Cultural Machines: Unlocking the power of digital methods and computational techniques for understanding socio-cultural processes in digital environments

Alessandro Caliandro
University of Pavia

Davide Bennato
University of Catania

Since the advent of big data, social scientists tried to 'unlock' the cultural power embedded in them, by extracting qualitative 'thick' data from huge amount of quantitative digital data (Ford, 2014). This ambitious methodological endeavour has been undertaken by several scholars from different fields in social science, giving birth to numerous innovative research approaches and techniques. See for example the work of STS scholars who adapted the language and methodological array of Actor-Network-Theory to the analysis of big data (Vertesi and Ribes, 2019) – including works on digital mapping of scientific controversies (Venturini, 2010; 2012; Marres, 2015), digital network analysis (Cambrosio et al., 2014; Venturini et al., 2021), or the application of co-word analysis on web content (Venturini and Guido, 2012; Eykens et al., 2021). Other notable contributions have been forthcoming from digital methods (Rogers, 2009), computational approaches (Giglietto, Rossi, Bennato, 2012), interface methods (Marres and Gertliz, 2015), and platform methods (Nieborg et al., 2020) to the exploration and understanding of the huge repositories of qualitative data on social media (Lewis et al., 2013; Niederer, 2016; Rieder et al., 2018). Similarly, various ethnographic approaches have tried to mix ethnographic observation with the use of digital tools for data collection and analysis, such as ethnomining (Aipperspach et al., 2006), trace ethnography (Geiger and Ribes, 2011), ethnography for the Internet (Hine, 2015), computational ethnography (Elish and boyd, 2017), digital methods for ethnography (Caliandro, 2018) – just to name a few.

   Notwithstanding the exceptional advancements in this direction, in our opinion, qualitative analysis of big data and the exploration of cultural processes within them are still

underdeveloped (Pedersen, 2021). On the one hand, so far one of the main (and probably the best) strategies to extract 'thick description' from big data consists in conducting manual analysis (via traditional qualitative techniques, such as qualitative content analysis or ethnographic observation) on small sample of digital data (Caliandro and Gandini, 2017), as adopted for example in social media research on public opinion (Dragotto et al., 2020), fandoms (Arvidsson et al., 2016), brand cultures (Schöps et al., 2020), micro-celebrity (Marwick and boyd, 2011), or platform vernaculars (Gibbs et al., 2015). For how insightful and innovative these studies could be, they are nonetheless difficult to reproduce on a large scale. On the other hand, there exist computational approaches that focus precisely on 'cultural data', like *cultural analytics* (Manovich, 2009), which use automated image recognition techniques to explore huge quantities of visual data (Manovich, 2017). Such approach considers images as data, meaning that it focuses more on structural characteristic of images (e.g., colours, filters, resolution, etc.), rather than the content per se (Niederer and Colombo, 2019).

As a matter of fact, it appears to be still quite arduous to take advantage of digital methods and computational techniques to automate the qualitative analysis and cultural interpretation of digital content on a very large scale - despite the attempts being made in this direction (among others Bennato, 2021; Caliandro and Anselmi, 2021; Cambrosio et al., 2020; Vicari and Kirby, 2022).

Anyhow, it seems that new methodological possibilities, helpful to fill this gap, are now emerging, especially referring to two current 'technical' trends.

First, several major digital platforms are starting to release ad hoc tools (seldom based on artificial intelligence) specifically meant to perform qualitative analysis on digital cultural objects as well as support researchers in their interpretation. These include Google Vision API, a machine learning-based image recognition toolkit (Mulfari et al., 2016) that provides modules to automatically analyse images based on: a) the content of the image itself (image-label); b) its specific 'audiencing' through references obtained from the web (image-web entities); and c) the sites of image circulation (image-domain) (Omena et al., 2021, p. 4). Or CrowdTangle – featured by Facebook – that in its dashboard includes a 'meme search' function (Fraser, 2021). Not to mention GPT-3 (now a product of Microsoft): an artificial intelligence application that is able to autonomously write a text provided a specific question from the user (Schmelzer, 2021) (e.g., 'can I get a poem please?' Or even better 'can I get a research report on the topic X?').

