JDSR 6(2) (2024) 46-65

JOURNAL OF DIGITAL SOCIAL RESEARCH

10.33621/jdsr.v6i2.214

Researching visual protest and politics with “extra-hard” data

Suay Melisa Özkula[1], Janna Joceli Omena[2] and Radhika Gajjala[3]

[1] University of Salzburg, Austria
[2] Universidade Nova de Lisboa, Portugal
[3] Bowling Green State University, USA

✉ suay.oezkula@plus.ac.at

Abstract

A range of scholars have criticised scholarly tendencies to focus on “easy” data such as provided by the low-hanging fruit of Twitter hashtag networks (Burgess & Bruns, 2015; Hargittai, 2020; Tromble, 2021). As a result, digital social research has been said to create a glut of studies that favour particular platforms, data forms, and networking dynamics, choices that may create ‘digital bias’ (Marres, 2017). These issues are particularly significant in visual data as the implicit nature of visuality means that platform spaces, text, and networked uses of visuals contribute to how visuals are interpreted in digital environments. In response to this issue, we present and critically reflect on *new potentialities* in software-based visual research on protest and politics, including: (1) rich cross-project comparisons; (2) complementing platform data with on-the-ground engagement, and (3) quali-quanti visual methods. These allow for rich data journeys through multi-modality, hybridity, comprehensive data curation, reiterative image data collection and interpretation, and the inclusion of contextual reflections in focused visual research, elements that provide meaning, texture, and context (= extra-hard data). We argue that visual digital methods consequently have the potential to provide nuanced, robust, and versatile analysis of visual data, if not necessitate these in a post-API age in which easy data access is no longer a given.

Keywords: visual methods; extra-hard data; quali-quanti methods; cross-platform studies; digital methods; feminist approaches

1. Introduction

Due to the increasing availability of research software, open source tools, code repositories, and opportunities for conducting research using computational tools, digital social research has been said to have undergone a “computational turn”, i.e. considerations of computational mediums as carriers of meaning (see Berry, 2011; Rieder, 2020; Omena, 2021). While these new opportunities have been very much embraced by the scientific community, concerns have been raised about scholarly tendencies to focus on “easy” data such as provided by the “low-hanging fruit” of hashtag networks (Burgess and Bruns, 2015; see also Hargittai, 2020; Tromble, 2021). As a result, digital social research has been claimed to create a glut of studies that favour particular platforms, data forms, tagging mechanisms, and networking dynamics. Such preferences may create partial views into digital practices or phenomena, also described as "digital bias" (Marres, 2017). As part of these concerns, software-based research has been the target of some critique. While some research has already discussed and shown more contextualised efforts such as cross-platform approaches (Rogers, 2018; Venturini et al. 2018; on visuals: Pearce et al. 2020, Colombo, Bounegru and Gray, 2023) and situational analytics (Marres, 2020), doubts remain as to whether easy data research necessarily or consistently produces data that is rich, contextualised, or nuanced (see, e.g., Özkula, Reilly and Hayes, 2022; Tromble, 2021).

In part, researchers' choices in the design of software-based research are grounded in the access provided by individual platforms or the funds provided by institutions. As such, scholars in the field have already amply addressed and acknowledged the critiques on the prevalence of certain types of research, above all, as emergent from easily accessible data such as has historically been the case with Twitter/X. Additional complications (i.e. beyond persistent API access restrictions) stem from differences in technological grammars, algorithmic influences, and the treatment of platforms as data sources rather than sociocultural (deep) vernacular environments (see Burgess et al. 2023; Gibbs et al. 2015; de Zeeuw and Tuters, 2020; Rogers and Giorgi, 2023). These issues are particularly salient in visual data as the implicit nature of visuality means that individual platform spaces, textual elements, tags/labels/hashtags, and networked uses of visuals all contribute to how visuals are interpreted in digital environments. As such, visual data are to a significant degree *context-dependent* and rely on contextual and technical readings.

This paper therefore presents and critically reflects on *new potentialities in visual research on protest and politics.* It will suggest that contemporary research presents numerous opportunities for **'extra-hard'** data in digital visual research, including: **(1)** hybrid engagements that allow for triangulation and contextualisation; **(2)** rich data journeys through reiterative image data collection and interpretation processes, comprehensive data curation, and multi-level enquiries; **(3)** comparative research such as multimodal, intersectional, and extended multi-sited or cross-platform visual research, and **(4)** the inclusion of contextual reflections in focused visual research, such as researcher positioning, platforms' (sub)cultures of use and their knowledge about computational affordances and limitations. This perspective challenges the notion of viewing data in isolation from its digital environment and socio-technical context. The paper therefore proposes the application of rich cross-project comparisons, complementing platform data with on-the-ground engagement, and quali-quanti visual methods, under the premise that, in the post-API age, 'easy' visual data on politics and protest is neither a viable possibility nor representative of visual cultures, above all when these are disembedded or de-contextualised from their site of circulation.

