HERMES
Journal of Language and Communication in Business

Chiara Polli & Carmen Serena Santonocito***

Reputation at Risk: Sentiment Analysis and Social Media Listening Tools under the Lens of Critical Multimodal Discourse Studies[1]

Abstract
In the field of online crisis management communication, AI-based social media listening tools (i.e., tools designed to track and monitor online conversations about a topic or brand) play a pivotal role in opinion mining and reputation audits. Compared to manual analyses, AI enables a faster large-scale collection and classification of vast amounts of data from several online platforms, thus facilitating the task of detecting and monitoring the sentiment linked to a brand and/or product. Nonetheless, AI-based analyses are far from unbiased. This paper adopts a critical multimodal perspective to explore the challenges linked to sentiment analysis in digital multimodal aggregations. It presents an empirical study in which the results of sentiment analysis performed by the social media listening tool Meltwater and manual tagging are compared to evaluate the efficacy of the tool in assessing online reputation damage following a crisis event. The findings of this study suggest that, in addition to possible incorrect classifications of texts (e.g. lack of understanding of pragmatic features or texts in languages other than English), the AI-based tool's misinterpretation of emotional cues also includes multimodal ensembles. This is due to social media listening tools' dependence on unimodal (verbal-only) classifiers that fail to produce reliable outputs. Despite the predictive power of these tools, the findings ultimately indicate that the accuracy of sentiment analysis is still affected by a hard-to-die bias concerning the primacy of language over non-verbal communication – a trend which is in contrast with the multimodal nature of semiosis and the proliferation of complex multimodal artifacts online.

Keywords
multimodality, sentiment analysis, critical multimodal discourse studies, artificial intelligence, online crisis management communication, social media listening tools.

1. Introduction
Today's digital mediation of knowledge and discursive practices (Moschini & Sindoni, 2022) implies that technologies provide tools and affordances to create and share a vast array of meaningful multimodal wholes (Martinec & Salway, 2005). Among the latter, representations, evaluations, and emotions are disseminated across diverse online platforms (e.g., social media, blogs, forums, e-commerce reviews) and, as such, play a crucial role in influencing how people evaluate the world and perceive companies and society as a whole (Liu, 2012). Since "the growing popularity of the Internet has lifted the web to the rank of the principal source of universal information" (Wankhade *et al.*, 2022: 5731), the proliferation of user-generated content related to personal opinions and emotions has been impacting on people's perceptions of reality and decision-making processes (Cui *et al.,* 2023). Given that emotions build social relationships, (Cambria *et al.*, 2017) gauging how these are expressed in discourse is central to monitor public opinion and decision-making processes *vis-à-vis* brand reputation and crisis management communication. Within this context, online crisis management communication (OCMC henceforth) refers to the strategic use of communicative resources that can prevent or limit the financial or reputational harm of a crisis event (Coombs, 2019). In a scenario that increasingly needs to improve corporate communication and reputation management, sentiment analysis (SA henceforth) is an application of text-based analytics that

[1] Although this article is the result of a collaborative research effort, authorship can be attributed as follows: Chiara Polli is responsible for Sections 3, 4, and 6; C. Serena Santonocito is responsible for Sections 1, 2, and 5.

* *Chiara Polli*
Department of Humanities - Ancient and Modern Languages, Literatures and Civilizations
University of Perugia
chiara.polli@unipg.it

** *Carmen Serena Santonocito*
Department of Ancient and Modern Civilizations
University of Messina
carmenserena.santonocito@unime.it

enables the processing of big data on people's attitudes, emotions, evaluations, and opinions (Liang *et al.*, 2022). SA builds on natural language processing and machine learning to respond to the increasing demand of harnessing haphazard data on people's sentiments on social media.

In the last decades, start-ups and corporations have invested in building or outsourcing a variety of tools for automated SA (Eriksson, 2018). These tools are a fundamental staple of current business and government intelligence as their application can monitor rapidly escalating hostile communication (Lanfranchi, 2017).

The continuous advancement of technology and participatory platforms (KhosraviNik & Unger, 2016) has spurred the popularity of SA in a variety of fields, such as business, finance, politics, and education (Cui *et al*., 2023). SA has garnered widespread attention from governments, institutions, and companies (Khatua *et al.*, 2020) because it accelerates decision-making processes, the identification of challenges, and public opinion hotspots. The result has been the proliferation of social media listening (SML henceforth) tools which harness AI-powered algorithms to monitor online sentiment in real time and to support decision-making in today's digital marketing campaigns (Perakakis *et al.*, 2019). Since these tools usually gauge spontaneous conversations on social media, their use is becoming crucial in OCMC to face crisis events which can potentially damage online visibility and reputation (Al-Ghamdi, 2021).

However, capturing potential customers' sentiment in the spontaneity of semantically and semiotically articulated social media conversations comes with a cost. In this respect, the majority of user-generated sentiment is made up of unstructured verbal text combined with different modalities, which makes AI-driven SA problematic. Recent studies describe different issues on the effective detection of sentiment, which is often compromised by socio-cultural aspects – e.g., opinion spam and fraudulent reviews (Jain *et al.*, 2021b) – and language and discourse-related aspects – e.g., code-mixing (Vijay *et al.*, 2018), use of languages other than English (Angel *et al*., 2021), and multimodal combination of different modes (Upadhyaya *et al.*, 2023). These studies have mainly tackled questions from the perspectives of computer science and business communication, but linguistic and discursive aspects have been overlooked, or have focused on, mostly unimodal SML algorithms, that are language-alone datasets with little or no training as far as other semiotic resources are concerned. Consequently, SML algorithmic models are likely to miss nuances in language use, as well as the full complexity of human discursive practices (Duarte *et al.*, 2017). The lack of a more wide-ranging approach to the pragmatics of language in use may in fact lead to misinterpretation of product evaluations and faults in web-care strategies (Catenaccio, 2021) .

This paper adopts a Critical Multimodal Studies (CMS henceforth) perspective (Djonov & Zhao, 2014) to analyze and discuss an empirical case study that employs Meltwater, an AI-powered SML tool, which performs, *inter alia*, AI-based SA. We report results of a case study (Santonocito & Polli, 2023) carried out during our research period (March-September 2022) at Digital Trails to address an OCMC event. Starting from real-world data, we compare Meltwater's SA results with our manual tagging. In particular, we focus on the analysis of sentiment conveyed by multimodal combinations of verbal, visual, and aural resources with the aim to provide evidence of problematic sentiment interpretation from a discourse-oriented perspective.

2. AI-powered SA and OCMC

A Critical multimodal approach may help capture the complexities of fluid digital platforms that incorporate a wealth of semiotic resources while, at the same time, evolving over short time spans. Indebted to Foucault (1972) and to the social semiotic tradition (Kress, 2010), the notions of texts and discourses are here intended as broadly encompassing semiotic modes that comprise but are not limited to verbal language (Jewitt *et al.*, 2016). The Critical approach addresses the continuous engagement of discourses in shaping collective understanding about social practices and power relations that regulate knowledge production (Fairclough & Wodak, 1997).

Discourses are ideologically driven and socially created in transient digital landscapes where appeal to emotions and intensified subjectivity infuse empowerment and like-mindedness to the point that affect and emotions in argumentation models have been shifting (KhosraviNik, 2018). This shift has been consistently discussed in affective communication studies in the last decade (Papacharissi, 2015). This is especially the case for social media, where the inventory of emotive impressions and subjectivity favours the creation of polarized sentiments, which may emerge in and across the combination of different semiotic resources (e.g., text, image, proxemics cues) orchestrated in meaningful patterns (Halliday, 1978). In this respect, the mono-modal training of SML algorithms is problematic because AI protocols tend to ignore the simultaneous combination of semiotic resources other than the verbal mode, as well as their inability to interpret context-bound discourse practices (Sunderland, 2004).

Notwithstanding such problematic aspects, SA is attracting attention in OCMC in response to the need to analyze what people think and feel (Kotras, 2020). In this respect, the large amount of scrutinizable digital data and the explosion of online expression of judgement put corporate image and brand reputation at risk (McDonald & Slawson, 2002). This means that understanding and measuring online opinions is of pivotal importance. To this end, SA makes use of text-based analytics to gauge people's opinions and attitudes (Liang *et al.*, 2022) as organized in the polarity of the sentiment, generally expressed as positive, neutral, or negative (Lei & Liu, 2021).

