the low number of retweets by the registered groups. Mentions declined by 23% from phase 1 (712) to phase 2 (550). If we consider the time span during which all groups participated (phase 2), out of the total number of active authors for each group, the percentage of those who mentioned someone else at least one time is quite low: 16% for A and 18% for B, OtherReg and NR. Mentions were more frequent in phase 1: 22% for B, 28% for OtherReg and 19% for NR. The loops in Figure 3 show that some users from group A and B mentioned themselves probably because they misinterpreted the instruction received about identifying the author of the comment.

5.3 Content analysis

We estimated that, overall, more than half of the tweets (61%) were text reuse, being either citations from the novel or paraphrases of it. However, there were important differences between the groups: group A generated 870 original tweets (35% of the group's total), group B 3,367 (43%) and OtherReg 2,012 (62%). In a sample of original tweets, we identified five types of content: generalization, interpretation, intertextuality, reader's reaction, and summary (Table 3). The majority of the original content generated is an interpretation of the text (32%) or a generalization based on it (25%). There is a slightly smaller amount of readers' reactions and summaries (19%) and a few intertextual references (5%).

A further confirmation that participants focused on the story events and on generalizations elicited by them comes from word frequency analysis. The name of the protagonist (Mattia) and the term "life" (vita) are the most frequent for all three groups, followed by the infinitive form "to be" (*essere*) and terms like "man" (*uomo*) and "freedom" (*libertà*) (Figure 4).

6. Discussion

Previous study found that participants' unfamiliarity with Twitter, information overload, and imbalance of contribution among users can be a challenging aspect for a successful use of Twitter in a learning activity (Tang and Hew, 2017). Nevertheless, most studies reported positive results regarding learners' interaction with the platform and its facilitation of

