

Bridging the tool gap in visual social media research: SMIQQDA, a free and open-source solution

Tiago Joseph¹, Michaël Stappers², Catherine Bouko³

¹Ghent University – tiago.joseph@ugent.be

²Independent engineer - stappers.michael@hotmail.com

³Ghent University – catherine.bouko@ugent.be

Abstract

The manual analysis of text-image social media posts has gained growing scholarly importance, yet suitable analytical tools remain scarce. Our review of 15 years of research (N = 477) reveals the limited development of tools and technological innovation. Despite its limitations for multimodal and visual content analysis, Excel remains the most widely used software. Against this backdrop of technological limitations, this article introduces SMIQQDA (*Social Media Interface for Quantitative and Qualitative Data Analysis*), a free and open-source tool designed for multimodal data analysis. SMIQQDA allows users to structure, enrich, and synthesize multimodal datasets within a unified database; visualize, sort, and annotate data; create text-image sub-corpora; integrate additional quantitative tools; and synthesize analytical results.

Keywords: discourse analysis; content analysis tool; visual analysis tool; qualitative method; multimodal content

Résumé

L'analyse manuelle des publications (texte-image) issues des réseaux sociaux a acquis une importance scientifique croissante, mais les outils analytiques adaptés restent rares. La revue de 15 années de recherche (N = 477) révèle un développement limité d'outils et d'innovations technologiques. Malgré ses limites pour l'analyse de contenu multimodal et visuel, Excel reste le logiciel le plus utilisé. Face à ces limitations technologiques, cet article présente SMIQQDA (*Social Media Interface for Quantitative and Qualitative Data Analysis*), un outil gratuit et open-source conçu pour l'analyse de données multimodales. SMIQQDA permet aux utilisateurs de structurer, enrichir et synthétiser des ensembles de données multimodales au sein d'une base de données synthétique ; de visualiser, trier et annoter les données ; de créer des sous-corpus texte-image ; d'intégrer des outils quantitatifs supplémentaires ; et de synthétiser les résultats analytiques.

Mots clés : analyse du discours ; outil d'analyse de contenu ; outil d'analyse visuelle ; méthode qualitative ; contenu multimodal

1. Methodological challenges and the need for integrated visual tools in social media research

The most popular social media platforms – YouTube, Facebook, Instagram, and TikTok (Kemp, 2025) – are dominated by (audio)visual content, if not exclusively focused on it. The explosion of visual content exemplifies the *visual turn*, in which images are both subject and medium of communication (Jay, 2002). This visual turn on social networks is also reflected in scientific research. Though, in the era of social media, methodological progress remains uneven. On one hand, enriched by new semiotic frameworks (Kress and van Leeuwen, 2020; Männistö et al., 2022), computational methods are advancing, for visualization techniques (Niederer and Colombo, 2024; Rogers, 2021), automated analysis (Joo and Steinert-Threlkeld,

2022) and computer vision (Pandiani and Presutti, 2023). On the other hand, there is a lack of suitable tools for manual digital research (see Section 3), which often relies on “methodological tinkering” (Julliard and Saemmer, 2022, translated). These practices typically take the form of screenshots, dispersed data storage, or manual annotations in Excel or research notebooks – methods far from ergonomic (Caldeira, 2024; Despontin Lefèvre, 2022; Millette, 2023).

This lack of (open access) suitable tools for manual analysis is not without consequences. Indeed, the technological limitations help explain the scarcity of studies in certain areas, such as visuals in misinformation or the visual framing of crises (Brennen et al., 2021; Omondi, 2025). This lack can be attributed to two main factors. First, the technical skills required may pose a barrier for researchers in the social sciences and humanities (Münster and Terras, 2020). Second, as in automation, standardized tools struggle to capture the semantic elusiveness of visual content, which needs approaches grounded in social context rather than solely structuralist theory. Structuralist approaches, still used in computer vision, focus on identifying pre-given “concepts encoded in an image” (Joo and Steinert-Threlkeld, 2022).