These arguments rest on the premise that (limited) data access and a myriad of open-source tools have distracted scholars from the foundations/tenets of the computational turn. It remains easy to divert attention from the language, practices, and technical substance of research software and how these are entangled with visuals and the environments they originate in. This is despite the potential of software-oriented research for situating data, which creates more nuance, robustness, and versatility towards contextualising visual politics and protest. To demonstrate this, the paper will start with a review of Bruns and Burgess' easy data hypothesis and a critical overview of schools of method (incl. visual methods) in digital social research. It will then briefly describe the data and methods that form the basis for the reflections presented here. Subsequently, it will reflect on the new potentialities of digital visual methods research. The paper closes with reflections, recommendations, and a discussion on extra-hard visual data in contemporary digital social research.

https://doi.org/10.33621/jdsr.v6i2.214

2. Literature review

2.1 Easy data & hard data critically reviewed

Despite the potential of digital methods, scholars have highlighted a range of concerns around trends towards "easy data", a hypothesis famously coined by Burgess and Bruns (2015). In their essay, Burgess and Bruns (2015, n.p.) problematise the ways in which "the Twitter API mediates access to Twitter data actually inscribes and influences the macro level of the global political economy of science itself, through re-inscribing institutional and traditional disciplinary privilege". They consequently argue that (1) methods have substantially changed through the computational turn in social science and humanities research, (2) that more easily available, accessible, and modifiable data have led to an increase in "easy data studies" in digital methods research (as linked to API access economies), and (3) that these socio-economic circumstances and the related choices in research design have consequently resulted in a prevalence of research on specific media dynamics, above all narrow-period or recent Twitter-based hashtag and @reply network research based on keyword searches. Twitter/X access has since been subject to a range of changes, first providing open academic access to Twitter data on a case-by-case request basis, and then the creation of a paywall for research data. These changes have been and still are affecting what platform spaces are studied and by whom. Other platforms such as Facebook and TikTok have similarly been subject to changes in API access for research purposes, developments that have been amply discussed and exploited by researchers. While this may mean that Twitter/X may no longer be the golden child of API access and a goldmine of social media data in future research, it suggests that "easy data" access (by whichever platform) will ultimately determine who and what digital methods researchers study.

In comparison, only a handful of scholars in digital social research are able to obtain "hard data", that is "more comprehensive, longitudinal data sets and/or any of the "missing" metadata" (Burgess & Bruns, 2015, n.p.) due to paywalls, and the lack of technical infrastructure, advanced access, or practical knowledge on handling these kinds of data. While Burgess and Bruns do not establish these concepts as a hard-and-fast dichotomy, their essay draws attention to how the *research access economy* and subsequent methodological choices produce new forms of digital bias, which (as we argue) may potentially ignore new potentialities afforded by digital methods. They also suggest that following what Bruns (2019) describes as the "APIcalypse", easy data access may no longer be an option, as all data become subject to socio-economic factors that require researchers to invest effort into critically reviewing what data have been obtained, omitted, and what exactly these data represent.

These methodological considerations have, at times, given rise to concerns around what may realistically constitute good, valid, or 'solid' digital social research (e.g. Burgess and Bruns, 2015; Özkula, Reilly and Hayes, 2022). While small-scale immersive research may be criticised for limits in scale or representativeness, the growing emphasis on large-scale statistical data may also signal a 'positivistic turn' that overlooks the benefits of qualitatively oriented research. In comparison, this paper suggests that there is both necessity and opportunity to engage with wider digital environments by drawing on the affordances of different research modalities, above all the case in visual artefacts due to their reliance on contextual readings. This includes considerations of a) who the actors are in individual digital communities, networks, and platforms, b) how these actors engage in those spaces, i.e. their individual practices and dynamics, c) how visual artefacts are produced, circulated, and read, as well as d) how these dynamics emerge or change in response to research software and platform affordances[1]. In accumulation, these considerations allow for conducting more nuanced and contextualised research that demarcates a new age of extra-hard visual data.

[1] I.e. "(...) the perceived actual or imagined properties of social media, emerging through the relation of technological, social, and contextual, that enable and constrain specific uses of the platforms" (Ronzhyn et al. 2022, 14).

https://doi.org/10.33621/jdsr.v6i2.214

2.2 'School of method': A brief review

Over the past decade, scholars from various disciplines have advanced digital humanities and social science methods repertoires with a focus on big data, computational techniques, and empirical evidence from these studies. To lay the foundation for this paper and clarify our understanding of digital methods, we will provide a 'quick-and-dirty' overview of three schools of software-based methods.

2.2.1 Methods school 1: Cultural analytics

The first school of method has broadly been described as "cultural analytics", a field spearheaded above all by Lev Manovich's introduction to the quantitative study of cultural patterns on different scales. In this methods school, research questions are raised after the mapping and measuring of fundamental characteristics associated with professional or user-generated datasets. Big data samples, such as cultural artefacts including digitised and digital images, are subject to different visualisation techniques, as demonstrated by seminal projects like Time Magazine Covers (2009)[2] and Selfiecity (2014)[3]. These cultural analytics projects provide insights into cultural practices, representations, and expressions through methods of media visualisation such as image montage or sampling versus data summarisation (Manovich, 2020).

2.2.2 Methods school 2: Computational social sciences

Similarly relying on computational methods and tools is Lazer's computational social sciences (CSS), where large-scale human behavioural data guide the study of social phenomena. CSS uses mathematical models and statistical analysis such as machine learning, natural language processing, and algorithmic or network analysis oriented methods (Lazer et al. 2020). Theories and methods from computer science feed into this school of method. While CSS provides valuable insights into social phenomena and cultural practices, it draws on techniques that a wide array of social scientists are not trained in. Data storage infrastructures, domain expertise, and techniques for processing big data are not yet part of everyday teaching practices of social sciences and humanities schools. As such, computational social sciences is a recently emerging and rapidly spreading field, for which the required research skills are not as salient (compared, for example, to cultural analytics) across disciplines.