From an architectural perspective, natural language processing and machine learning are used to extract information on polarized sentiment from digital data. These learn from models constituted by labelled training data that harness, among others, word ngrams, valence association lexicons, negations lists, and embeddings-based features (Mohammad, 2017). Since the more affective data they receive the more accurate they become (Calvo & D'Mello, 2010), natural language processing and machine learning algorithms are often trained with verbal data from online reviews to determine the polarity of specific objects or events (Liu, 2012).

Unlike traditional surveys, SML tools consist of in-built software infrastructures, which parse web content on the basis of customizable instructions and Boolean searches (Hayes *et al.*, 2021). The latter contain simple operators, such as 'and', 'or', and 'not', used to combine or exclude keywords in a search. Generally speaking, the AI-driven sentiment features real-time dashboards to monitor online visibility (Avery, 2017) and to detect online criticism and potential reputation crises (Kotras, 2020).

From an OCMC perspective, SML tools can be crucial in the development of adequate web-care strategies intervening in every phase of crisis management. Attending to one or more of such phases, SML tools may enable monitoring of online interaction between (potential) clients and the company, tracing possible crisis triggers, and managing online reputation and visibility (Buzoianu & Bîră, 2021).

As argued beforehand, the vast amount of meaning-making resources in multimodal digital texts (Sindoni, 2013) constitutes a further challenge to the current lack of multimodally-informed protocols for SA in OCMC. In this respect, various studies (Mohammad, 2017; Majumder *et a*l., 2018; Schwaiger *et al.,* 2021) highlight the problematic aspects of SML platforms in determining the sentiment of digital data. According to these studies, such complexities include, but are not limited to, sentiment assessment at various textual levels (e.g., terms, sentences, argumentative structure etc.); distinction between sentiment expressed by text producers *vs.* other actors; various pragmatic and rhetorical levels used to convey sentiment (e.g., implicitness, sarcasm, metaphors). Overall, these complexities stem from the unstructured nature of social media discourses, which occur in free-flowing environments, where users are not constrained in opinion making and expression (Levashina *et al.*, 2014). Consequently, the detection of the real sentiment may be obscured because such unstructured data contain more information than the mere polarized sentiment the algorithm is trained to detect.

Multimodal fusion models may counter present flaws with a view to reaching a satisfactory recognition of the multimodal component of sentiment formation. Unlike traditional models that take as input single modalities (e.g., verbal, visual, or aural resources) in multimodal fusion models different modes are considered holistically in the training process (Majumder *et al.*, 2018) for SA. Although these models are a rapidly expanding field, they present a number of challenges that require further exploration since their success rate is in current need of consistent improvement (Lai *et al.,* 2023).

Given the current challenges faced by SML tools in the interpretation of multimodal contents the research questions (RQs henceforth) of this study are as follows:

RQ1: what are the main interpretation issues of Meltwater when SA is applied to multimodal contents and therefore depends on the multimodal orchestration of different semantic resources?

RQ2: what are current challenges that need to be addressed for improving the output of AI-based SA?

3. Case-study construction and methodological framework

To tackle our RQs, we designed a case-based empirical study (Jewitt, 2016; Bateman *et al*., 2017) in which we adopted a blended methodological approach that combined an AI-based SA with manual tagging, as well as qualitative observations applied to *real-world* data.

We collected our dataset during a 6-month fieldwork experience at Digital Trails, an international SME based in the Netherlands that provides PR coverage and reputation management services for business-to-business and business-to-consumer clients. We collaborated with Digital Trails's core team of marketing specialists to perform several online visibility and reputation management activities, including some OCMC tasks. Our case study is part of an online reputation audit conducted between April and June 2022. The client was an international company operating in the fermentation product market (here anonymized). The crisis event involved the sale of a defective batch.

In light of the potential reputational damage, Digital Trails was asked to overview the sentiment of relevant online mentions of the client and find any potential negative comment about the incident in its three main markets (i.e., UK, US, France). Based on the findings of this analysis, Digital Trails's ultimate goal was to propose a plan for future actions aimed at recovering the potential damage and improving the client's online visibility. To do so, Digital Trails needed to accurately scrutinize online conversations about the client's brand and products and assess possible reputational risks.

For this task, Digital Trails employed the AI-powered SML tool Meltwater, listed among the business analytics instruments most frequently employed (Ziora, 2016). Meltwater can extract data across a wide variety of publicly available sources, including blogs and news outlets (e.g., New York Times, BBC, CNN, local newspapers, online journals), social media (e.g., Instagram, Facebook, Reddit, TikTok, Pinterest, and Twitter – now rebranded as X), and the video live streaming and sharing platforms YouTube and Twitch. Searches can be refined by using Boolean operators and the filtering bar, which singles out results according to source types (e.g., online newspapers like the New York Times *vs* social media platforms like Facebook), language (language detection for 242 languages and full SA for 28 languages are provided),[2] location, multiple keyword combinations, sentiment, author (e.g., lists of journalists or influential social media users can be created), and other customized categories. Digital Trails followed a four-step procedure to conduct the online reputation audit:

1. Keyword selection, including the brand's and the faulty products' names, along with some generic search terms (e.g., *yeast for beer/levure de bière pour brassage*) in the official languages of the target markets (i.e., English and French).

[2] https://help.meltwater.com/en/articles/4064558-how-is-sentiment-assigned (last access 12.01.2024)

2. Data gathering and filtering via Meltwater. Data were filtered according to time (12-month period after the crisis event) and location (UK, US, France). No language filter was applied to incorporate potential mentions in languages other than English and French.
3. SA via Meltwater and manual tagging.
4. Manual search engine result page analysis to detect potential negative mentions in prominent positions on Google search's first two pages.

The resulting online reputation audit was then employed by Digital Trails's experts as a starting point to elaborate a plan to boost the client's online visibility. In this respect, a fine-grained analysis of the data collected was required. For this reason, we were directly involved in the SA task (Step 3), which we used as a testbed to evaluate the performance of an AI-based tool (Meltwater) in comparison to manual tagging and finetuning of the final assessment. Meltwater's SA uses natural language processing to classify items into 'positive', 'negative', 'neutral', or 'not-rated' if the system fails to recognize them. Meltwater's classification process is based on a polarity score: "If something is both positive and negative, but more positive, the polarity score would rank it as positive" (help.meltwater.com). It is also trained for emoticon and emoji detection. Owing to its proprietary nature, further information on data processing and sentiment modelling cannot be disclosed.

Regarding Digital Trails's task, Meltwater collected and automatically categorized 4,804 entries according to their sentiment. We then manually refined the results by eliminating double entries. The resulting dataset consisted of 2,567 unique items in 15 languages. Most entries were in English (95.2%) and only 0.7% of results were in French. The remaining items (3.4%) were in Spanish, Portuguese, Moldavian, German, Russian, Swedish, Italian, Chinese, Polish, Czech, Finnish, Gujarati, and Turkish. The language was unknown (0.7% items) in case of contents that Meltwater failed to recognize.

The results were retrieved from a wide range of sources and exemplified different digital text types, spanning from Twitter posts, YouTube videos and vidcasts, 6parks.news and Player.fm podcasts, recipe and personal blog posts, brewing industry-related magazines and journals, suppliers' webpages, buying guides, review websites, and the client's website pages.

We were asked to manually review and tag for sentiment all items retrieved and automatically classified by Meltwater. In addition to the categories of 'positive', 'negative', and 'neutral', Digital Trails instructed us to use two additional labels that Meltwater was unable to discern automatically: 'not accessible' in case of broken links, and 'not relevant' for spam or mentions related to contents that did not refer to the client's brand and products. To conduct our tagging and improve inter-rater reliability, we split our task into further three steps:

A. Independent tagging performed in parallel by us and a third rater from Digital Trails's senior team of OCMC experts.
B. Comparison of the three manual classifications and creation of a unique manually-tagged dataset (in cases of disagreement between individual taggings, Digital Trails's expert attribu-ted the sentiment according to the agency's in-house decision-making strategies).
C. Comparative analysis of the final manually-tagged dataset with the output of AI-based SA and qualitative content analysis (Schreier, 2012) aimed at disambiguating and categorizing potentially recurring criticalities.

Our manual tagging and content analysis draw observations on social semiotics (Kress, 2010; Kress and van Leeuwen, 1996/2020), multimodal critical discourse analysis (Machin, 2013), and multimodal pragmatics (O'Halloran *et al.,* 2014).

An overview of the main findings of this process is provided in Section 4, where criticalities linked to multimodality are also analyzed (Section 4.1.).