This discussion highlights the challenge of capturing visual meaning, even with methods moving beyond structuralism. Addressing both commonality and difference requires attention to *co-text* (captions, hashtags) and the broader social *context* of visual production and sharing on social media. Consequently, beyond combinations between manual and computational methods, it is crucial to develop tools that support – but do not replace – manual analysis. Such tools should facilitate access to an image’s co-text, context, production, reception, and circulation, enabling qualitative and quantitative manual approaches to conduct close readings at every stage (Baker, 2018; Bouté et al., 2024; Pincemin, 2020; Rogers, 2021). Access to the digital context mitigates the limits of purely quantitative extraction and fosters an ecological, holistic approach to the web (Develotte and Paveau, 2017; Souchier and al., 2019), aligning with Rose’s (2022) call for empirically grounded analyses that consider the social modalities of image production, viewing, and circulation. Recognizing how each social media platform and its affordances influence (visual) communication (Pearce et al., 2018), cross-platform analysis also address a methodological gap by highlighting the need for context-adapted tools.

To ensure data access, verification, and reuse, and to encourage visual and cross-platform analysis, we believe that promoting intuitive, user-friendly tools within an open science framework is essential (Gomez-Diaz and Recio, 2022). Yet despite being integral to research, software remains undervalued in the social sciences and humanities (Di Cosmo, 2022; Gomez-Diaz and Recio, 2022).

In summary, even though research on multimodal data from social media has grown substantially, it continues to face major methodological limitations. Qualitative analysis of digital images often seems to rely on fragmented and ergonomically limited practices. To test this hypothesis, we conducted a systematic review of 477 articles published between 2010 and 2025, which reveals a predominance of a persistent dependence on software ill-suited for visual analysis. We present the results in the following section.

In response to these challenges, Tiago Joseph and Michaël Stappers designed *SMIQQDA*, *Social Media Interface for Quantitative and Qualitative Data Analysis* (Joseph et al., 2025). This software provides a free and open-source solution for structuring, enriching, and analysing textual and visual corpora derived from social media platforms. Developed to meet the needs of researchers in the humanities and social sciences, SMIQQDA makes it possible to centralise, annotate and categorise data in a unified database, and facilitates mixed-methods research and interoperability with other tools. These features and their relevance to critical visual analysis

are discussed in the third section, while the fourth section covers types of data that can be analyzed in SMIQQDA. The fifth section presents the software through a case study, and the conclusion outlines future developments.

2. Review of 15 years of manual analysis for image-based social media research

To objectify the research trends and needs in the field of visual analysis through evidence-based insights, we conducted a systematic review following PRISMA standards (Page et al., 2021). We constructed a Boolean search query incorporating terms related to social media, to the visual dimension of social media posts, and to manual methods for analyzing image-based social media posts. We restricted the search to full-length articles which were accessible either through open access or through institutional access provided by Ghent University. We limited our queries to specific disciplinary fields (Social Sciences, Arts & Humanities, and Business in Scopus) and indexes (Social Sciences Citation Index, the Arts & Humanities Citation Index, and the Emerging Sources Citation Index in Web of Science). The search was conducted on March 1, 2025. This process yielded 1,035 articles in Scopus and 2,664 articles in WoS. To remove duplicates, we utilized JabRef, Rayyan, and Zotero, resulting in a dataset of 3,007 articles. We manually selected only those that met the following four inclusion criteria: the study must be empirical (not conceptual), analyze social media content (not networks), employ manual analysis (not automated) and the dataset must be image-based content (not videos). Rayyan was used to highlight keywords related to the four inclusion criteria in the titles, keywords, and abstracts. We then manually selected the articles according to these inclusion criteria. We did not use Rayyan's AI-powered selection tool, as it produced poor results. After applying these inclusion criteria, the dataset was reduced to 477 articles.