Second, the current processes of platformization of the web (Helmond, 2015), and the subsequent platformization of culture (Duffy et al., 2019), put into motion a progressive standardization of web content as well as their production; a condition that (hypothetically) makes the tracing, analysis, and interpretation of cultural content easier and possibly scalable on larger datasets. A sheer example of such standardization of cultural processes and social dynamics within social media platforms is represented by Internet memes (Shifman, 2014). Memes are "collections of standardized multimodal texts spreading rapidly across digital networks, which consist in user-created derivatives that stem from an original piece of content" (Caliandro and Anselmi, 2021, p. 5; see also Milner, 2016). Memes are becoming very common and widespread means of interaction and communication among

the massive and dispersed publics populating social media. Thanks to memes the dispersed users of social media can 'simulate' conversations on disparate topics (e.g., politics, gossip, social trends, etc.) (Rogers, 2019). Furthermore, memes amount to be not only cultural objects or means of communication, but also provide the structural and semantic template that inform other forms of cultural production on social media. Consider for example the memetic logics informing the visual production of brands' fans on Instagram (Caliandro and Anselmi, 2021), or teenagers' self-presentation strategies on TikTok (Zulli and Zulli, 2020). Framed in this way, Internet memes are not mere cultural object to study, but also methodological resources that work as a heuristic to identify, follow, analyse, and interpret processes of platformization of culture in social media (Nieborg et al., 2020). In this sense, memes look very much like Tardian's monads (Tarde/Clark, 2010), which, in the epistemological elaboration of Bruno Latour, can be considered as single (digital) entities that contain "a representation, a reflection, or an interpretation of a whole set of other elements borrowed from the world around [them]" (Latour 2010, p. 57; Latour et al., 2012).

Therefore, more generally speaking, we can argue that a sort of new infrastructure of 'cultural machines' is emerging from digital platforms – made by an assemblage of platforms, pieces of software, content, and participatory cultures –, which can be used as a methodological resource to extract 'thick descriptions' from big data as well as interpret and analyse complex cultural processes on a very large scale. Given this emerging scenario, brand new methodological challenges pose to social researchers dealing with digital sociology and digital methods. To this purpose, the present special issue includes four stimulating articles that try to address the above-mentioned methodological challenges. The contributions are the following: *Fandom and pop politics* (Claudia Cantale); *Memes as socio-narrative representations* (Elisabetta Risi, Riccardo Pronzato, Guido Di Fraia); *Platformization hate* (Miriam Di Lisio, Rosa Sorrentino, Domenico Trezza); *Russian-Ukraine war and institutional use of memetic communication: methodological opportunities and challenges* (Gabriella Punziano, Domenico Trezza, Suania Acampa). Moreover, the special issue contains an invited article by Richard Rogers (the father of *Digital Methods*), who reflects on the ontology of the meme 'as technical object'. The article is the transcription of the opening speech that Professor Rogers gave at the 2021 ILIS Conference (Salerno, Italy) - which was dedicated to the topic of '*Research Methods in the Digital Society*'. All these five contributions offer to the reader innovative research strategies, techniques, and instruments to follow, analyse, and interpret digital cultural objects and processes unfolding within digital environments.

Nota biografica

Alessandro Caliandro is an Associate Professor in Sociology of Culture and Communication at the Department of Political and Social Sciences at the University of Pavia (where he teaches Digital Communication, Digital Journalism and Digital Consumer Culture). His current research focuses on digital methods, consumer culture, platformization of culture, surveillance capitalism, and ageing & digital technologies. Among his most recent

publications: Audy Martínek, P., Caliandro, A., & Denegri-Knott, J. (2022) 'Digital practices tracing: studying consumer lurking in digital environments' *Journal of Marketing Management*; Caliandro, A., & Anselmi, G. (2021) 'Affordances-based brand relations: An inquire on memetic brands on Instagram' *Social Media+ Society*. His forthcoming book is titled 'The Platformization of Consumer Culture: A Digital Methods Guide' (Amsterdam University Press – with A. Gandini, L. Bainotti, G., Anselmi). He is research coordinator of V-Data Project: 'The value of digital data: enhancing citizens' awareness and voice about surveillance capitalism' (https://vdataresearch.com/).