4. Data analysis and key insights

An overview of the results obtained through Meltwater's SA and our manual tagging is provided in Table 1. Meltwater identified 161 positive, 2,358 neutral, and 12 negative results, while 36 items were not rated. Conversely, in our manual tagging, we detected 179 positive, 2,069 neutral, 4 negative, and 59 irrelevant results, while 256 items were not accessible. Positive mentions comprised reviews, promotional contents, YouTube videos, vidcasts, podcasts, blog entries and Twitter posts in which users exchanged information on the benefits of the client's products and favourable experiences with the brand. Neutral mentions mainly referred to recipes in which the client's products were listed among the ingredients. Negative mentions included Twitter posts and online magazine entries which explicitly referred to the crisis event. Our decision was guided by the context which helped disambiguate the products' and the client's mentions. As mentioned beforehand, in ambiguous cases, Digital Trails's expert made the final decision.

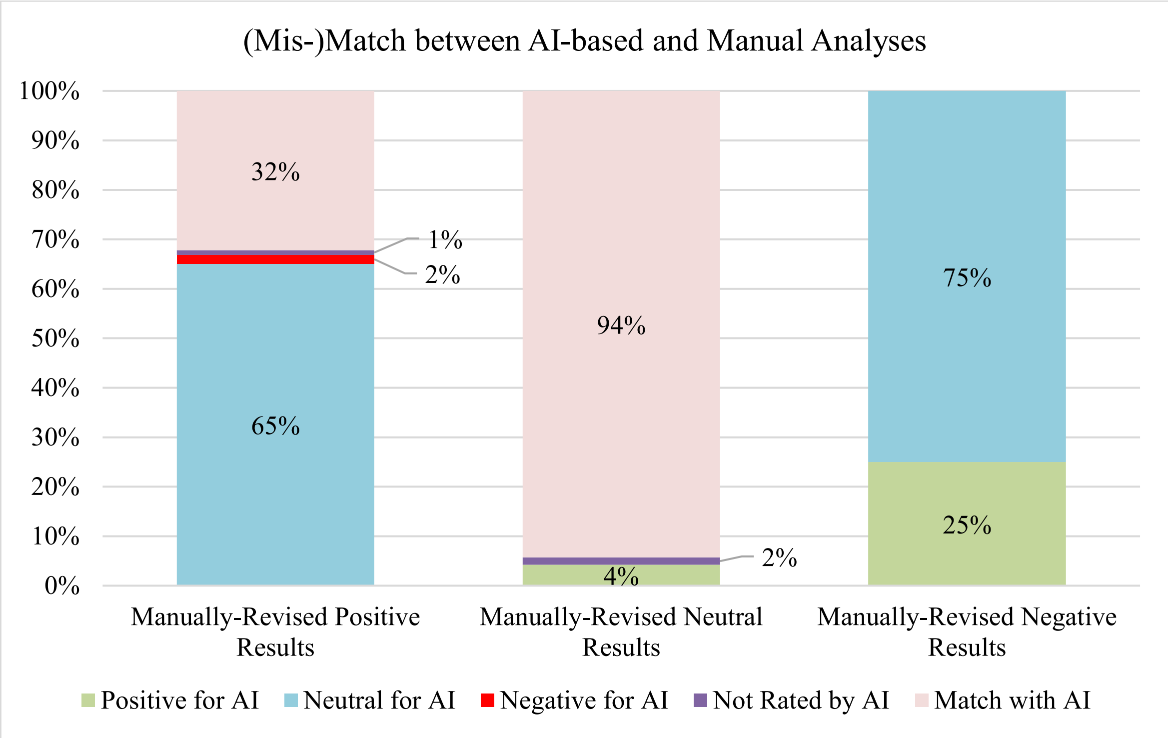
Label	Meltwater's Results	Manual-Tagging Results
Positive	161	179
Neutral	2,358	2,069
Negative	12	4
Not Rated	36	-
Irrelevant	-	59
Not Accessible	-	256

Table 1. Overview of the results of AI-based SA and manual-tagging

If the results of Table 1 are compared solely from a quantitative viewpoint, the differences between Meltwater's performance and the manual tagging appear negligible. In terms of OCMC, the findings seem to indicate minor concerns. However, to plan future actions aimed at improving the client's online reputation and visibility, Digital Trails required a precise and detailed classification of the entries. In this respect, the comparative analysis of the output of AI-based SA and our manually-tagged dataset revealed several discrepancies as there was no full agreement between what was labelled positive, negative, and neutral by Meltwater and by us.

The partial mismatch between AI-based and manually-tagged results is shown in Graph 1.[3] Meltwater performed poorly for both positive and negative mentions, whereas the neutral mentions coincide for the most part with our tagging (i.e., 94%) because they mainly come from referential mentions of the product's names in recipes. Meltwater misclassified 68% of positive mentions for complex pragmatic features (e.g., implicit messages) or lack of context, use of slang or languages other than English, and multimodal emotional cues (e.g., in the presence of emojis, pictures, videos, and podcasts). Negative mentions were even more problematic on the same grounds as the classification of positive mentions. For example, the four negative mentions we identified (including explicit references to the crisis event) were all misclassified by Meltwater. This was especially important in the context of brand reputational recovery, as we were asked to look into potential negative mentions. The discrepancy between our tagging and Meltwater's classification highlighted that the SML tool is partially unreliable and still requires a manual revision of the results. In the next Section, we will give some qualitative examples of recurring criticalities in the misinterpretation of emotional cues conveyed by multimodal ensembles.

[3] Results were rounded up to the nearest whole number. Non-accessible and irrelevant results were excluded from the count.



Graph 1. (Mis-)Match between the results of AI-based SA and manual tagging

4.1. Critical multimodal analysis of results

The dataset was qualitatively analyzed by focusing on the entries where a mismatch between the manual and SA classifications emerged. Criticalities in SA can be loosely grouped into the following areas:

1. contents in languages other than English,
2. contents that require contextual and pragmatic knowledge, and
3. multimodal contents.

The areas in 1) and 2) have been discussed in Santonocito & Polli (2023), hence in this paper, we address the criticalities in the interpretation of emotional cues in multimodal contents. In particular, Section 4.1.1. discusses cases in which the sentiment is conveyed by word-image combinations, while section 4.1.2. focuses on aural and video resources and the client's website pages.

4.1.1. Sentiment cues conveyed by word-image combinations

In our dataset, the sentiment was frequently expressed by a combination of verbal texts and visual resources, particularly pictures, emojis, and star-rating systems. The following examples show how the interplay between the visual and verbal modes may prove problematic for AI systems that rely on unimodal encoders (Polli & Sindoni, 2024), thus leading to potential errors in the interpretation and classification of multimodal items. In particular, these instances illustrate how Meltwater ultimately failed to properly categorize the sentiment of a series of positive tweets, which were labelled as neutral or negative.

4.1.1.1. Word-picture combinations

The Twitter post in Figure 1 exemplifies how interpreting meanings conveyed by an interplay of verbal texts and pictures may be tricky for an AI that is not trained to analyze multimodal ensembles.

In Figure 1, a user describes their home-brewing experience with and without the client's product. While verbally recounting the experience, they show the benefits of the client's product by posting three pictures comparing the beers crafted by using such product and those without it (e.g., the difference in foam). Although the only potential keyword for polarity in the text is "benefit", the series of pictures contribute to amplifying the positive sentiment of the post as they display a clear difference in the features of the two beers. The combined interpretation of text and images is straightforward for humans, who can easily classify the message as positive. However, Meltwater failed to recognize such interaction and categorized the item as neutral.

Figure 1. Word-picture combination: positive sentiment rated neutral by AI

4.1.1.2. Word-emoji combinations

Our dataset analysis indicated that sentiment was frequently conveyed by emoji/word combinations. Emojis are considered crucial emotional cues (Evans, 2017). In our case, their presence was fundamental to determining the sentiment when verbal messages were ambiguous, had no polar words, or required the understanding of implicit meanings. Since Meltwater is trained for emoji detection, we expected them to be considered in SA. Nonetheless, the qualitative examination of results suggested that Meltwater in fact performed poorly in emoji categorization. In particular, the example shown in Figure 2 is useful to illustrate how the AI failed to recognize all visual elements providing sentiment cues and even classified the mention as negative, despite the absence of polar words indicating negativity.