We manually screened the insights generated by Scispace from the PDFs (PubGenius Inc., 2026). To obtain reliable results from SciSpace, we prompted the request for each paper individually. While SciSpace allows up to ten papers to be uploaded at once, we found that result quality declined as more papers were submitted. Uploading one paper at a time ensured accurate outputs, which we validated through a comparative manual screening of 10% of the dataset.

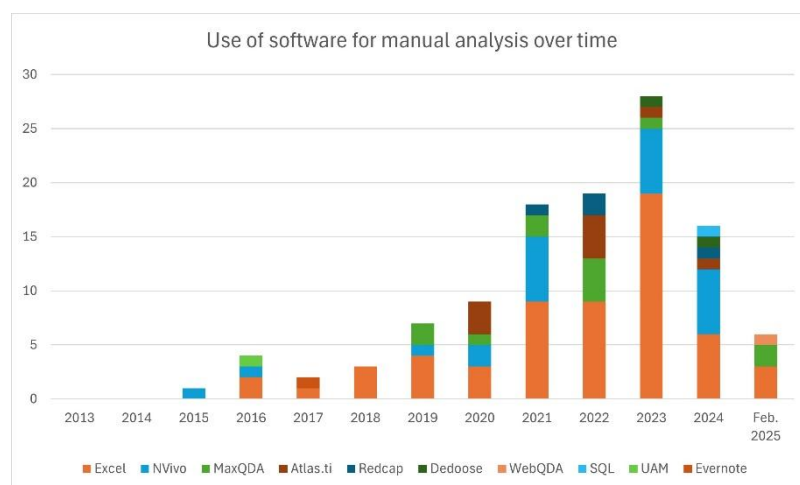


Figure 1. Use of software for manual analysis over time ($N = 477$)

In total, 24% of the 477 articles mention the use of an analysis software (see Figure 1). Excel stands out as the most frequently used tool, consistently appearing across the years and reaching

its highest usage in 2023. NVivo ranks as the second most commonly used software, particularly prominent from 2021 to 2023. MaxQDA has also gained popularity in recent years, especially in 2022 and 2023. Redcap and Dedoose were used more sporadically. The AI tool Atlas.ti saw notable use in 2020 and 2022. The increase in articles mentioning analysis software parallels the rise in the overall number of articles published, although there was a decline in 2024. Moreover, the fact that Excel remains the most used software over time is somewhat concerning, given its limited ergonomics for analyzing visual or multimodal content. To address this persistent dependence on software ill-suited for visual analysis, we introduce SMIQQDA, a free and open-source alternative designed for more effective multimodal analysis.

3. SMIQQDA, a researcher-friendly tool

3.1. Development context, access, and references

SMIQQDA was programmed by Michaël Stappers as part of Tiago Joseph's doctoral research in discourse analysis on a text–image corpus from feminist and queer Instagram accounts. Its first version and source code are freely available on [GitHub](#) and [Zenodo](#) under the free and open-source GPLv3 license, alongside a bilingual (English/French) user manual. Built on Python (3.11.6) and libraries such as Flask, pandas, and XlsxWriter, SMIQQDA is a locally installed application and runs on all major operating systems (Windows, macOS, Linux).

3.2. Building a database: collecting, importing, and linking different corpora

Data collection (preliminary step). Once social media data is scraped, researchers usually obtain two types of data: textual and visual. They are then faced with a large amount of data, which is difficult to access because it is scattered across different formats (textual or visual) and lacks structure (other than by image name, for example). SMIQQDA provides a concise way to organize this diverse data.