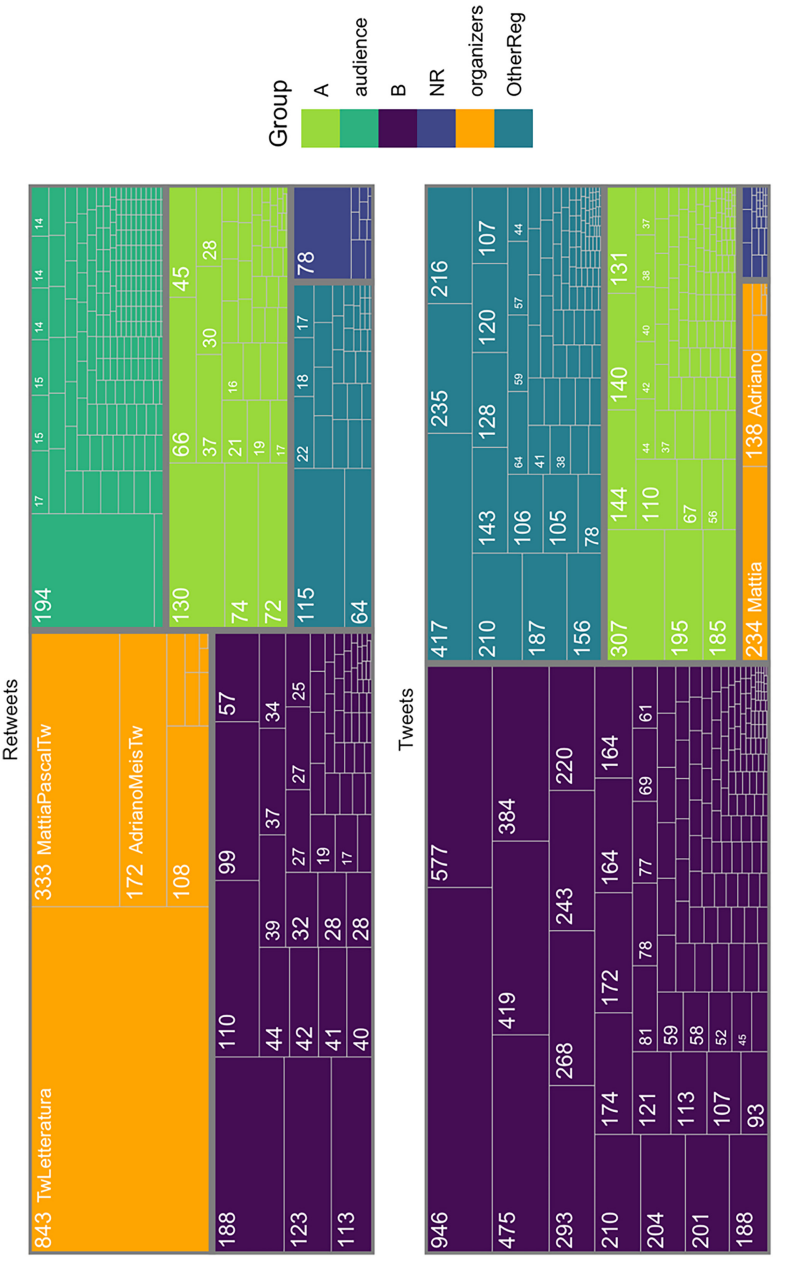

Figure 2.
Number of tweets and retweets generated by each user, hierarchically ordered within the group to which they belong