2.2.3 Methods school 3: Digital methods

The third school of method relates to what has been dubbed "digital methods" (DM), a term famously heralded by Richard Rogers (Rogers, 2019; Rogers and Lewthwaite, 2019) and his establishment of associated research labs at the University of Amsterdam, which we term "the Amsterdam School". It is characterised as "a research practice crucially situated in the technological environment it explores and exploits" (Omena, 2012, p. 24), and draws conceptually from Science and Technology Studies, Actor-Network Theory, and Software Studies. In DM, the web is both a data source and a research site, a place to ground findings (Rogers, 2013), focusing on the technicity of research tools and software (Omena, 2021). For example, in his early writings, Rogers (1996) suggests repurposing digital objects and web cultures for understanding social and cultural phenomena, and shows how digital objects, web data, and practices can ground research findings. Therefore, according to Rogers (2013), before prioritising computational techniques and data (not necessarily 'big' data), there is an invitation to develop ways of thinking that work in tandem with the medium and what it has to offer.

While all three schools of software-based social research methods have made considerable contributions to empirical, methodological, and epistemological research practices across disciplines, this paper focuses primarily on the third school of method (DM). DM do not necessarily stand in separation to the other schools of method or qualitatively oriented internet research methods (examples: digital

[2] http://lab.culturalanalytics.info/2016/04/timeline-4535-time-magazine-covers-1923.html
[3] https://selfiecity.net/#

https://doi.org/10.33621/jdsr.v6i2.214

ethnography, online diary-keeping, over-the-shoulder interviews). In fact, recent digital methods research has sought to combine diverse approaches for triangulation and contextualisation. These trends form the premise for the arguments presented here. The DM School was also chosen as the primary focus due its media-ecological approach that considers practices of data production in web environments. In comparison, CSS approaches rely strongly on computing backgrounds and the possibilities afforded by independent computing. As such, the wide application of DM stems in part from it not de facto requiring advanced computational skills such as programming.

2.2.4 Mapping visual methods

Visual methods have developed alongside these schools of method. There has been a noticeable shift in image analysis, moving from individual or qualitatively sampled images to collections, involving both quantitative and qualitative methods[4] (see Manovich, 2008; Rose, 2016; Ricci et al. 2017; Colombo, 2019). Central to this transformation is the advancement of software and algorithmic techniques for image processing, visualisation, and interpretation, which, in turn, has reshaped how we conceptualise and study visuals. In light of these developments, this section provides an overview of interpreting visuals in digital environments and delves into three critical aspects of visual method development over the past decades. To support this discussion, we introduce the visual methods grid (see figure 1). It presents ways of interpreting image collections using digital methods, and indexes the methods, associated software, and visual models used for that purpose.

(1) Software-making for image analysis

Software, especially with the advent of macros, network visualisation software, and plugins in the early 2000s, has played a pivotal role in mechanising analytical processes for visual methods. In 2007, ImageSorter by the Visual Computing Group facilitated the *zooming in and out* technique by arranging images according to colour, name, size, and date. In 2011, the emergence of macros by the Software Studies Initiative, designed to run within ImageJ (Rasband, 1997; Schneider, Rasband and Eliceiri, 2012), marked a significant step towards employing computational methods for understanding digital media culture, i.e. Cultural Analytics (Manovich, 2020). ImageJ, acquired new capabilities with macros like ImagePlot and Image Montage (Software Studies Initiative, 2011; Manovich, Gianchino and Chow, 2012). These macros assist visual methods in extracting meaning from images and their metadata (see figure 1). In 2012, the Yale Computer Graphics Group developed a plugin for the then newly born network visualisation and exploration software Gephi (Bastian, Heymann and Jacomy, 2009). This plugin –ImagePreview– has opened up new horizons for visual methods, facilitating, for example, network vision analysis (see Omena et al., 2021) and influencing the development of other research software, like Memespector Graphical User Interface (Chao, 2021).

Even so, software for image analysis may offer only specific perspectives on an image collection, including an understanding of what is *in* the images, their relational, cultural, temporal and technical aspects (using metadata), and also their web context (see figure 1). Classic examples include grouping images by colour (the image itself) or by associated web entities (utilising web detection algorithms) and using engagement metrics (image metadata) as a measure of significance (see figure 1). Each of these perspectives offers a snapshot of what might represent an image collection, as its meaning-making depends on the specific groups, networks, and web environments in which images are shared (see Burgos-Thorsen and Munk, 2023; Rogers and Giorgi, 2023).