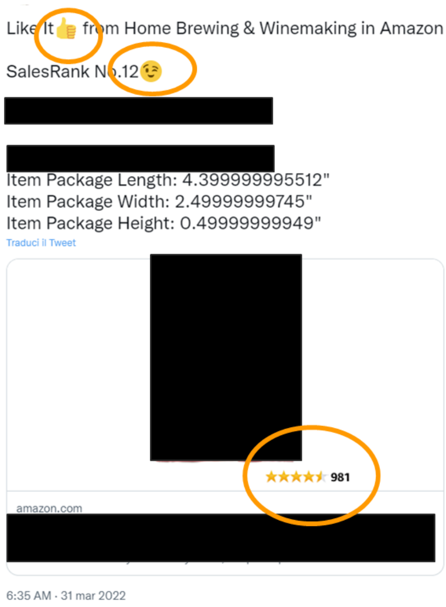


Figure 2. Word-emoji-star-rating combination: positive sentiment rated negative by AI

Figure 2 exemplifies a recurring multimodal combination of our dataset, that is the use of words, emojis, and star-rating system (here highlighted in orange circles). The latter was frequently found in buying-guides, e-commerce retailer pages, and promotional tweets, either to inform on the users' appreciation of a certain product or to convey personal opinions and product rankings. As for Figure 2, a positive sentiment is conveyed by the 'like' (combined with the "like it" text) and the winking-face emojis, as well as by the 4+ star-rating. Hence we considered the mention positive. However, we found that Meltwater classified all the occurrences of star-rating systems as neutral or negative mentions. Unfortunately, this is the only figure we can include in this paper since all other tweets and web pages are no longer available. In several cases, Twitter users deleted their accounts and reviews, promotions were erased, or websites and blogs were shut down.

Figure 3. Word-emoji-picture combination: positive sentiment rated neutral by AI.

In this study, we also observed that Meltwater's poor performance in emoji recognition was often connected to difficulties in interpreting subtle and implicit references to a given sentiment. In Figure 3, for instance, a user (User1) retweets another user's (User2) post. The latter is written in an informal register and includes a mention of the client's brand and products (defined as "goodies") and a picture of the specific "goodies" employed. User2's message ends with "Let's hope that means I'm a busy boy!", thus relating the purchase of the products with his brewing activity. User1's message consists of a single smiling face emoji which conveys a positive sentiment. Therefore, we tagged the item as positive. Conversely, Meltwater classified it as neutral, thus indicating that the emoji was disregarded and the implicit positiveness underlying the retweeted post was not identified.

4.1.1.3. Combinations of emojis and verbal texts in languages other than English

SA proved even more challenging in cases where visual devices like emojis combined with verbal texts in languages other than English. In particular, we found that the combined use of French (i.e., the official language of one of the client's target markets) and non-verbal sentiment cues were not recognized in their reciprocal interplay by Meltwater, which was unable to classify the overall message properly.

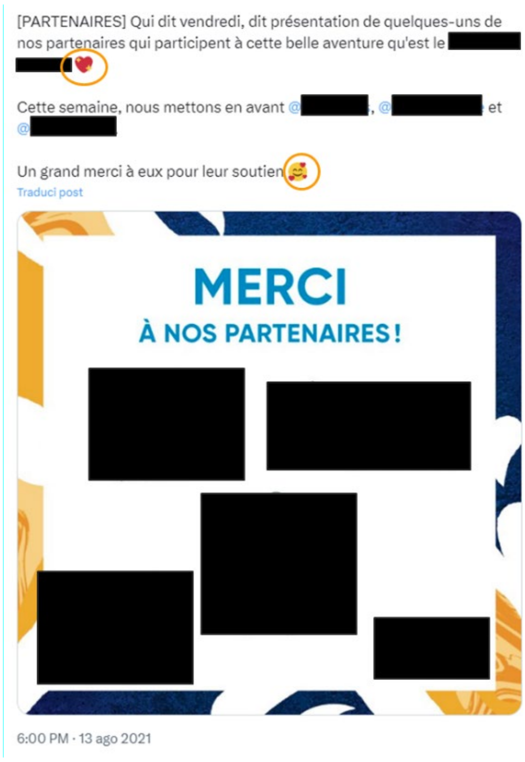


Figure 4. French text-emoji combination: positive sentiment rated neutral by AI.

For instance, Figure 4 displays a tweet in French, in which the user thanks the client for sponsoring a brewing festival. The positive sentiment is conveyed by the two emojis with a sparkling heart and a smiling face with hearts and the verbal message in French included in the tweet ("[PARTENAIRES] Qui dit vendredi, dit présentation de quelques-uns de nos partenaires qui participent à cette belle aventure qu'est le ***/// Cette semaine, nous mettons en avant ***, *** et ***. /// Un grand merci à eux pour leur soutien"; English translation: "[PARTNERS] Fridays are synonymous with the presentation of some of our partners involved in the wonderful adventure that is the ***. // This week, we highlight ***, *** and ***. A big thanks to them for their support") and in an image attached to the tweet ("Merci à nos partenaires!"; English translation: "Thanks to our partners!"). The classification of the user's positive attitude is not particularly challenging, since the emojis are unambiguously positive and keywords such as "Merci" ("Thanks"; repeated twice) and "belle aventure" ("wonderful adventure") facilitate the interpretation. Nonetheless, the tweet is categorized as neutral by Meltwater.

4.1.2. Sentiment cues conveyed by aural and video resources

In our dataset, the sentiment of a given entry frequently depended on the interaction between verbal, visual, and aural modes. The following examples show how the interpretation of such multimodal ensembles proved problematic for AI in several cases, including podcasts, vidcasts, and videos on platforms such as Player.fm, 6park.news, and YouTube or incorporated in Twitter posts, blogs, and web pages.

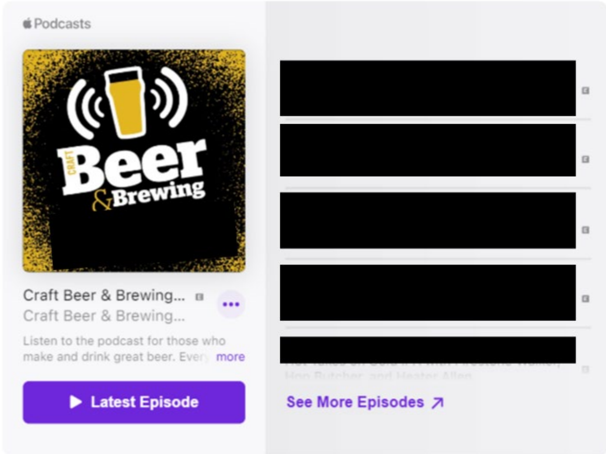
4.1.2.1. Word-emoji-video combinations

Figure 5 shows a multimodal combination in which, in addition to the verbal text and the sequence of emojis, a video is incorporated into a tweet. The video is a neutral shooting of a fermentation process. The sound is verbally described in the text of the post in positive terms (i.e., "the sound of little yeastie hard at work"). The hypocoristic "little yeastie" and the slang expression "hard at work" guide SA’s output. The positive rating is also reinforced by emojis (the smiling face with sunglasses and the hang loose emoji combined with that of the beer). In addition to the pragmatic complexities of the text and its reference to the video, Meltwater did not detect the sequence of positive emojis and ultimately categorized the post as negative – probably considering “little” and “hard” as negative keywords and neglecting their context of use and the overall multimodal nature of the post.

Figure 5. Word-emoji-video combination: positive sentiment rated negative by AI.

4.1.2.2. Podcasts

As for podcasts, Digital Trails’s client sponsored several podcast series about the brewing world on Player.fm and 6park.news. These entries were characterized by an interplay of verbal and aural resources. We labelled all these results as positive, in light of the positive sentiment conveyed both orally (in the podcast) and in writing (in the captions). Figures 6 and 7 display an example of a podcast interface on 6.park.news and its caption endorsing the client as one of the sponsors.



This episode is hosted by:

Figure 6. Podcast interface on 6pack.News and caption introducing the sponsors.

Support for this episode comes from [redacted] the obvious choice for beverage fermentation. [redacted] is a new all-in-one blend of enzymes and Saccharomyces Pastorianus yeast designed to produce dry, low-carb lager beers. [redacted] is capable of high attenuation at low temperatures, making it the perfect choice for clean blonde beers with very low residual sugars. To harness the power of [redacted] or to find out more about [redacted] range of fermentation solutions, contact [redacted]

Figure 7. Podcast caption: positive sentiment rated neutral by AI.

In the audio, the content-creators positively promoted the client's brand, e.g., by describing their yeast as the "perfect solution" for homebrewers. Likewise, in the caption, the brief promotional text includes explicit endorsements of the client's brand and products, such as "the obvious choice for beverage fermentation" and "perfect choice for clean blonde beers". As for SA , it is unclear whether Meltwater had been trained to analyze both spoken and written messages or only focused on the captions. However, our analysis showed that it failed to detect positive sentiment in any of the podcast pages (including the explicit endorsement shown in Figure 7) and classified all these mentions as neutral.