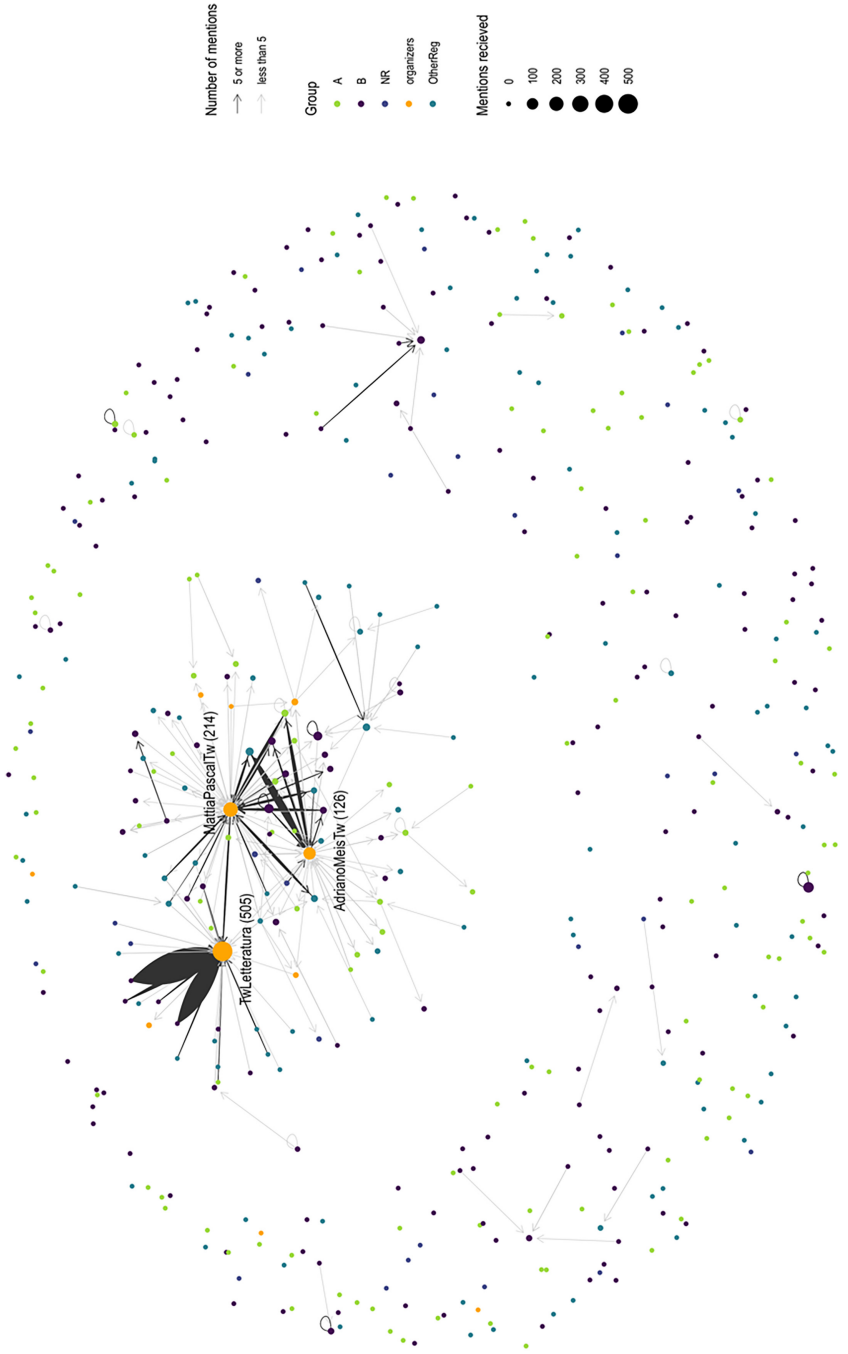

Figure 3.
Social network of #MattiaTw interactions for 1,262 tweets with mentions of users. The size of the nodes is based on the number of mentions received by each author. The width of the connecting edges is proportional to the number of times the originating node mentioned another user

Table 3.
Examples of content categories identified in a sample of tweets ($n = 80$), which were automatically detected as not being text reuse

Category	Frequency	Example
Generalization	25%	Non possiamo comprendere la vita, se in qualche modo non ci spieghiamo la morte. [We cannot understand life if we do not somehow make sense of death]
Interpretation	32%	Adriano si sente in trappola nella gabbia della libertà. [Adriano feels trapped in the cage of freedom]
Intertextuality	5%	#MaledettoSiaAmleto: Copernico ci ha dato la consapevolezza di essere niente, Shakespeare ci ha dato quella di essere nessuno [#DamnedBeAmleth: Copernicus made us realize that we are nothing, Shakespeare that we are nobody]
Reader's reaction	19%	Provo angoscia poichè a breve distanza sono morte le figlie e la madre, Mattia mi fa davvero pena [I'm distressed because in a short time both daughters and mother died, I feel pity for Mattia]
Summary	19%	Adriano Meis sceglie Roma come città dove vivere. Una stanzetta in affitto e poco altro. [Adriano Meis chooses to live in Rome. A rented snuggery and little more]

learner-content and learner-learner interactions (e.g. Tur and Marín, 2015; Luo, 2015). With respect to this, #MattiaTw resulted to be an activity dominated by the participation of the registered classes, with almost no interaction outside the classroom, neither with other students nor with Twitter users. The social dimension of DSR – involving reading other people's response to a text and replying to them – has thus been neglected, probably because of the separation between discussion (in the classroom) and written output (on Twitter). Depending on how the DSR was managed, the sharing of other people's comments could have been either part of the classroom activity or done independently by the teacher managing the account. Therefore, we do not know whether reading other people's comments was part of the learning activity. Relatedly, the decoupling (both in space and time) of text and comments is likely one of the reasons that led users to share for the largest part text excerpts that they deemed significant – either quoting them or trough paraphrases – rather than sharing their reactions to the story or discussing it with others, like it frequently happens with other kinds of DSR (Blyth, 2014; Pianzola *et al.*, 2020; Sun and Gao, 2017; Thoms and Poole, 2018). Moreover, the fact that the majority of the original content generated is an interpretation of the text or a generalization based on it seems to reflect the style and goals of teaching literature in Italian secondary school and, specifically, the novel "The Late Mattia Pascal". Indeed, literary critics and textbooks often focus on Pirandello's ideas about human existence and identity in an epoch of crisis, like it was the beginning of the 20th century in Europe. On a different level, Twitter has been used as a tool to publicly share the outcome of reading acts performed individually or in small groups, similar to sharing highlights or personal annotations (Barnett, 2014; Rowberry, 2016). In contrast, previous research on a similar learning activity, namely asking students to reflect on and tweet about course content, found that 47% of posts were relevant to the content, and 59% were about connecting to other students (Domizi, 2013). However, the small number of retweets and mentions is consistent with other studies about hashtags used for events somehow similar to #MattiaTw, in which the most relevant activity is the posting of "original commentary" and not "information sharing". This kind of social reading is thus very similar to other uses of Twitter that can be considered a "shared experience of 'audiencing'" media events (Bruns and Stieglitz, 2012, p. 177; Fiske, 1992) rather than a proper learning activity. In the case of DSR, the media event is the reading of the original literary text: all classes were reading the same book independently and used Twitter to publicly share their comments.