(2) AI methods and visual models for making sense of image collections

Another significant development in visual methods have been pathways towards understanding AI methods for interpreting image collections and defining visual models for analysing them. Algorithmic

[4] For the theoretical foundations of quali-quantitative methods, see the work of Venturini (2024).

techniques, the cultural context of image production, and medium-technicity features facilitate visual methods yet carry significant epistemological implications for scholarship (Rieder and Röhle, 2012; Omena, 2021). While adopting a discerning approach to algorithmic techniques for interpreting image collections is paramount (see Burgess et al. 2021), it is imperative to recognise the limitations of software and data visualisation in providing comprehensive perspectives of visuals. Figure 1 illustrates common methodological choices for studying images *en groupe* (Colombo, 2019), drawing on emerging literature on visual methods. Here, different arrangements of images enable distinct analytical procedures, with visual models as research devices and initial points of exploration to gain insights into data (Gray et al. 2016; Colombo et al. 2023; Rogers, 2021). For instance, in image grouping methods, image clusters can be analysed by colour pattern or the outputs of AI methods. Exploring image metadata with methods like audience site analysis provides additional aspects such as engagement metrics or temporal insights. As such, grouping and classifying images based on their colour patterns, associated digital objects, and web detection algorithms enables: a) the exploration of imagery linked to specific places, issues, or events; b) conducting research on the audience response to images, including an analysis of the practical impact of online images; c) comparing competing, controversial, or antagonistic visual spaces, as observed in program and anti-program research; and d) advancing a form of image circulation research (see Colombo et al. 2023).

(3) New analytical techniques, new critical considerations

Understanding emerging analytical techniques is imperative for visual methods. The complexity of visual methods evolves from speculating about algorithms for image analysis to additional efforts required before delving into visual exploration, as observed in lexicon-demarcated image collections. Situational Analytics, as outlined by Marres et al. (2023), proposes semi-automated methods that aid in uncovering situations (i.e. context) from vast social media datasets. Another challenge arises, for instance, in networked image analysis (Niederer and Colombo, 2019), where formatting, filtering, and recommendation systems of image sources become crucial. Similarly, when visualising images and associated vision AI outputs as networks, epistemological concerns about the process of network building and visualisation emerge (Omena, 2021). This impacts the interpretation of the network based on its shape, as well as the positioning, colour, and size of nodes (see Venturini, Jacomy and Jensen, 2021). Whether it's visual situational analytics, networked image analysis, or network vision analysis, each of these methods presents unique challenges and epistemological considerations for navigating visual methods. They account not only for the contextual background of images but also for the research software, AI methods, and visual models that researchers must choose to make sense of them (see figure 1).

https://doi.org/10.33621/jdsr.v6i2.214

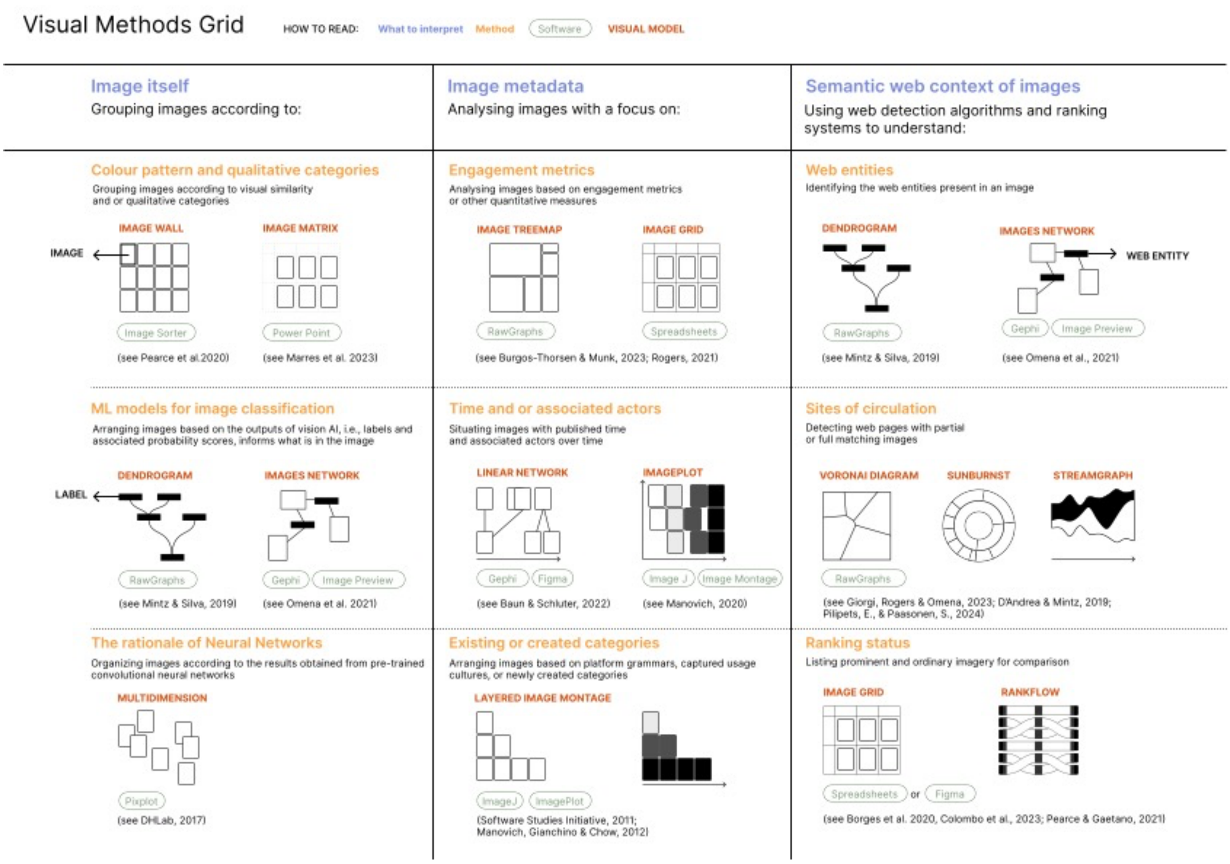

Figure 1. Visual methods matrix (created by © Janna Joceli Omena and Beatrice Gobbo).