4.1.2.3. YouTube contents

Our tagging indicated that vidcasts and videos – in this case, all shared on YouTube – proved especially challenging for Meltwater. Studies on the social semiotic affordances of video-sharing communities (e.g., Sindoni, 2023) highlighted how the analysis of these platforms needs to encompass their complex interweaving of multiple semiotic resources (e.g., speech, gesture, moving images, music, color, among others) as well as to acknowledge their *hyper-intertextuality* (i.e., an ongoing dynamics of video-sharing, commenting, and cross-referencing; Sindoni, 2013). In our manual analysis, we thus examined the contents of videos, captions, and comments to look for sentiment cues. However, in the case of Meltwater's sentiment analysis, we found that all YouTube contents were not rated, thus indicating that the AI struggled with the recognition of such multimodal resources.

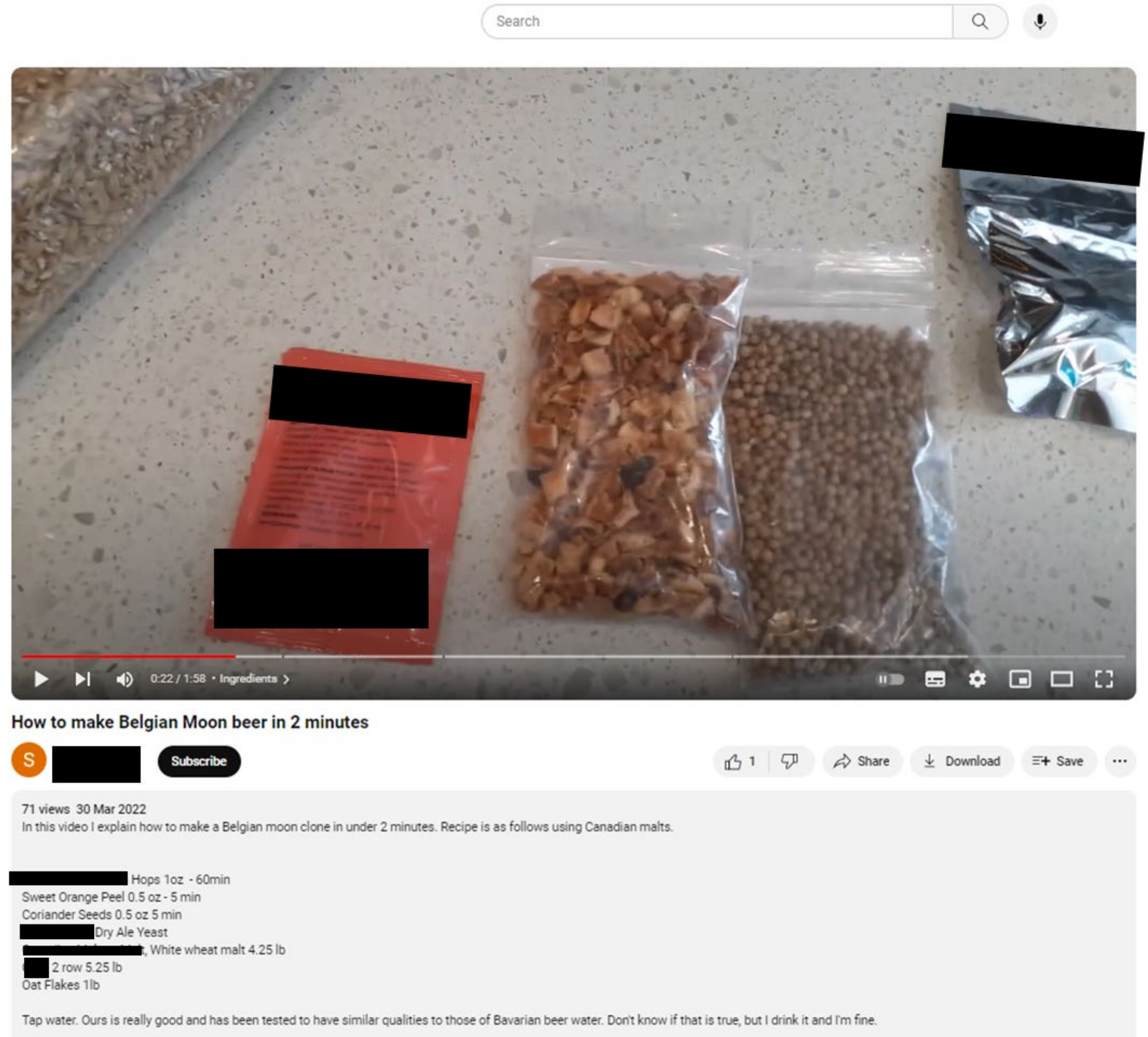


Figure 8. YouTube contents: neutral sentiment not rated by AI.

In our dataset, we found three main types of mentions of the client's brand and products, always included in both the videos and the captions: the brand was mentioned as a sponsor of vidcast series about the brewing world and promoted by the content-creators, while the client's products were listed among the ingredient used in video recipes (e.g., Figure 8). These occurrences were labelled as 'positive' and 'neutral', respectively. In one case, both brand and products were positively mentioned in the recording of a symposium (Figure 9) in which a brewing specialist (collaborating with the client) provided a "practical sensory session to highlight the impact of yeast selection on hop flavour and aroma" (see caption Figure 10).

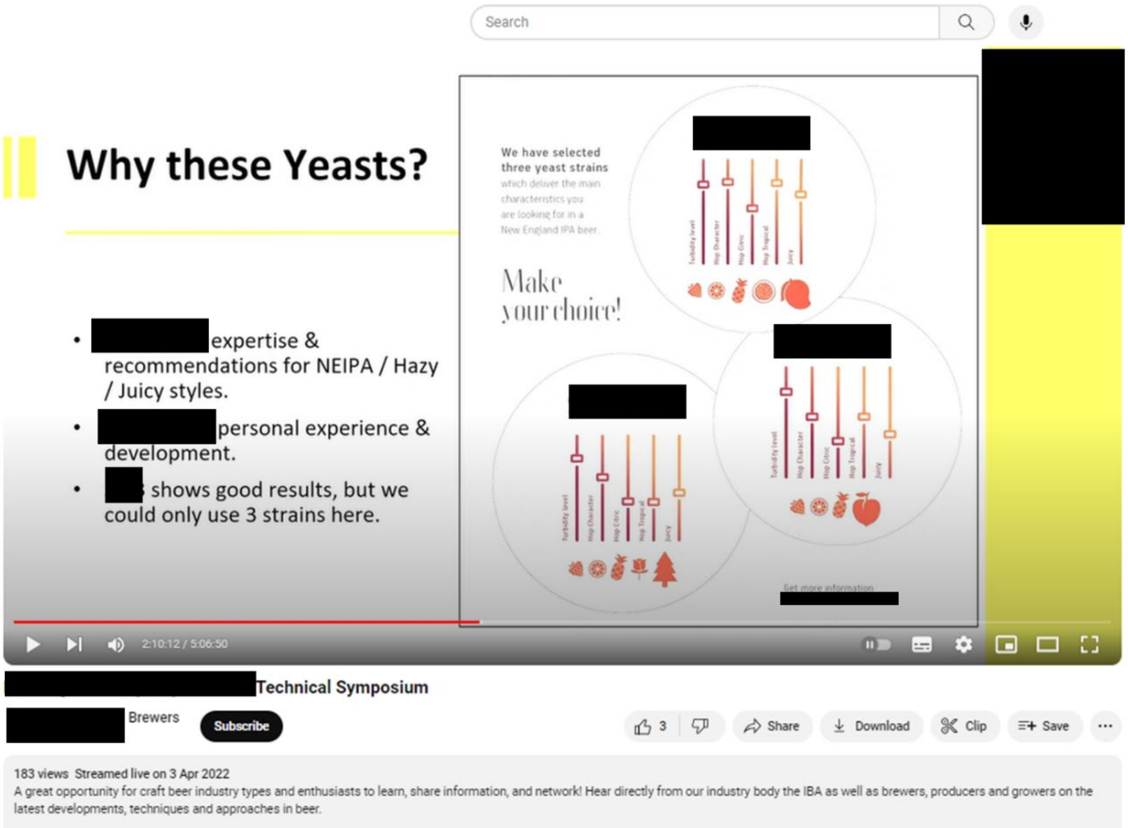
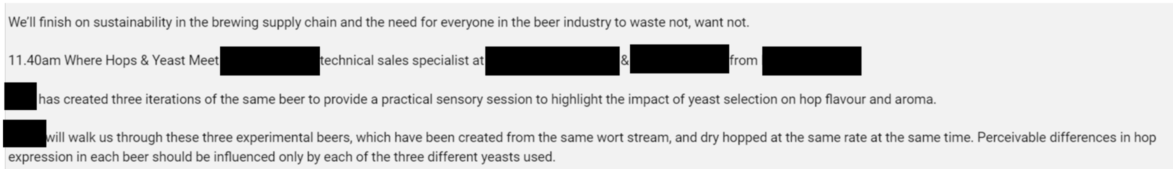


Figure 9. YouTube contents: positive sentiment not rated by AI.