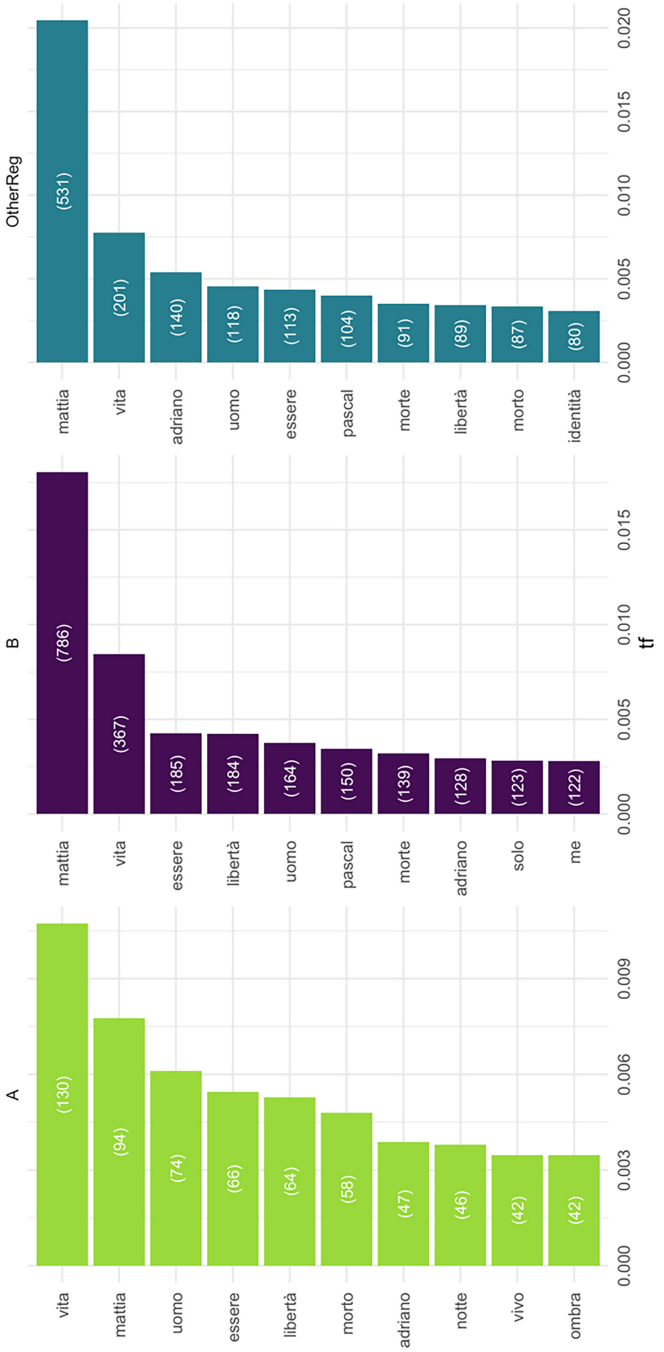

Figure 4.
Most frequent words in the not-text-reuse tweets by groups A, B and OtherReg. Occurrences are expressed as standardized (tf) and absolute values (within brackets). Total tweets are 870 (A), 3,367 (B), 2,012 (OtherReg)

The lack of social interaction can also explain why many users from groups B and Other Reg abandoned the DSR: without feedback from other readers and the related feeling of participating in a communal activity, the motivation to keep tweeting was very feeble. Teachers more familiar with Twitter have probably been more involved from a social point of view and enjoyed the project more. However, considering the 20% reduction of participants, phase 2 was still quite productive in terms of tweets, with only a reduction of sociability for group B (weighted −20% of retweets per chapter and −4% of mentions). But group B was also the most engaged throughout the whole social reading, with an average of tweet per user that increased from 26.9 to 29.6 in phase 2. It is very likely that its users have been influenced by the initial hype for this uncommon learning activity, and this helped to keep the level of engagement higher for three months (cf. Junco *et al.*, 2013). In particular, its performance is mostly due to a few very active teachers and classes. The two-month delay seems to have affected group A, whose number of tweets per user was always lower than groups B and OtherReg. Teachers and students of group A registered at the beginning of the school year but then had to wait until January to take part in it. This might have reduced the initial motivation linked to the engagement with the discovery of a new kind of activity.