2.2.5 Extra-hard visual data?

The different schools of method provide background for concerns around easy data, digital bias, and the historical Twitter/X reign (arguably a field in transition). These paved the way to a new critical approach to research methods as influenced by access economies that have come to affect, if not at times dominate, researchers' methodological decision-making. In some ways, computational methods as per Lazer have created a partial solution to these issues through a radical hacktivism-based approach that overcomes API limitations. Even so, hacking may not be a necessary or the only solution to these issues, and the same may apply to digital policies that seek to remedy these influences on the research community, as there are many ways to enrich software-based research - the endeavour of this particular article. As shown in the different schools of method, some research has already exploited these new opportunities, for example through (1) cross-platform research; (2) wider mapping efforts; (3) longitudinal, multi-sited, multimodal, and creative research; (4) approaches that consider contextual aspects of digital methods research, e.g. Marres' (2020) "situational analytics" that applies situated research (i.e. understanding a situation) to media-ecological phenomena; (5) triangulation with qualitative methods; and (6) collaborative research that includes industry and platform provider perspectives (see Burgess and Matamoros-Fernández, 2016; Marres, 2020; Rogers, 2018; Venturini, et al., 2018; on visuals: d'Andrea & Mintz, 2019; Pearce et al. 2020, Colombo, Bounegru and Gray, 2023; Puschmann, 2019). While some of these works centre on images, visuality has hitherto been of comparatively little concern for these considerations (compared to software-based methods more widely).

This paper builds on these works through a consideration of data interpretation in association with vernacular (sub)cultures of use and not purely or predominantly guided by top-ranking content or engagement metrics. While some research has already applied such comprehensive approaches, the specific complexities of visual data and methods suggest that holism is a necessary part of visual research. Even where data may be easy to obtain (i.e. easy data *access*), the contextual dependence of visual data requires hard data *readings*. This proposal of a holistic view therefore questions the 'spirit of easy data' in visual research and aims to develop an awareness of the pillars of the DM approach (Omena, 2021) through three distinct but related facets of digital fieldwork: an understanding of platform grammatisation, (sub)cultures of use, and the affordances and limitations of computational tools (ibid). Through this approach, we hope to introduce new potentialities that move away from "methodological archetypes" in research that centre on network-based Twitter data (Özkula, Reilly and Hayes, 2022) towards more diverse visual research.

3. Note on case studies

The methodological reflections presented here stem from several standalone research projects conducted by the authors between 2020-2023. The projects combine a range of different methodological approaches that include platform-based data collection, ethnographic observation, qualitative semi-structured interviews, and diverse forms of both quantitative and qualitative data analysis. Protest case studies were chosen for two reasons: (1) the significance of sociocultural and socio-political contexts in how visual data on politics and protest need to be read and interpreted, and (2) based on the critical and controversial visualities political imagery produces. The case studies also reflect aspects of the new potentialities we list earlier in this paper.

These projects (see figure 2 for visual protocol) include (1) data visualisations of the 2020 'Shaheenbagh protests' **(SBP)**[5] that show the amplification of the presence of local community women in Shaheenbagh, Delhi, India; (2) a multi-author comparative endeavour on digital-visual misogyny that combines insights from four distinct projects with individual research designs **(FemVis)** - for full methodology see Özkula et al. (2024); and (3) visual quali-quanti methods to infer insights from bot-following networks regarding Brazilian political antagonist debates **(Bolsobots)** - for full methodology see Omena et al. (2024). These projects (cf. figure 2) address four challenges: comparing like-for-like data, validating information and contextualising data by connecting these with offline activities, studying deleted YouTube content, and moving beyond bot-score detector analysis by operationalising image collection-based analysis and techniques. The three cases presented here address these issues as follows: the SBP case study shows that a flat and easy reading of the surface quantitative indicators without algorithmically contextual understanding risks giving more importance to the social media users as activists than to the local women who were actually on the streets offline; the FemVis projects highlights opportunities for triangulating contexts for visual phenomena across different cultural and medial contexts; the Profiling Bolsobot Networks project offers a new methodological protocol that moves beyond bot-score detector analysis, reimagining bot studies with digital methods.

[5] A full methodological account can be found in Gajjala et al.'s article in this special issue.