We'll finish on sustainability in the brewing supply chain and the need for everyone in the beer industry to waste not, want not.

11.40am Where Hops & Yeast Meet [redacted] technical sales specialist at [redacted] & [redacted] from [redacted]

[redacted] has created three iterations of the same beer to provide a practical sensory session to highlight the impact of yeast selection on hop flavour and aroma.

[redacted] will walk us through these three experimental beers, which have been created from the same wort stream, and dry hopped at the same rate at the same time. Perceivable differences in hop expression in each beer should be influenced only by each of the three different yeasts used.

Figure 10. YouTube contents: positive sentiment not rated by AI – Caption detail

In the video (Figure 9), the brand is described by the speaker as a “long-standing” company in the field, producing “excellent-quality” and “great” yeast. The speaker describes their personal experience with the client’s products with the support of a PowerPoint presentation. The slide in Figure 9 further refers to the client’s “expertise” and “good results”, and shares details about the client’s website and its products’ features. We interpreted this content as an example of good publicity for the client and labelled the item as positive. This is the only item in which we expected potential challenges in the automatic interpretation and classification of the contents.

In all the other cases, since YouTube is included among the sources scrutinized by Meltwater and the brand’s and the products’ names were always referred to in the captions, we initially thought that the mentions could be easily detected even without multimodally-informed modelling. However, as anticipated, Meltwater ultimately failed to recognize such mentions, and all YouTube contents were not rated.

4.1.2.4. Client’s website

Meltwater also struggled with the classification of the client’s website pages, since they display manifold combinations of visual, verbal, and aural semiotic resources that acted as sentiment cues.

As a part of its marketing strategy, the client created several content pages in which they provided general - seemingly neutral - information about a given topic related to the beer and wine-making world (e.g., “Low-alcohol beer - everything you need to know”), combined with the presentation of their products (e.g., a product to make low-alcohol beer), by using promotional videos and texts in combination with images evoking positive emotions.

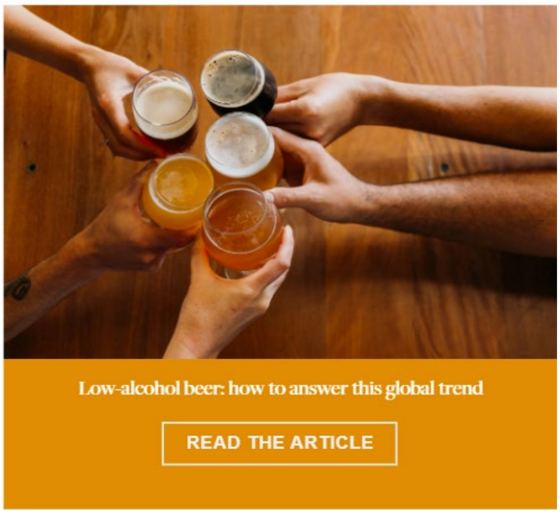


Figure 11. Client’s website: positive sentiment rated neutral by AI.

Figure 11 shows an image used on the client’s website page on low-alcohol beer. Five people are making a toast with different types of beers, saliently positioned at the centre-left of the picture while the people’s arms function as vectors (Kress & van Leeuwen, 2020). The image establishes a positive symbolic association with ideals like friendship, closeness, and happiness and implicitly suggests that people can have fun while drinking low-alcohol beer, while also showing the wide variety of beers that can be created with the client’s products. Below the image, a link to an article about low-alcohol beer is included to provide the illusion of neutrality but also position the low-alcohol beer as a “global trend” the client is in step with. The website also displays a sophisticated use of colours (Kress & van Leeuwen, 2020): in particular, the recurrence of warm colours (orange, yellow, and red for the background and caption boxes) creates a visual association with the colours of beer and wine. While we can easily detect this intricate web of interrelated semiotic resources and the pragmatic allusions at play, this task is extremely complex for a machine to perform. Thus, we consistently classified these contents as positive, whereas Meltwater considered them neutral. This suggests that images and videos were disregarded and the promotional nature of the pages was misunderstood.

5. Discussion

Section 4 confirms that applications of SA to AI models present a number of issues when linguistic and multimodal perspectives are considered for OCMC. Despite acknowledging Meltwater’s benefits of providing a quantitative generalized indication of sentiment distribution (Santonocito & Polli, 2023), a high qualitative mismatch emerges in our comparison between AI-based SA *vs.* our manual tagging. From an OCMC perspective, the SML tool is valuable for identifying potential threats, crisis response opportunities, and best practices. The classification of negative sentiment is a case in point. In our dataset, the four items we manually tagged as negative mentions neither correspond to nor are included within the twelve items related to the negative mentions detected by Meltwater (Table 1). In this respect, our manual qualitative tagging detected all negative sentiment (four cases) in the dataset. Even if it can be argued that the number of negative mentions detected qualitatively is negligible, an OCMC perspective cannot ignore their damage potential (Lei & Liu 2021).

The percentages of Graph 1 show a general comparison between the AI-based and the manual analyses. The highest degree of matching is with neutral mentions (i.e., 94%). Conversely, problematic aspects emerge in all cases of negative mismatch, and, to a lesser extent, in cases of positive mentions (i.e., 68%). The SML detection of negative sentiment has been particularly poor in this study, even though the data at hand was not statistically significant. When negative mentions are not properly classified, reputation cannot be recovered successfully. Additionally, two positive and three neutral misclassifications have been found as actually negative, further complicating intervention areas to restore reputation. The same can be said about positive mentions misclassified as neutral and neutral mentions misclassified as positive, because these do not consider actual good marketing practices.

To address RQ1, when the detection of sentiment depends on the multimodal orchestration of diverse semiotic resources, Meltwater fails to detect the sentiment produced by multimodal ensembles. As for the first issue, the machine struggles to make sense of the combination of the verbal and the visual modes when the latter is represented by different types of static images which interact with the verbal text to produce a specific sentiment. This is exemplified in Figures 1 and 11, where the sentiment is rated as neutral, and in Figure 2, where the combination of verbal text, pictures, star-rating system, and emojis is rated as negative. The slight quantitative mismatch between AI detection *vs.* manual tagging, as shown in Table 1, is actually significant in terms of reputational audits. In addition, when the SML tool misclassifies as negative those mentions that are actually positive (see, Figure 2), the assessment of potential reputational damage is even more serious.

Misclassification problems entail also multimodal contents which feature languages other than English. In this respect, Figure 4 exemplifies how Anglo-centric training data (i.e., the prevalence of English-alone datasets, discussed in Santonocito & Polli, 2023) affects the misrecognition of the verbal text in French. When languages other than English are accompanied by multimodal resources (e.g., emojis, pictures, as in the example discussed in Figure 4) the SML tool performs poorly.

The second issue in the interpretation of multimodal ensembles relates to Meltwater's criticalities in understanding hypertextuality and hypermodality (Lemke, 2002) of digital artifacts. The SML tool fails to successfully interpret visual and typographic features, such as emojis, hearts, and bold capitalization. In Figure 3, Meltwater neither detects the smiling face emoji, nor the interactions (i.e., two 'likes'), nor the reposting (i.e., retweet). Verbal cues conveying a positive sentiment, such as slang expressions and irony, are disregarded as well. Therefore, Meltwater classifies the sentiment as neutral, when this is not the case.

The third issue involves Meltwater's failure to correctly rate combinations of verbal, visual, and aural modes. In particular, sentiment is not correctly detected or not rated by the AI when its production depends on the complex multimodal combination of moving images, audio, spoken and written language. In cases featuring the combination of videos with emojis and verbal text (Figure 5), AI-based SA is incorrect because the machine ignores the pragmatic function of informal register, as instantiated by slang expressions and emojis. Therefore, the friendly tone and ensuing positive sentiment discussed in Figure 5 is misclassified into negative. Figures 6 and 7 further show misclassification due to incorrect interpretation of combinations of aural and verbal resources. In fact, despite the presence of aural and verbal indicators of promotional and positive messages, the SML tool rates these contents as neutral. The case of vidcasts, entirely constituted by YouTube contents, is even more emblematic. Despite the fact that YouTube features as a scrutinizable channel for SA[4], all YouTube vidcasts were not rated by Meltwater. In contrast, in our manual tagging we labelled some contents as neutral (Figure 8), while others as positive (Figures 9 and 10).