Davide Bennato is Associate Professor of Digital Media Sociology at the Department of Humanities of the University of Catania. He is president of the Degree Course in Sciences and languages for the communication of the University of Catania, member of the board of professors of the Ph.D. in Complex Systems for Physical, Socio-economic and Life Sciences at the Department of Physics and Astronomy of the University of Catania and board member of the Center for Digital Humanities (CINUM) of the University of Catania. His research topics are digital sociology, collective behaviors in social media, big data ethics, science and technology communications models in the internet, computational social science. Among his most recent publications: 'The Digital Traces' Diamond. A Proposal to Put Together a Quantitative Approach, Interpretive Methods, and Computational Tools', *Italian Sociological Review*, 11(4S), pp. 207-224 (2021); *Being digital media. From the separation to the confusion between mass media and personal media*, in D. Bruni (ed.), Media, Power and Public Opinion. Essays on Communication and Politics in a Historical Perspective, Peter Lang, Bruxelles (2022).

References

Aipperspach, R., Rattenbury, T. L., Woodruff, A., Anderson, K., Canny, J. F., & Aoki, P. (2006). Ethno-mining: integrating numbers and words from the ground up. Electrical Engineering and Computer Sciences University of California at Berkeley Tech Report UCB/EECS-2006, 125. Retrieved from https://www2.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-124.html.

Arvidsson, A., Caliandro, A., Airoldi, M., & Barina, S. (2016). Crowds and value. Italian directioners on Twitter. *Information, Communication & Society*, 19(7), 921-939.

Bennato, D. (2021). The Digital Traces' Diamond. A Proposal to Put Together a Quantitative Approach, Interpretive Methods, and Computational Tools. *Italian Sociological Review*, 11 (4S), 207-224. http://dx.doi.org/10.13136/isr.v11i4S.432

Caliandro A. & Gandini A. (2017), *Qualitative Research in Digital Environments: A Research Toolkit*. Routledge, London.

Caliandro, A. (2018). Digital methods for ethnography: Analytical concepts for ethnographers exploring social media environments. *Journal of Contemporary Ethnography*, 47(5), 551-578.

Caliandro, A., & Anselmi, G. (2021). Affordances-Based Brand Relations: An Inquire on Memetic Brands on Instagram. *Social Media+ Society*, 7(2), https://doi.org/10.1177/20563051211021367.

Cambrosio, A., Cointet, J. P., & Abdo, A. H. (2020). Beyond networks: Aligning qualitative and computational science studies. *Quantitative Science Studies*, 1(3), 1017-1024.

Cambrosio, A., Bourret, P., Rabeharisoa, V., & Callon, M. (2014). Big data and the collective turn in biomedicine. How should we analyze post-genomic practices?. *TECNOSCIENZA: Italian Journal of Science & Technology Studies*, 5(1), 11-42.

Dragotto, F., Giomi, E., & Melchiorre, S. M. (2020). Putting women back in their place. Reflections on slut-shaming, the case Asia Argento and Twitter in Italy. *International Review of Sociology*, 30(1), 46-70.

Duffy, B. E., Poell, T., & Nieborg, D. B. (2019). Platform practices in the cultural industries: Creativity, labor, and citizenship. *Social Media+ Society*, 5(4), https://doi.org/10.1177/2056305119879672.

Eykens, J., Guns, R., & Engels, T. C. (2021). Fine-grained classification of social science journal articles using textual data: A comparison of supervised machine learning approaches. *Quantitative Science Studies*, 2(1), 89-110.

Ford, H. (2014). Big Data and Small: Collaborations between ethnographers and data scientists. *Big Data & Society*, 1(2), 1-3. https://doi.org/10.1177/2053951714544337.

Fraser, L. (2021). What data is CrowdTangle tracking? https://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking

Geiger, R. S., & Ribes, D. (2011). Trace ethnography: Following coordination through documentary practices. In 2011 44th Hawaii international conference on system sciences (pp. 1-10). IEEE.

Gibbs, M., Meese, J., Arnold, M., Nansen, B., & Carter, M. (2015). # Funeral and Instagram: Death, social media, and platform vernacular. *Information, Communication & Society*, 18(3), 255-268.

Giglietto, F., Rossi, L., & Bennato, D. (2012). The open laboratory: Limits and possibilities of using Facebook, Twitter, and YouTube as a research data source. *Journal of technology in human services*, 30(3-4), 145-159.

Helmond, A. (2015). The platformization of the web: Making web data platform ready. Social *Media+ Society*, 1(2), https://journals.sagepub.com/doi/pdf/10.1177/2056305115603080.