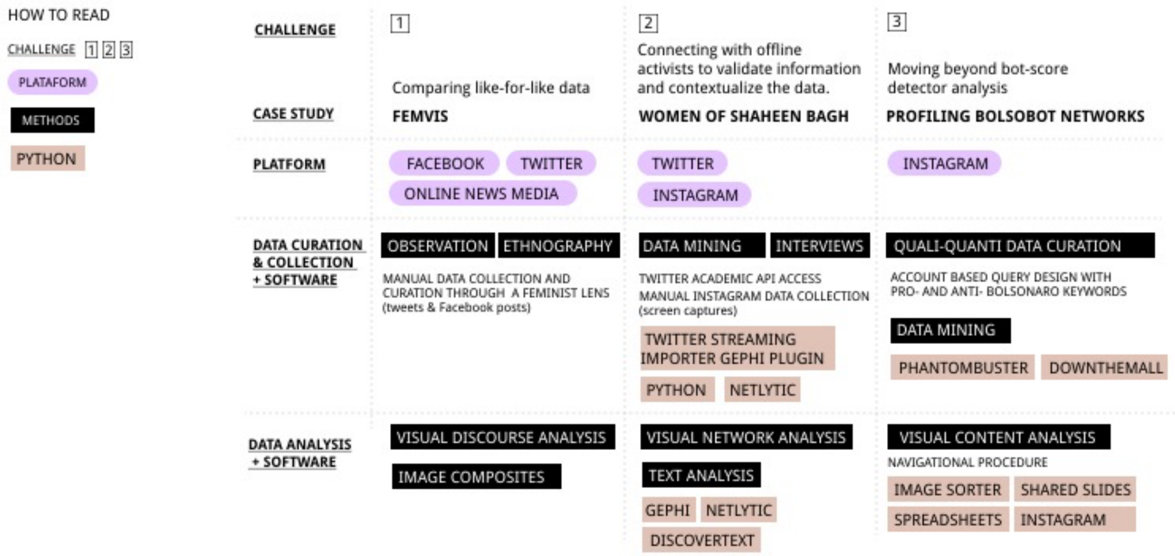

Figure 2. Visual protocol of the individual case studies (diagram created by Janna Joceli Omena and Beatrice Gobbo)

The projects presented here were conducted individually, but interlink through their shared focus on visual protest and/or politics. While visual protest and politics are not subject to a common definition or political category, we broadly define these as digitally mediated political activity that draws on and is to a significant part narrated by visual artefacts such as, for example, digitally native or digitised photographs, cartoons, memes, videos, gifs. This definition follows on from Karatzogianni's (2015: 1) definition of "digital activism' as digital technology use for "political conduct aiming for reform or revolution by non-state actors and new socio-political formations such as social movements, protest organisations and individuals and groups from civil society, that is by social actors outside government and corporate influence", but considers state-actors in formal politics as part of the visual ecology of protest and politics (see also Bennett and Segerberg, 2012). Visual politics and protest may respectively include, for example, the spreading of political messages on various social media platforms and their sub-spaces through visual artefacts, visual commentary in original digital posts as well as in responses to these, visual-performative aspects of political movements that are digitally generated, archived, or (re-)shared, visualisations of political information and/or mis-/disinformation, or visual aesthetics of political action (see Baun et al. 2022; Geboers et al. 2020a, 2020b; Nissenbaum and Shifman, 2017; Philipps, 2012).

Projects on politics and protest were chosen for this particular endeavour for reasons tied to their critical social impact as well as their dependence on contextual factors. First, political images may contain both explicit and implicit political messages that are encoded (social semiotics), but not textually articulated. They are, in that sense, employed towards instigating specific political ideas, attitudes, or behaviours, within certain geopolitical contexts that require *decoding*. Second, the interpretation of these visuals is network- or community-dependent, particularly where these are polysemic (e.g. political memes and cartoons). As such, the interpretation of political or activist visuals depends on the original meanings encoded by their creators, their reframing, editing, or subversion by users who share or circulate them, as well as their contextual readings by the audiences within the specific political narrative spaces they engage in.

4. New potentialities for rich digital visual research

4.1 Comparative research (FemVis)

Traditionally, the label of comparative research has been applied to research applying the same research design to comparable entities (i.e. 'like-for-like') such as countries or distinguishable social groups (e.g. on the basis of their demographics). Digital social and visual research challenge comparability to a certain extent, in part due to the at times missing informational insights that allow for regional or demographic markers to be inferred, but also due to differing platform grammars and affordances that conflict with the 'like-for-like' principle. Due to these differing logics, in the context of comparative digital social research, triangulation additionally relates to (1) cross-platform research (rather than simply multi-platform; with platforms as locational or demographic proxies), (2) intersectional research and studies that apply positionality and 'demographic sensitivity' (i.e., who is studied in what kind of demographic space, and how do researchers position themselves in these spaces), as well as (3) multiple forms of analysing and visualising digital data (since digital data collection and analysis may overlap). These bases for comparison embed visual data within location- or demographic-based user practices.

To illustrate, in the FemVis project, standalone projects were compared on the basis of a specific digital-visual practice - visual misogyny, which was expressed through gender-ideological visuals or visuals recontextualised to serve hate, abuse, and political violence. The case studies (Greta Thunberg memes in the DENY Facebook group; Fanquan Girls meme-wars in the social movement on the Hong Kong Anti-Extradition Law Amendment Bill; visual artefacts shared on Twitter under the hashtag #SisterIDoBelieveYou; and cartoons of Grace Mugabe relating to presidential succession across seven African countries), had been designed and conducted in isolation. The basis of comparison consequently lay in their joint focus on a visual practice and grounding in feminist methods with differences in the individual methods as well as their individual media contexts, for example the individual platforms on which research was conducted. The latter was taken as a point of comparison in light of differing visual platform affordances.

In doing so, the collaborative efforts allowed to identify visual misogyny as a cross-platform phenomenon that (a) took place across different cultural and medial contexts, (b) was expressed in different formats, but (c) was also shaped by different platform affordances that produced varying kinds of visibilities, and (d) analysed through different methodologies that drew on software in myriad ways depending on what the specific research design required. For example, the cross-project comparison produced insights into how visual misogyny was negotiated and made visible to either a wider public through hashtags or limited to a more closely knit community gatekept by Facebook group moderators. It also allowed for views into differing sub-cultures of use in these platform spaces, i.e. horizontally comparative data, including different aesthetic practices and the limitations of single-modal research (e.g. single-platform). That is not to say that research can become entirely limitation-free if proper procedure is followed, nor that purely quantitative or qualitative research is by its very nature limited. Despite calls for quali-quanti approaches, 'quanti' is not by default objective nor is 'quali'' necessarily subjective as they are both situated, nor does the method by itself determine rigour. However, the cross-project comparison allowed for an observation of contextual differences such as platform, region, and aesthetic.