[4] https://www.meltwater.com/en/suite/social-listening-analytics (last access: 12.01.2024)

Despite the proprietary reasons that make information on Meltwater's algorithmic infrastructure inaccessible, our analysis indicates that in cases where the sentiment is conveyed by the multimodal combination of different semiotic resources the SML tool does not perform successfully: it tends to either accomplish poor SA, or to not accomplish the task at all. This happens regardless of the digital platform and the complexity of the activated modes. Figures 2, 3, 4, and 5 showcase the interplay of complex pragmatic and multimodal components. However, all other examples show that sentiment is not successfully detected by the AI in less problematic multimodal combinations. Our analysis indicates that despite the presence of linear verbal text containing unambiguous polar words, the combination with another semiotic resource different from the verbal one leads to AI-based sentiment misclassification. This tends to happen irrespective of the level of semiotic complexity. In the case of vidcasts, the machine even stops working (Figures 8, 9, and 10 are not rated by Meltwater).

Therefore, our analysis demonstrates that sentiment in digital discourses is not successfully detected because it mostly depends on the verbal mode, even in cases when the verbal mode alone does not provide enough information for the detection of polar sentiment. Just as digital discourses produce meaning by accounting for the complexity of new genres and interactional patterns (Sindoni, 2013), so we argue the case for a more all-encompassing view of sentiment in digital discourse and practices.

As for RQ2, in order to improve AI-based SA, current challenges in misleading negative, neutral, and positive classification need to be more comprehensively addressed. Having recognized AI challenges in SA from multimodal resources, it is undeniable that one way to address such challenges could be the implementation of multimodal fusion models. However, since at the moment these models are still in the process of being fully developed, we maintain our focus on how CMS and discursive perspectives can improve the current output of AI-based SA. For this reason, we draw on CMS by Jewitt and Kress (2010) to emphasize that online sentiment is the product of multimodal discourses where the different semiotic resources are orchestrated with hypertextual and hypermodal interactions (Lemke, 2002), which need to be interpreted with the aid of human evaluation. Since online sentiment derives mainly from human interactions, no matter how these could be mediated or affected by technology and automation, we argue that it is human awareness that should engage with the SML tools in order to improve the output of AI-based SA. It is ultimately human meta-discursive and meta-pragmatic competence that recognizes discursive materializations and establishes genres, roles, and correlations among cognitive, rhetorical, and linguistic features (Catenaccio, 2023). These features tend to create a situated understanding and specific meanings for every given situation, thus contributing to successful SA.

Our CMS perspective allows to challenge the logocentric view on discourses and sentiment. By decentralizing the role of the verbal mode in the production of meaning we are indebted to the social semiotic tradition (Kress, 2010) in acknowledging that human affective discourses stem from a holistic process whereas every semiotic resource is intentional to the same degree in producing a certain emotion. Being constantly shaped and re-negotiated according to the dominant multimodal orchestrations and to speakers' and participants' intentions, digital discourses – together with emotions – become even more transient and unstable. In this study, this is testified by the inaccessibility of certain contents: some were not accessible since the beginning, while others have become inaccessible at later stages of our analysis. Additionally, the regimes of dynamic manipulation of digital texts (KhosraviNik, 2020), such as content editing and/or deletion, must be considered. We therefore draw on KhosraviNik and Unger (2016) to suggest that the a-historicity of online contents be considered as a distinct affordance of digital discourses for it favours or restricts what users can do.

Our Critical take challenges the monolithic view on discourses by focusing on users' motivated choices in the materialization of sentiment, which, in turn, may surface ideological representations and power relations. Just as emotions can vary over time and according to context, so discourses supporting such emotions are fluid and influenced by techno-discursive dynamics (i.e., the internal structure of the message, intended circulation, strategic design for affective engagement, see KhosraviNik, 2020). Trying to interpret our findings with Critical lenses, we contend that sentiment supported by transient multimodal combinations must be tackled with the awareness that the naturalization of certain multimodal patterns can allow the surfacing of ideological assumptions and power relations that materialize distorted discursive and social practices, such as strategically orchestrated negative reviews, click-baiting strategies, and fake news dissemination (KhosraviNik, 2018).

Given the complex discursive and semiotic dynamics at stake in OCMC, we acknowledge the usefulness of SML tools like Meltwater. However, we wish to draw attention to the current issues in classifying AI-based sentiment derived from multimodal ensembles. For this reason, we emphasize that the holistic process of interpreting affective discourses and sentiment therein conveyed is possible if the product of AI-powered SML tools is reviewed by human judgment, which is capable of identifying emotions within the combination of different semiotic modes and resources and within and across discourses realized by interlaced pragmatic and discursive features.

6. Conclusions

This paper has investigated the criticalities linked to the interpretation of sentiment conveyed by multisemiotic aggregations by using a dataset gathered from an empirical OCMC fieldwork experience. The dataset included online conversations about a brand and its products referring to the year following a crisis event. By using the SML tool Meltwater, items were gathered from a wide range of online sources in three geographical areas (UK, USA, and France). They encompassed different text types (e.g., social media posts, blog entries, journal articles, YouTube contents, podcasts, and the firm's website pages) in 15 languages. Data were categorized according to their sentiment via Meltwater and manually. The comparison and qualitative examination of results indicated that AI-based SA was problematic because algorithms were not designed to recognize 'multimodal sentiment'. Meltwater used only unimodal (i.e., verbal) classifiers to interpret the results, while disregarding non-verbal (visual, aural, video) emotional cues, even in cases in which the machine was supposedly trained to recognize them (e.g., in the case of emojis). Criticalities linked to multimodality combined with, and possibly intensified, further challenges related to the interpretation of non-English texts (e.g., French) and complex pragmatic and discursive features (e.g., the use of slang or subtle persuasive strategies).

In contemporary cyberspace, where digital fluidity requires the ability to embrace recurrent changes in communication means, modes, and affordances, such unimodal approach (seemingly dominated by a 'standardized' and Anglo-centric view of communication, in which verbal language is the most prominent mode) to digital text classification and interpretation appears to be bucking the trend. We believe that a praxis based on the integration of CMS insights into research on AI and SML technologies may pave the way for a renegotiation and possible overcoming of such in-built biases to improve the outputs of AI-based SA. In this context, although we acknowledge that our reflections are based on a small sample and a single SML tool, the fieldwork experience described here represented a first step towards the creation of a joint research pathway. Further empirical inquiries may prove beneficial to extend and refine the classification of criticalities of AI-based SA tools, for example by including larger datasets or SML tools other than Meltwater to create a detailed taxonomy of the challenges that AI still needs to be addressed to mimic human intelligence.

Acknowledgements:

This work was supported by the Italian National Ministry of University within the PON Programme ("Programma Operativo Nazionale Ricerca e Innovazione") 2014-2020, Azione IV, 4, "Dottorati e contratti di ricerca su tematiche dell'innovazione"; Project: "Development of linguistic and semiotic protocols for Artificial Intelligence models informed by multimodality".

The Authors wish to thank the team of Digital Trails (https://www.digitaltrails.co.uk/), and especially Dr. Bogdan Marinescu, for hosting them and generously sharing their expertise.

References

Al-Ghamdi, L. M. (2021). Towards adopting AI techniques for monitoring social media activities. *Sustainable Engineering and Innovation*, 3(1), 15-22.

Angel, S. O. *et al.* (2021). Systematic literature review of sentiment analysis in the Spanish language. *Data Technol Appl*, 55(4), 461-479. https://doi.org/10.1108/DTA-09-2020-0200.

Avery, E. J. (2017). Public information officers' social media monitoring during the Zika virus crisis, a global health threat surrounded by public uncertainty. *Public Relations Review*, 43(3), 468-476.

Bateman, J. A. *et al.* (2017). *Multimodality: Foundations, research and analysis: A problem-oriented introduction*. De Gruyter Mouton.

Boianu, C., & Bîră, M. (2021). Using social media listening in crisis communication and management: New methods and practices for looking into crises. *Sustainability*, 13, 13015. https://doi.org/10.3390/su132313015.

Calvo, R., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18-37.