More generally, a social reading activity with a predefined shared calendar using a novel seems not to be the best option to involve a general audience because it requires a long time to complete the reading. To maximize the possibility that more users will actively follow the DSR, the shared calendar should not be too tight, but #MattiaTw had to consider also that the main participants were high school students, reading and commenting in the classroom for 1–2 h per week. The result was a three-month long activity for a 250 page novel, which usually can be read in a few weeks. It is quite hard to keep the attention of a group of users who are interacting only via social media on such a long activity. In this regard, the role of TwLetteratura was to encourage the students' participation by directly quoiting them in their tweets and to increase the reach of the comments produced by the students. In order to do that, they commented using two fictional characters' accounts and retweeted a lot of the comments. Their high retweeting activity is coherent with their role as organizers, who had an interest in amplifying the reach of the user-generated content in order to potentially attract more participants and enlarge the community. Indeed, when they retweeted less and focused more on interacting with participants (phase 2), the number of interactions (retweets and mentions) between all participants decreased.

According to Preston *et al.* (2015), students feel more engaged and develop digital identities and digital literacy by sharing via Twitter their reflections on course content. Instead, #MattiaTw was more focused on the critical reading of a literary text, neglecting that digital literacy and communication skills could have been leveraged for engaging students, and in turn be improved by such engagement. Indeed, even though the behavior of the organizers was in line with Tang and Hew's (2017) recommendations, other factors interfered in the engagement process. For instance, providing access to a broader community was not enough because most of the students interacted with them only through the teacher's mediation. Being able to directly interact with participants outside the classroom is an additional guideline to be considered when planning educational uses of Twitter. Otherwise, Twitter's beneficial effect of expanding learning communities and increasing participation and engagement (Gao *et al.*, 2012) may be hindered.

7. Conclusions, limitations and future research

The #MattiaTw project involved many high school classes in what is probably the biggest DSR activity organized in an educational context. However, its outcome in terms of social interactions and reader response was rather limited. Our analysis was able to highlight how an activity presented as educational and highly structured prevented the involvement of spontaneous users and apparently did not change the way students engaged with the text in

comparison to more traditional ways of teaching literature. The participation of nonregistered readers was extremely low, but it suggests that familiarity with the medium (Twitter) and spontaneous participation are related to higher interest toward other readers' response (retweets) and discussion (mentions). However, we are unable to say whether this is an uncommon situation due to the educational context or something recurring on other platforms as well. More studies are needed to evaluate the extent of social interaction, both for other projects by TwLetteratura and for DSR in general. For instance, checking the proportion of replies to Goodreads reviews or Wattpad comments with respect to the total number of reviews or comments.

Another possible limitation is that students and teachers of each class commented the book mainly using a single Twitter account; thus, it was not possible to observe the individual readers' behaviors but only aggregated ones. Qualitative methods would be better suited to explore how this kind of DSR can affect individual reader response.

Engaging in DSR and using social media are two different kinds of activity, although their convergence can affect reading in many ways. Future research will need to pay attention to the difference between social reading happening within content communities (Wattpad, fanfiction, etc.) and social networking sites/applications (Goodreads, Twitter, etc.) (cf. Kaplan and Haenlein, 2010, pp. 62–64). Another important aspect to be considered is that "readers may write reviews and comments in order to achieve social goals rather than to accurately record their opinions" (Thelwall and Kousha, 2017, pp. 982–983). In all aspects, DSR is an extremely interesting phenomenon reflecting the evolution of reading practices in a digital age, and digital methods of analysis could give a crucial contribution to understanding the best way to make sense of this change (cf. Rebora *et al.*, 2021).

Notes

1. http://www.fondazionecariplo.it/en/index.html
2. http://usr.istruzione.lombardia.gov.it
3. Available at https://github.com/spatie/twitter-streaming-api

References

Barnett, T. (2014), "Social reading: the kindle's social highlighting function and emerging reading practices", *Australian Humanities Review*, No. 56, pp. 141-162.

Blyth, C.S. (2014), "Exploring the affordances of digital social reading for L2 literacy: the case of eComma", in Pettes Guikema, J. and Williams, L. (Eds), *Digital Literacies in Foreign and Second Language Education*, Calico, San Marcos, *Texas*, pp. 201-226.