A more developed research design would in principle have allowed for the consideration of all these projects as part of a more unified comparative research design, albeit also subject to different time and funding constraints. It may, as such, have been considered the ideal case. Even so, the distinct research designs with different research teams allowed for considerations of positionality and access across different cases, as well as for a comparison of data produced through platform-tailored methods. For example, in some of the visual misogyny case studies such as the Greta Thunberg hate speech in the DENY Facebook group, the researchers became vulnerable due to their specific positioning (e.g. female, international) in relation to the researched group (male, US-based) as well as the platform's informational insights (Facebook, in overt research) into the researcher background (see figure 3). As such, benefits may also be gained in digital-visual research that triangulates cases post-data collection. This is especially the case with visual research since the context-dependent nature of visual data means that contexts such as individual platforms, media ecologies, and connected geopolitical regions provide lenses for comparison in the first place.

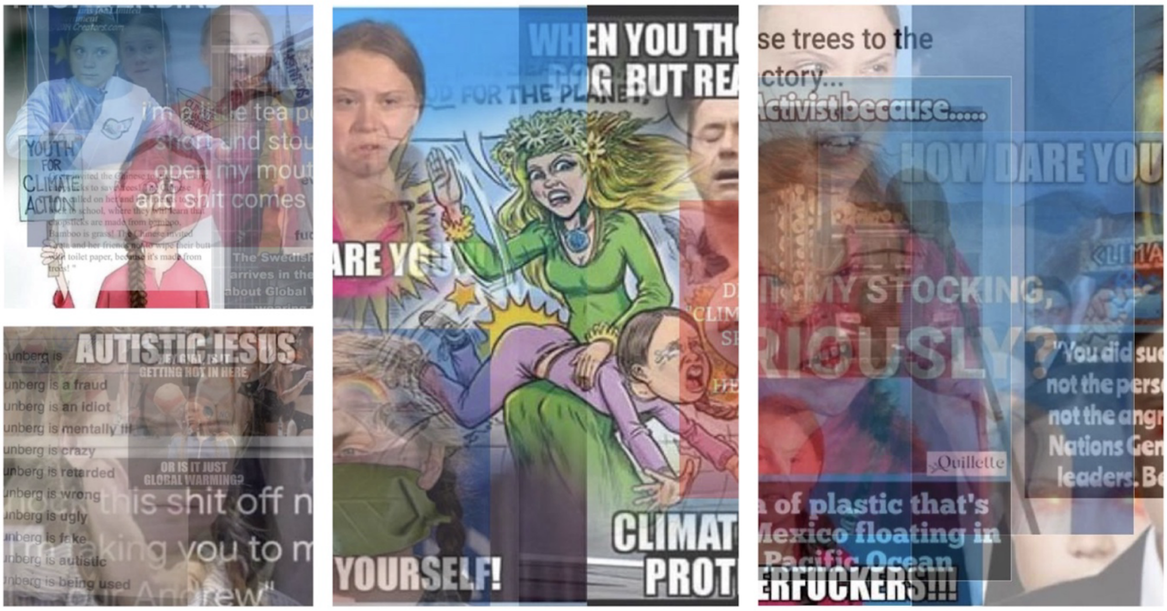

Figure 3. Example of unmoderated visual misogyny in the DENY Facebook group (image composite created by Suay Melisa Özkula)

A similar basis for comparison was Pearce and colleagues' (2020) "visual cross-platform analysis" (VCPA). There, visuals relating to the keyword 'climate change' were scraped from a set of five platforms: Instagram, Facebook, Tumblr, Reddit, and Twitter. Images were mined in separate but platform-appropriate research designs that followed the individual platform affordances and logics, for example hashtags on Twitter versus pages on Facebook. As such, platforms became the basis of comparison. Along similar logics, the visual misogyny project utilised the differences in research designs for comparisons of visual misogyny across different contexts. This means that separate situated and context-specific designs (i.e., spatially anchored, e.g. nation or platform) provide opportunities for triangulating contexts for visual phenomena (here: platformed visual misogyny).

4.2 Social media users amplifying a protest movement in India (SBP)

Other potentialities for extra-hard data were identified in the combination of platform data with direct user engagement, which allowed for the collection of contextual information on the digital dataset. In this case, the research team examined how transnational Twitter/X users contributed to the amplification of a protest movement started by a group of female activists who were staging a protest offline in India. The amplification of their concerns created global visibility, made possible because of the involvement of social media users on-site and the strategies adopted by Twitter/X activists both on site and transnationally. The multiple sites, researcher positionalities, and access options offline and on other platforms provided views into different aspects. Thus, this approach benefitted strongly from the inclusion of contextual reflections in focused visualisations of the digital data scraped including researcher positioning, platform affordances, and information gathered through several offline contexts.

https://doi.org/10.33621/jdsr.v6i2.214

The study used data network visualisation tools, such as Gephi and Netlytic, to mine data and examine smaller sections of larger datasets, and found tweets that cross-referenced hashtags such as #shaheenbaghprotests, #dadisofshaheenbagh, #womenofshaheenbagh, and #shaheenbaghdadis. These were hashtags that amplified the offline presence of women from the community who were protesting against the CAA/NRC (Citizenship Amendment Act/National Registry of Citizens) policies being voted on at that time in India. However, the tweet network by itself could not provide insights into how the visibility was produced and how various actors engaged with the algorithmic interface to amplify the movement. The research team therefore carefully examined each networked group of interactions to identify patterns and then conduct interviews with offline participants. It was only by doing careful offline work that allowed for the interpretation of the Twitter/X activity in context and thus how multiple factors contribute to how a movement - externally visible or not - plays out.