Latour, B. (2010). *Tarde's idea of quantification* (pp. 161-178). Routledge.

Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (2012). 'The whole is always smaller than its parts'–a digital test of Gabriel Tardes' monads. *The British journal of sociology*, 63(4), 590-615.

Lewis, S. C., Zamith, R. and Hermida, A. (2013). Content analysis in an era of Big Data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), pp. 34–52.

Manovich, L., (2009), How to Follow Global Digital Cultures, or Cultural Analytics for Beginners, http://manovich.net/index.php/projects/how-to-follow-global-digital-cultures

| Alessandro Caliandro, Davide Bennato

Manovich, L. (2017). Instagram and contemporary image. Manovich.net. http://manovich.net/index.php/projects/instagram-and-contemporary-image.

Marres, N. (2015). Why map issues? On controversy analysis as a digital method. *Science, Technology, & Human Values*, 40(5), 655-686.

Marres, N. and Gerlitz, C. (2015). Interface methods: renegotiating relations between digital social research, STS and sociology. *The Sociological Review*, 64(1), pp. 21–46.

Marwick, A., & Boyd, D. (2011). To see and be seen: Celebrity practice on Twitter. *Convergence*, 17(2), 139-158.

Milner, R. M. (2016). *The world made meme: Public conversations and participatory media*. MIT Press

Nieborg, D. B., Duffy, B. E., & Poell, T. (2020). Studying platforms and cultural production: Methods, institutions, and practices. *Social Media+ Society*, 6(3), https://doi.org/10.1177/2056305120943273

Niederer, S. (2016). *Networked content analysis: The case of climate change*. (Doctoral dissertation, Universiteit van Amsterdam).

Niederer, S., & Colombo, G. (2019). Visual methodologies for networked images: Designing visualizations for collaborative research, cross-platform analysis, and public participation. *Diseña*, 14, 40–67.

Omena, J. J., Pilipets, E., Gobbo, B., & Chao, J. (2021). The Potentials of Google Vision API-based Networks to Study Natively Digital Images. *Diseña*, (19), Article.1. https://doi.org/10.7764/disena.19.Article.1

Pedersen, A. (2021). Machine Anthropology, *Big Data & Society*, https://journals.sagepub.com/page/bds/collections/machineanthropology?M_BT=54541027589991.

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to 'ranking cultures' Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50-68.

Rogers, R. (2009). *The End of the Virtual*. Amsterdam: Amsterdam University Press.

Rogers, R. (2019). *Doing digital methods*. SAGE

Schmelzer, R. (2021). GPT-3, https://searchenterpriseai.techtarget.com/definition/GPT-3.

Schöps, J. D., Kogler, S., & Hemetsberger, A. (2020). (De-) stabilizing the digitized fashion market on Instagram–dynamics of visual performative assemblages. *Consumption, Markets & Culture*, 23(2), 195-213.

Shifman, L. (2014). The cultural logic of photo-based meme genres. *Journal of Visual Culture*, 13(3), 340–358.

Tarde, G. and Clark, T.N. (2010). *Gabriel Tarde on communication and social influence: Selected papers*. University of Chicago Press.

Venturini, T. (2010). Diving in magma: how to explore controversies with actor-network theory. *Public understanding of science*, 19(3), 258-273.

Venturini, T. (2012). Building on faults: how to represent controversies with digital methods. *Public understanding of science*, 21(7), 796-812.

Venturini, T., & Guido, D. (2012). Once upon a text: an ANT tale in text analysis. *Sociologica*, 6(3), 0-0.

Venturini, T., Jacomy, M., & Jensen, P. (2021). What do we see when we look at networks: Visual network analysis, relational ambiguity, and force-directed layouts. *Big Data & Society*, 8(1), 20539517211018488.

Vertesi, J., & Ribes, D. (2019). *digitalSTS*. Princeton University Press.

Vicari, S., & Kirby, D. (2022). Digital platforms as socio-cultural artifacts: developing digital methods for cultural research. Information, Communication & Society, 1-23, https://doi.org/10.1080/1369118X.2022.2027498.

Zulli, D., & Zulli, D. J. (2020). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, https://doi.org/10.1177/1461444820983603.

Alessandro Caliandro, Davide Bennato