Bruns, A. and Stieglitz, S. (2012), "Quantitative approaches to comparing communication patterns on twitter", *Journal of Technology in Human Services*, Vol. 30 Nos 3–4, pp. 160-185.

Costa, P. (2016), *Il Futuro Della Lettura. L'esprienza Del Testo Nell'era Postmediale*, Egea, Milano.

Dimitrov, S., Zamal, F., Piper, A. and Ruths, D. (2015), "Goodreads versus Amazon: the effect of decoupling book reviewing and book selling", *Proceedings of the Ninth International AAAI Conference on Web and Social Media*, pp. 602-605.

Domizi, D.P. (2013), "Microblogging to foster connections and community in a weekly graduate seminar course", *TechTrends*, Vol. 57 No. 1, pp. 43-51.

Driscoll, B. and Rehberg Sedo, D. (2018), "Faraway, so close: seeing the intimacy in Goodreads reviews", *Qualitative Inquiry*, Vol. 25 No. 3, pp. 248-259, doi: 10.1177/1077800418801375.

Fiske, J. (1992), "Audiencing: a cultural studies approach to watching television", *Poetics*, Vol. 21 No. 4, pp. 345-359.

Gao, F. (2013), "A case study of using a social annotation tool to support collaboratively learning", *The Internet and Higher Education*, Vol. 17, pp. 76-83.

Gao, F., Luo, T. and Zhang, K. (2012), "Tweeting for learning: a critical analysis of research on microblogging in education published in 2008-2011: tweeting for learning", *British Journal of Educational Technology*, Vol. 43 No. 5, pp. 783-801.

Hegel, A. (2018), *Social Reading in the Digital Age*, UCLA, Los Angeles.

Junco, R., Elavsky, C.M. and Heiberger, G. (2013), "Putting twitter to the test: assessing outcomes for student collaboration, engagement and success", *British Journal of Educational Technology*, Vol. 44 No. 2, pp. 273-287.

Kalir, J.H. and Garcia, A. (2019), *Annotation*, MIT Press Open, Boston, available at: https://mitpressonpubpub.mitpress.mit.edu/annotation (accessed 24 July 2020).

Kalir, J.H., Morales, E., Fleerackers, A. and Alperin, J.P. (2020), "'When I saw my peers annotating': student perceptions of social annotation for learning in multiple courses", *Information and Learning Sciences*, Vol. 121 Nos 3/4, pp. 207-230, doi: 10.1108/ILS-12-2019-0128.

Kaplan, A.M. 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, Elsevier.

Kutzner, K., Petzold, K. and Knackstedt, R. (2019), "Characterising social reading platforms. A taxonomy-based approach to structure the field", *Proceedings of the 14th International Conference on Wirtschaftsinformatik*.

Luo, T. (2015), "Instructional guidance in microblogging-supported learning: insights from a multiple case study", *Journal of Computing in Higher Education*, Vol. 27 No. 3, pp. 173-194.

Marshall, C.C. and Brush, A.J.B. (2004), "Exploring the relatinship between personal and public annotations", *Proceedings of the 4th ACM/IEEE Joint Conference on Digital Libraries*, June 2004, pp. 349-57.

Murray, S. (2018), *The Digital Literary Sphere: Reading, Writing, and Selling Books in the Internet Era*, Johns Hopkins University Press, Baltimore.

Nakamura, L. (2013), "'Words with friends': socially networked reading on Goodreads", *Pmla*, Vol. 128 No. 1, pp. 238-243, doi: 10.1632/pmla.2013.128.1.238.

Page, R. (2012), *Stories and Social Media*, Routledge, London.

Pianzola, F., Rebora, S. and Lauer, G. (2020), "Wattpad as a resource for literary studies. Quantitative and qualitative examples of the importance of digital social reading and readers' comments in the margins", *PloS One*, Vol. 15 No. 1, doi: 10.1371/journal.pone.0226708.

Pianzola, F., Toccu, M. and Viviani, M. (2021), "MattiaTW digital social reading", *OSF*, doi: 10.17605/osf.io/mfqh6.

Preston, J.P., Jakubiec, B.A.E., Jones, J. and Earl, R. (2015), "Twitter in a Bachelor of Education course: student experiences", *LEARNing Landscapes*, Vol. 8 No. 2, pp. 301-317, doi: 10.36510/learnland.v8i2.710.