SMIQQDA processes textual and visual data in specific formats: jpg for images, xlsx (Excel) and xml for text. These formats are widely supported, sustainable, and easily convertible. xlsx was preferred over csv for its flexibility with encodings and mixed data (text, numbers, dates, URLs), as well as its advanced features (multi-sheets, formatting, calculations, charts). Based on Open XML, it remains open, free, and well documented. To enable joint processing of text and images, SMIQQDA relies on systematic image naming, where each file receives a unique metadata identifier. Profile pictures are named after the account, while post images combine account name, timestamp, and image number (if the post contains multiple images). This naming process is automated by a Python script provided in the documentation.

Importing textual and visual data. Once collected, visual data is copied into a local SMIQQDA folder, while textual data is imported into an SMIQQDA-generated .sqlite database. This database centralizes textual corpora for operations such as synthesis, annotation, or filtering. Import is “non-destructive” (Bouté et al., 2024, p.3, translated): the original Excel files remain intact and unmodified. This feature is essential for text retrieval, scientific evaluation, and data archiving. Furthermore, SMIQQDA accepts any type of textual (meta)data (text, dates, numbers, URLs, etc.), ensuring flexibility across platforms. This adaptability makes it suitable for analyzing heterogeneous social networks and enables cross-platform studies.

Associating Different Corpora. SMIQQDA allows for the association of corpora at two levels. In this way, SMIQQDA contributes to the structuring of rich, complex, and heterogeneous digital data. First, SMIQQDA automatically associates the corresponding textual and visual

data. Second, when different corpora concern a specific set of actors, SMIQQDA allows them to be associated based on common metadata. Indeed, profile, post, and OCR corpora are automatically associated:

- **Profiles and Posts:** Only posts associated with profiles already imported into SMIQQDA can be imported. This setting helps exclude posts that were mistakenly collected (noise) or incorrectly attributed (collection errors), which represents a significant gain in data cleaning. The association between profiles and posts also serves two other functions: display and filtering.
- **Posts and OCR:** OCR results are automatically associated with the posts from which they originate. This task is difficult to carry out manually when the number of social media posts differs between the post corpus and the OCR corpus: manual copying and pasting becomes impossible.

However, SMIQQDA does not prevent from working exclusively with a post corpus if the researcher does not have a profile corpus (e.g., hashtag corpus) or has not conducted OCR, with only a slight technical adjustment required.

3.3. *Cleaning the data: synthesizing and verifying*

Synthesizing the data. The import of textual data has been designed to be synthetic, meaning it is relevant (only posts from imported profiles are included), comprehensive (all textual data is imported), and non-redundant (duplicates are automatically removed). As highlighted in the characteristics of digital data collection (Barats, 2024), cleaning these duplicates is essential, provides a significant time-saving benefit, and strengthens the results of textometric analyses (Pincemin, 2020), which can be combined with the use of SMIQQDA.

Verifying data quality. Several features allow for checking the quality of the data collection. During navigation, the absence of images or other textual data is immediately visible – which is not the case when visual and textual data are separated. Furthermore, access to the source text allows for verifying the alignment between visual and textual data and for completing missing data (e.g., when images from a post were not collected due to a technical issue).

When exporting the SMIQQDA database, the ***automated calculation of uncollected posts and images*** (expected minus collected) allows for evaluating the quality of the data collection.

3.4. *Exploring the data: display, search, sort, filter, return to source text*

Displaying data by simulating the digital environment: profiles and accounts. SMIQQDA replicates Instagram’s visual layout, arranging images, text, and metadata. This recontextualization addresses the semiotic loss from extracting data (Emerit, 2016) and allows posts to be analyzed under conditions similar to typical user reception (Barats, 2024). The ‘profile’ and ‘posts’ tabs enable joint analysis of posts and their associated profiles, enhancing contextual understanding. SMIQQDA allows for the visualization of profiles in a dedicated tab. The profile picture is displayed alongside several metadata (see Figure 2, section 5). SMIQQDA also allows users to view posts in a dedicated tab. When a post contains multiple images, they can be browsed one by one — a feature of SMIQQDA that significantly facilitates analysis. Each post is presented with its images alongside profile and post metadata (see Figure 3, section 5).