Cambria, E. *et al.* (2017). Affective Computing and Sentiment Analysis. In Cambria, E. et al. (eds) *A practical guide to sentiment analysis* (pp. 1-10). Springer.

Catenaccio, P. (2021). Discursive strategies of legitimation on the web: Stakeholder dialogue in the agri-biotech industry. *Discourse, Context & Media*, 43, 100535. https://doi.org/10.1016/j.dcm.2021.100535

Catenaccio, P. (2023). Large Language models and discourse analysis. An introduction. *Paper presented at: International Symposium AI: Challenges and potentialities. A Discourse Analysis Perspective*. 19th-20th October 2023. University of Bergamo, Italy.

Coombs, W. T. (2019). Crisis Communication. The best evidence from research. In Gephart, R. et al. (eds) *The Routledge companion to risk, crisis and emergency management* (pp. 51-66). Routledge.

Djonov, E., & Zhao, S. (2014). From Multimodal to Critical Multimodal Studies through Popular Discourse. In Djonov, E. and Zhao, S. (eds) *Critical multimodal studies of popular discourse* (pp. 1-16). Routledge.

Duarte, N. *et al.* (2017). Mixed messages? The limits of automated social media content analysis. Center for Democracy and Technology. https://cdt.org/insights/mixed-messages-the-limits-of-automated-social-media-content-analysis/ (last accessed 14.07.2023).

Evans, V. (2017). *The emoji code: The linguistics behind smiley faces and scaredy cats*. Picador.

Eriksson, M. (2018). Lessons for crisis communication on social media: A systematic review of what research tells the practice. *International Journal of Strategic Communication*, 12, 526-551.

Fairclough N. (1992). *Discourse and social change*. Polity Press.

Fairclough, N., & Wodak, R. (1997). Critical discourse analysis. In T. A. van Dijk (ed), *Discourse as social interaction* (pp. 258-284). Vol. 2. Sage.

Foucault, M. (1972). *The archaeology of knowledge and the discourse on language*. Pantheon.

Halliday, M. A. K. (1978). *Language as social semiotic: the social interpretation of language and meaning*. Edward Arnold.

Hayes J. L. *et al.* (2021). Can social media listening platforms' artificial intelligence be trusted? Examining the accuracy of Crimson Hexagon's (Now Brandwatch Consumer Research's) AI-driven analyses. *Journal of Advertising*, 50(1), 81-91.

Jain, P. K. *et al.* (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer Science Review*, 41, 100413.

Jewitt *et al.* (2016). *Introducing Multimodality*. Routledge.

Jewitt, C. (2016). Multimodal analysis. In Georgakopoulou, A. and Spilioti, T. (eds) *The Routledge handbook of language and digital communication* (pp. 69-84). Routledge.

Jewitt, C., & Kress, G. (2010). Multimodality, literacy and school English. In D. Wyse, R. Andrews and J. Hoffman (eds), *The Routledge international handbook of English, language and literacy teaching* (pp. 342-352). Routledge.

Joshi *et al.* (2017). Sentiment resources: lexicons and datasets. In E., Cambria *et al.* (eds), *A practical guide to sentiment analysis* (pp. 85-106). Springer.

Khatua A. *et al.* (2020). Predicting political sentiments of voters from Twitter in multi-party contexts. *Applied Software*

Computer Journal, 97, 106743. https://doi.org/10.1016/j.asoc.2020.106743.

KhosraviNik, M. (2018). Social media techno-discursive design, affective communication and contemporary politics. *Fudan Journal of the Humanities and Social Sciences*, 11, 427-442. https://doi.org/10.1007/s40647-018-0226-y

KhosraviNik, M. (2020). Digital meaning-making across content and practice in social media critical discourse studies. *Critical Discourse Studies*. https://doi.org/10.1080/17405904.2020.1835683.

KhosraviNik, M., & Unger, J. (2016). Critical discourse studies and social media: power, resistance and critique in changing media ecologies. In R. Wodak and M. Meyer (eds), *Methods of Critical Discourse Studies* (pp. 206-233). Sage.

Kotras, B. (2020). Opinions that matter: The hybridization of opinion and reputation measurement in social media listening software. *Media Culture and Society*, 42 (7-8), 1495-1511.

Kress, G., & van Leeuwen, T. (1996/2020). *Reading images: The grammar of visual design*. Routledge.

Kress, G. (2010). *Multimodality. A social semiotic approach to contemporary communication*. Routledge.

Lai, S. *et al.* (2023). Multimodal Sentiment Analysis: A survey. https://arxiv.org/abs/2305.07611 .

Lanfranchi, V. (2017). Machine learning and social media in crisis management: agility vs ethics. *WiPe Paper – Ethics Legal and Social Issues Proceedings of the 14th ISCRAM Conference*.

Lei, L., & Liu, D. (2021). *Conducting Sentiment Analysis*. Cambridge University Press.

Lemke, J. L. (2002). Travels in hypermodality. *Visual Communication*, 1(3), 299-325.

Levashina, J. *et al.* (2014). The structured employment interview: narrative and quantitative review of the research literature. *Pers Psychol*, 67(1), 241-293.

Liang, B. *et al.* (2022). Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowledge-Based Systems*, 235, 107643. https://doi.org/10.1016/j.knosys.2021.107643

Liu, B. (2012). Sentiment analysis and opinion mining. Vol. 16. *Synthesis Lectures on Human Language Technologies*. Morgan & Claypool. https://doi.org/10.2200/S00416ED1V01Y201204HLT016.

Machin, D. (2013). What is multimodal critical discourse studies? *Critical Discourse Studies*, 10(4), 347-355. https://doi.org/10.1080/17405904.2013.813770.

Majumder, N. *et al.* (2018). Multimodal sentiment analysis using hierarchical fusion with context modeling. *Knowledge-Based Systems*, 161, 124-133.

Martinec, R., &Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication,* 4, 337. https://doi.org/10.1177/1470357205055928.

McDonald, C.G., & Slawson, V. C. (2002). Reputation in an Internet auction market. *Economic Inquiry*, 40(4), 633-650.

Mohammad, S. M. (2017). Challenges in Sentiment Analysis. In Cambria, E. *et al.* (eds). *A Practical guide to sentiment analysis* (pp. 61-84). Springer.

Moschini, I., & Sindoni, M. G. (2022). *Mediation and multimodal meaning making in digital environments*. Routledge. https://doi.org/10.4324/9781003225423.

O'Halloran, K. L. *et al.* (2014). Multimodal pragmatics. In Schneider, K. and Barron, A. (eds) *Pragmatics of discourse* (pp. 239-268). De Gruyter Mouton.

Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.

Perakakis, E. *et al.* (2019). Social media monitoring: an innovative intelligent approach. *Designs,* 3(2), 1-12.

Picard, R. (2000). *Affective computing*. Boston: The MIT Press.

Polli C., Sindoni, M.G. (2024). Multimodal computation or interpretation? Automatic vs. critical understanding of text-image relations in racist memes. *Discourse, Context & Media*, 57, 100755. https://doi.org/10.1016/j.dcm.2024.100755

Santonocito C.S., Polli C., (2023). An experimental critical multimodal discourse study to the AI-driven sentiment analysis of online crisis communication. *Lingue e Linguaggi*, 59, 333-349. https://doi.org/10.1285/i22390359v59p333

Schreier, M. (2012). *Qualitative content analysis in practice*. Sage.

Schwaiger, J. *et al.* (2021). UR: SMART–A tool for analyzing social media content. *Information Systems and e-Business Management*, 19, 1275-1320.

Sindoni M. G. (2013). *Spoken and written discourse in online interactions: A multimodal approach*. Routledge.

Sindoni M. G. (2023). *Multimodality and translanguaging in video interactions*. Cambridge University Press.

Sunderland, J. (2004). *Gendered discourses*. Palgrave Macmillan.

Upadhyaya, A. *et al.* (2023). Towards sentiment and Temporal Aided Stance Detection of climate change tweets. *Information Processing & Management*, 60(4). https://docs.tweepy.org/en/stable/

Vijay, D. et al. (2018). Corpus creation and emotion prediction for Hindi-English code-mixed social media text. *In: Proceedings of the 2018 conference of the North American chapter of the Association for Computational Linguistics: student research workshop*, pp 128-135.

Wankhade, M. *et al.* (2022). A survey on sentiment analysis methods, applications, and challenges. *Artif Intell Rev*, 55, 5731-5780. https://doi.org/10.1007/s10462-022-10144-1

Ziora, L. (2016). The sentiment analysis as a tool of business analytics in contemporary organizations. *Studia Ekonomiczne*, 281, 234-241.