R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, available at: http://www.R-project.org/.

Rebora, S. and Pianzola, F. (2018), "A new research programme for reading research: analysing comments in the margins on Wattpad", *DigitCult - Scientific Journal on Digital Cultures*, Vol. 3 No. 2, pp. 19-36, doi: 10.4399/97888255181532.

Rebora, S., Boot, P., Pianzola, F., Gasser, B., Herrmann, J.B., Kraxenberger, M., Kuijpers, M., Lauer, G., Lendvai, P., Messerli, T.C. and Sorrentino, P. (2021), "Digital humanities and digital social reading", *Digital Scholarship in the Humanities*. doi: 10.1093/llc/fqab020.

Rowberry, S.P. (2016), "Commonplacing the public domain: reading the classics socially on the Kindle", *Language and Literature*, Vol. 25 No. 3, pp. 211-225.

Smith, D. (2012), "Dasmiq/passim", *Java,* available at: https://github.com/dasmiq/passim (accessed 14 August 2020).

Smith, D.A., Cordel, R., Dillon, E.M., Stramp, N. and Wilkerson, J. (2014), "Detecting and modeling local text reuse", *IEEE/ACM Joint Conference on Digital Libraries, Presented at the 2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL), IEEE*, London, United Kingdom, pp. 183-192.

Sun, Y. and Gao, F. (2017), "Comparing the use of a social annotation tool and a threaded discussion forum to support online discussions", *The Internet and Higher Education*, Vol. 32, pp. 72-79.

Tang, Y. and Hew, K.F. (2017), "Using Twitter for education: beneficial or simply a waste of time?", *Computers and Education*, Vol. 106, pp. 97-118.

Thelwall, M. (2019), "Reader and author gender and genre in Goodreads", *Journal of Librarianship and Information Science*, Vol. 51 No. 2, pp. 403-430.

Thelwall, M. and Kousha, K. (2017), "Goodreads: a social network site for book readers", *Journal of the Association for Information Science and Technology*, Vol. 68 No. 4, pp. 972-983.

Thomas, B. (2014), "140 characters in search of a story: twitterfiction as an emerging narrative form", in Bell, A., Ensslin, A. and Rustad, H. (Eds), *Analyzing Digital Fiction*, Routledge, New York, pp. 94-108.

Thomas, B. (2020), *Literature and Social Media*, Routledge, London and New York.

Thoms, J.J. and Poole, F. (2017), "Investigating linguistic, literary, and social affordances of L2 collaborative reading", *Language Learning and Technology* `, Vol. 21 No. 2, pp. 139-156.

Thoms, J.J. and Poole, F.J. (2018), "Exploring digital literacy practices via L2 social reading", *L2 Journal*, Vol. 10 No. 2, pp. 36-61.

Tur, G. and Marín, V.I. (2015), "Enhancing learning with the social media: student teachers' perceptions on Twitter in a debate activity", *Journal of New Approaches in Educational Research*, Vol. 4 No. 1, pp. 46-53, doi: 10.7821/naer.2015.1.102.

TwLetteratura (2016a), *Pirandello, Finalmente*, Twletteratura.Com, available at: https://www.twletteratura.org/2016/05/pirandello-finalmente/ (accessed 20 February 2018).

TwLetteratura (2016b), *#MattiaTw, Come Funziona Il Progetto*, Twletteratura.Com, available at: https://www.twletteratura.org/2016/12/valutazione-mattiatw-cariplo/ (accessed 20 February 2018).

TwLetteratura (2017), *The TwLetteratura Method*, Twletteratura.Com, available at: https://www.twletteratura.org/2017/03/the-twletteratura-method/ (accessed 18 February 2018).

Vázquez Cano and Esteban (2012), "Mobile learning with twitter to improve linguistic competence at secondary schools", *New Educational Review*, Vol. 29 No. 3, pp. 134-147.

Vesanto, A., Nivala, A., Rantala, H., Salakoski, T., Salmi, H. and Ginter, F. (2017), *Applying BLAST to Text Reuse Detection in Finnish Newspapers and Journals, 1771–1910*, p. 5.

Appendix

The appendix files are available online for this article.

Corresponding author

Federico Pianzola can be contacted at: federico.pianzola@unimib.it

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**