Exploring the data: navigate, search, sort, and filter. To facilitate data exploration, SMIQQDA integrates four key features. Users can **navigate** from one profile to another and browse through posts and switch between accounts. The *Posts* tab allows posts to be **sorted**

chronologically by account or by date. A **search** function by account name is available in both the *Profiles* and *Posts* tabs. Finally, all imported metadata can be converted into properties during import. These properties are integrated into a filtering function that enables researchers to select a subset of profiles or posts. Profile properties are linked to posts and can also be used to filter them. By providing both transversal access to the data (through navigation) and targeted one (via sorting, searching, and filtering), these features expand the ways in which researchers can view and analyze their corpora.

Returning to the source text. SMIQQDA integrates URLs of profiles and posts, enabling direct access to the original Instagram content. This allows users to explore comments, interactions, tags, hashtags, popularity dynamics, edits, and removals, supporting maximal semiotic contextualization and an ecological approach to digital data (Develotte and Paveau, 2017). This approach facilitates access to an image’s co-text, context, production, reception, and circulation, enabling close readings at every stage of the visual analysis.

3.5. Analyzing the data: annotate, categorize, combine with other tools

Annotation. SMIQQDA supports analytical work through two core features: free-text annotation and categorization. These functionalities streamline, standardize, and accelerate the research process (Lejeune, 2010; Page et al., 2022), positioning SMIQQDA within a mixed-methods research framework. The *Profiles* and *Posts* tabs include a free-text annotation space, which is automatically saved.

Categorization. SMIQQDA allows for the creation of clickable categories to code profiles and posts (see Figure 3, section 5). This makes it particularly well-suited for manual quantitative content analysis. The variables can be either mutually exclusive or not.

Combining different analysis tools. SMIQQDA facilitates the integration of both manual and automated tools, aligning with the mixed-methods approach. SMIQQDA is compatible with the textometry software **TXM** (Heiden et al., 2010). When exporting data, it generates image paths and converts timestamps to TXM’s date format. The *xlsx* export can be transformed into *xml* via a TXM macro for import into TXM. Conversely, SMIQQDA can import *xml* files with automated TXM annotations, converting them into properties that researchers can verify, correct, and interpret in context. In addition, exporting the SMIQQDA database in **xlsx format** allows users to manage large datasets in Excel (for instance, converting dates into months), perform statistical calculations, and integrate with automated visual analysis tools such as Panoptic (Bouté et al., 2024). Added or modified data can then be re-imported into SMIQQDA, updating the existing textual (meta)data, which is especially useful for large-scale corrections or reviewing automated analyses.

3.6. Managing data: exporting, creating sub-corpora, saving, archiving

Exporting and creating sub-corpora of texts and images. SMIQQDA enables exporting the database (*.sqlite*) to *xlsx* files, including all textual data, annotations, and categories. Users can define precise selections using properties and SQL queries to create specific sub-corpora. Each sub-corpus export automatically generates an *xlsx* file and optionally a folder with the corresponding visual data, organized by account, saving considerable time and avoiding duplication when sharing subsets of the dataset.

Saving. An automatic backup of the SMIQQDA database and previous exports is generated with each new export. This feature helps maintain a record of the data, analyses, and sub-corpora over time.

Archiving. Data from social media platforms are particularly unstable (Bouko, 2024; Emerit, 2016): they constantly evolve, can be modified, deleted, or censored. Platforms themselves may also change or disappear. By freezing and structuring the data, and making it readable independently of the platforms, SMIQQDA, thanks to its open-source nature, contributes to the long-term preservation of digital data and their analysis, ensuring they remain accessible and interpretable over time.

4. Types of data

Cross-platform appropriations. SMIQQDA was originally designed for Instagram data, which include specific metadata such as date, URL links, and engagement metrics (e.g., number of followers or likes). However, because these affordances are also found on other platforms (such as Facebook, Twitter, Reddit, LinkedIn, Telegram, and Tumblr), SMIQQDA can equally be used to analyze data from these social media. SMIQQDA supports the analysis of posts without images, making it applicable to platforms where visuality remains central but is not always present. SMIQQDA is also flexible enough to function without a profile-based corpus: it can be used solely to analyze collections of posts.

Extending SMIQQDA beyond social media posts. SMIQQDA is also compatible with other types of image-based corpora beyond social media. It can be used to analyze a wide range of text-image materials, including web pages, comics, illustrated books (such as children's books or art books), ephemeral media (collages, posters, flyers), news articles, and more. Given the central role of visuals in everyday life, the possibilities for adapting and repurposing SMIQQDA are extensive. The main challenge lies upstream, in the construction of the corpus, though this is by no means an insurmountable obstacle.

5. A SMIQQDA case study in multimodal analysis

This case study demonstrates how SMIQQDA functions as an integrative platform linking data collection, cleaning, annotation, multimodal analysis, and visualization. This platform enables constant back-and-forth between quantitative and qualitative methods. More specifically, it highlights the methodological contribution of SMIQQDA to the processing and analysis of a large, heterogeneous digital corpus drawn from Instagram, focusing on feminist and queer discourses, and notably in relation to colonialism. The corpus consists of three datasets: (1) approximately 1,500 profiles, (2) nearly 300,000 posts combining images (jpg) and textual (meta)data (xlsx), and (3) an OCR corpus extracted from images (around 270,000 entries). These data, collected through scraping, are organized into Excel files and image folders.

	Account : smash_HRP (Smash the heteracispatriarchy) 11/11 (11) Biography : I am here to party and to smash the heteronormative racist patriarchy © Picture Raphael Renter @raphi_rawr (https://unsplash.com/fr/@raphi_rawr) Url : https://research.flw.ugent.be/nl/tiago.joseph Posts : 500 Posts actually collected : 1 Posts missing : 499 Followers : 150 Follows : 26
---	--

Figure 2. Display of a profile in SMIQQDA

In a first step, SMIQQDA is used to clean and organize the corpus. All profiles are imported and manually coded using clickable categories in order to distinguish between individual and collective accounts, as well as their antiracist political positioning. This process makes it possible to filter the corpus and export only the relevant profiles into a new database. The posts associated with these profiles are then imported into SMIQQDA, where duplicates are automatically removed.

The integration of the OCR corpus constitutes a crucial step. Due to a mismatch between the total number of posts (300,000) and OCR results (270,000), manual matching would have been impractical. SMIQQDA allows for the automatic linking of datasets via shared identifiers. These consolidated data are then exported for lexicometric processing in TXM, where CQL queries enable the automatic annotation of posts related to the concept of colonialism. These annotations are reimported into SMIQQDA as properties that can be used as filters. A manual content analysis is then conducted to refine the TXM annotations and distinguish reported speech from direct speech.

SMIQQDA subsequently facilitates the articulation between quantitative text and visual analysis tools. Direct speech on colonialism is exported from SMIQQDA as a new subcorpus, which is then imported into the Panoptic tool to annotate visual features, such as the presence of specific colors. These annotations are reimported into SMIQQDA (see Figure 3). At this stage, the researcher cross-references textual properties, visual features, and metadata to analyze the corpus. For instance, they examine how the frequency of the lexeme “colonialism” and the presence of the color red evolve over time.

Account : [smash_HRP](#) (Smash the heteracistpatriarchy) 5/5 filtered accounts with 29 posts (total of accounts: 5; total of posts: 29)
 Post : 1/1 filtered posts → Go
 Number of images : 1/1 (Image)
 Posts : 500
 Posts actually collected : 1
 Posts missing : 499
 Followers : 150
 Follows : 26

	Date : 16/12/2022 12:31:49 Localisation : n/a Link to the post 120 👍 21 ✎ Legend : Feminisme must speak against colonialism © Picture Mihai Surdu mihaisurdu https://unsplash.com/fr/@mihaisurdu	ColorRed : ☒ RedYes ☐ RedNo
---	---	---

Figure 3. Display of a post in SMIQQDA, with properties imported from Panoptic (i.e. "ColorRed")

In addition to these automated processes, SMIQQDA supports in-depth qualitative analysis. Posts are explored through close reading, freely annotated, and linked to their original context via URLs, allowing the researcher to identify the symbolic meanings associated with the color red (love, danger, blood, violence, etc.).

Finally, once exported, the data and analyses enable advanced statistical processing. A hierarchical cluster analysis is carried out to identify which types of feminist and queer profiles (individual or collective, popular or not, identified as intersectional or not, etc.) contribute to the politicization of colonialism, and through which visual framings.

6. Conclusion

This article has highlighted the persistent tool gap in multimodal social media research, where manual approaches remain constrained (Baker, 2018; Rogers, 2021; Bouté et al., 2024). By introducing SMIQQDA, we address these limitations, offering an open-source tool that structures, enriches, and allows analysis for text-image datasets within a user-friendly environment.

As evolving software, SMIQQDA's free and open-source distribution invites collaborative development (Di Cosmo, 2022; Heiden et al., 2010). We encourage researchers to contribute to its future, with possible directions including: an OCR editing interface, support for additional multimodal formats (e.g., audio, video), broader import/export compatibility, enhanced annotation options (including localized tags), improved usability and installation processes, improved image naming schemes (facilitating image-text association), and multilingual documentation.

References

- Baker P. (2018). Language, sexuality and corpus linguistics: Concerns and future directions. *Journal of Language and Sexuality*, 7(2): 263-279.
- Barats C. (2024). *Corpus. Publicationnaire. Dictionnaire encyclopédique et critique des publics*. <https://publicationnaire.huma-num.fr/notice/corpus>
- Bouko C. (2024). *Visual Citizenship: Communicating political opinions and emotions on social media*. Routledge.
- Bouté É., Julliard V., Alié F., Gödicke D., Pailler F. and Ecrement V. (2024). PANOPTIC, un outil d'exploration par similarité de vastes corpus d'images. *Humanistica*, 2024: 1-7.
- Brennen J.S., Simon F.M. and Nielsen R.K. (2021). Beyond (Mis)Representation: Visuals in COVID-19 Misinformation. *The International Journal of Press/Politics*, 26(1): 277-299.
- Caldeira S. (2024). Exploring feminisms on Instagram: Reflections on the challenges and possibilities of incorporating digital methods strategies in feminist social media research. *Journal of Digital Social Research*, 6(1): 74-89.
- Despontin Lefèvre I. (2022). *Stratégies de communication et pratiques militantes dans le mouvement féministe en France au début du 21ème siècle: Étude de cas du collectif #NousToutes (2018-2021)*. PhD Thesis, Université Paris-Panthéon-Assas, France.
- Develotte C. and Paveau M.-A. (2017). Pratiques discursives et interactionnelles en contexte numérique. Questionnements linguistiques. *Langage et société*, 160-161(2-3): 199-215.
- Di Cosmo R. (2022). Construire le pilier logiciel de la science ouverte. In Open Science European Conference, editor, *Actes des Journées européennes de la science ouverte*, pp. 185-194.
- Emerit L. (2016). La notion de lieu de corpus: Un nouvel outil pour l'étude des terrains numériques en linguistique. *Corela. Cognition, représentation, langage*, 14(1): 1-28.
- Gomez-Diaz T. and Recio T. (2022). Research Software vs. Research Data I: Towards a Research Data definition in the Open Science context. *F1000Research*, 11(118): 1-37.
- Heiden S., Magué J.-P. and Pincemin B. (2010). TXM: Une plateforme logicielle open-source pour la textométrie – conception et développement. In Angermüller J., editor, *10th International Conference on the Statistical Analysis of Textual Data*, pp. 1021-1032.
- Jay M. (2002). Cultural relativism and the visual turn. *Journal of Visual Culture*, 1(3): 267-278.
- Joo J. and Steinert-Threlkeld Z.C. (2022). Image as Data: Automated Content Analysis for Visual Presentations of Political Actors and Events. *Computational Communication Research*, 4(1): 11-67.

- Joseph T., Stappers M., Gent University (2025). *SMIQQDA, Social Media Interface for Quantitative and Qualitative Data Analysis (Version 1.0)*. Zenodo, <https://doi.org/10.5281/zenodo.15642544>.
- Julliard V. and Saemmer A. (editors) (2022). *Étudier les communautés interprétatives et émotionnelles. Bricolages méthodologiques. Communication & langages*, 2(212).
- Kemp S. (2025). *Digital 2025: Top social platforms in 2025*. DataReportal, 2 April.
- Kress G. and van Leeuwen T. (2020). *Reading Images: The Grammar of Visual Design*. Routledge.
- Lejeune C. (2010). Montrer, calculer, explorer, analyser. Ce que l'informatique fait (faire) à l'analyse qualitative. *Recherches qualitatives*, 9: 15-32.
- Männistö A., Seker M., Iosifidis A. and Raitoharju J. (2022). Automatic Image Content Extraction: Operationalizing Machine Learning in Humanistic Photographic Studies of Large Visual Archives. *arXiv*: 1-25.
- Millette M. (2023). Théorie et méthodologie féministes pour la recherche en contexte numérique. *Communication. Information médias théories pratiques*, 40(2): 1-20.
- Münster S. and Terras M. (2020). The visual side of digital humanities: A survey on topics, researchers, and epistemic cultures. *Digital Scholarship in the Humanities*, 35(2): 366-389.
- Niederer S. and Colombo G. (2024). *Visual Methods for Digital Research: An Introduction*. Polity.
- Omondi G. (2025). Visual framing during crisis: A 10-year systematic review. *Corporate Communications: An International Journal*, 30(1): 139-161.
- Page M.J., McKenzie J.E., Bossuyt P.M., Boutron I., Hoffmann T.C., Mulrow C.D., Shamseer L., Tetzlaff J.M., Akl E.A., Brennan S.E., Chou R., Glanville J., Grimshaw J.M., Hróbjartsson A., Lalu M.M., Li T., Loder E.W., Mayo-Wilson E., McDonald S. and Moher D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, 372(71): 1-9.
- Page R., Barton D., Lee C., Unger J.W. and Zappavigna M. (2022). *Researching Language and Social Media: A Student Guide*. Routledge.
- Pandiani D.S.M. and Presutti V. (2023). Seeing the Intangible: Surveying Automatic High-Level Visual Understanding from Still Images. *arXiv*: 1-29.
- Pearce W., Özkula S.M., Greene A.K., Teeling L., Bansard J.S., Omena J.J. and Rabello E.T. (2018). Visual cross-platform analysis: Digital methods to research social media images. *Information, Communication & Society*, 23(2): 161-180.
- Pincemin B. (2020). La textométrie en question. *Le Français Moderne – Revue de linguistique Française*, 88(1): 26-43.
- PubGenius Inc. (2026). SciSpace. Suite # 217 691.
- Rogers R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*, 8(1): 1-23.
- Rose G. (2022). *Visual Methodologies. An Introduction to Researching with Visual Materials* (5th ed.). Sage.
- Souchier E., Candel É., Gomez-Mejia G. and Jeanne-Perrier V. (2019). *Le numérique comme écriture. Théories et méthodes d'analyse*. Armand Colin.
- Valyaeva A. (2023). AI Has Already Created As Many Images As Photographers Have Taken in 150 Years. *Everypixel Journal*, 15 August.