To illustrate, figure 4 presents insights into the specifics of digital practices by highlighting how network visualisation can inadvertently magnify the impact of certain actions or individuals. In this bipartite network, the nodes represent Twitter/X accounts, connected through mentions. The blue node represents a Bollywood actress, @reallyswara, who has been mentioned by several other accounts (green nodes), thereby increasing her visibility within the network. Despite her lack of active tweeting during the offline event, she was still featured among the top ten. The methodological limitation here lies in understanding that @reallyswara received such a significant number of mentions online even though at this time she did not seem to be present on social media or on-site offline. However, background research to unveil the role of this particular user within the offline context showed that her role as celebrity amplifier extended beyond social media and that the amplification of her Twitter/X handles was the result of public comments of support made by her. While this accidentally contributed to the global amplification of the SBP, it was not part of the planned strategy of the offline activists. Yet her own location as a visible bollywood actor contributed to her being visible as well. When the team of researchers reached out to various people who were present on-site either as activists or onlookers, it transpired that there were many interventions by people who did not belong to the local community that were getting highlighted transnationally because of the way social media algorithms create visibility.

Figure 4. Network of Twitter/X mentions of @reallyswara that served to amplify this user's presence as one of the top ten in the particular time span

https://doi.org/10.33621/jdsr.v6i2.214

4.3 Quali-quanti visual methods (profiling bolsobot networks)

Another example of how digital visual research can be enhanced is through the application of visual quali-quanti methods (as developed by Colombo, Bounegru and Gray 2023; Rabello et al. 2022; Omena et al. 2024; Venturini, Jacomy and Jensen 2021), which embrace a triple principle consisting of rich data curation, reiterative image data collection and interpretation (i.e. extra-hard data). This potentiality is exemplified by the Public Data Lab project Profiling Bolsobot Networks. The project[6] focuses on investigating the activities of bolsobots, social media accounts that promote or demote Jair Bolsonaro's political agenda, in the context of online and offline political debates in Brazil. The project utilises quali-quantitative methods to define bolsobots as accounts operated by specialised marketing teams, hackers/activists, campaign supporters, or paid workers, that are either partially or fully automated. These accounts often use profile pictures of Bolsonaro that depict him as a (un)likeable persona, as well as national, political, and cultural symbols.

From the visual methods implemented in this project, image analysis follows colour cluster similarity analysis. The first step involves designing the queries to scrape Instagram data by searching for and testing keywords used in online and offline conversations about Bolsonaro. The initial list of keywords indicate positioning efforts (Akrich and Latour, 1992), reflecting “the connections people are currently making of a word or phrase, whether established or neologistic” (Rogers, 2019, p.37). For instance, ‘Bolsonaro mito’ (myth) is a form of praise that depicts him as a heroic figure, while ‘Bolsonaro genocida’ (genocidal) is a clear accusation (figure 5). The project enhances the list of keywords by incorporating an interactive approach, enabling the researcher to generate new search queries to refine the keyword selection, e.g. in figure 5, ‘memes mito’ (on the left) and ‘anti bozo’(on the right).

In this context, ‘mito’ is a nickname adopted by Bolsonaro supporters and is associated with his image as a tough-talking, anti-establishment figure who promised to combat corruption, reduce crime, and restore Brazilian traditional values. Nevertheless, the former president came to be referred to as ‘Genocida’ because of his controversial handling of the health crisis during the pandemic. Healthcare professionals and various sectors of society accused him of failing to respond to necessary health measures, downplaying the severity of the virus, promoting unproven treatments, and showing resistance to vaccination efforts. Figure 5 demonstrates that these contextual keywords also ressonate on Instagram search results, indicating specific usage cultures on account creation. When searching for terms such as ‘Bolsonaro’ and ‘mito’, or ‘Genocida’, a substantial number of memetic and bot accounts are recommended. After verifying and accepting the platform’s recommendations, 70 pro- and anti-Bolsonaro accounts were listed. They serve as entry points to building the project datasets with Phanthombuster: who these accounts follow and their public profile description. Downloading the images and knowing in advance the short life span of Instagram image URLs[7] was therefore a priority. On the same day that the project succeeded in scraping the following networks (with a maximum of seven thousand accounts per seed), all images were accessed and downloaded through their uniform resource locator (URL). This process, from query design to dataset building, reflects the qualitative fronts of digital methods and what we describe here as a comprehensive image data curation process (i.e. extra-hard data).

[6] https://publicdatalab.org/projects/profiling-bolsobot-networks/
[7] The lifespan of image URLs can vary between different platforms. Platforms may remove or expire image URLs to free up storage space or manage their content or they can use dynamic URLs generated on the fly rather than being assigned a permanent URL. Also, users can delete online content.

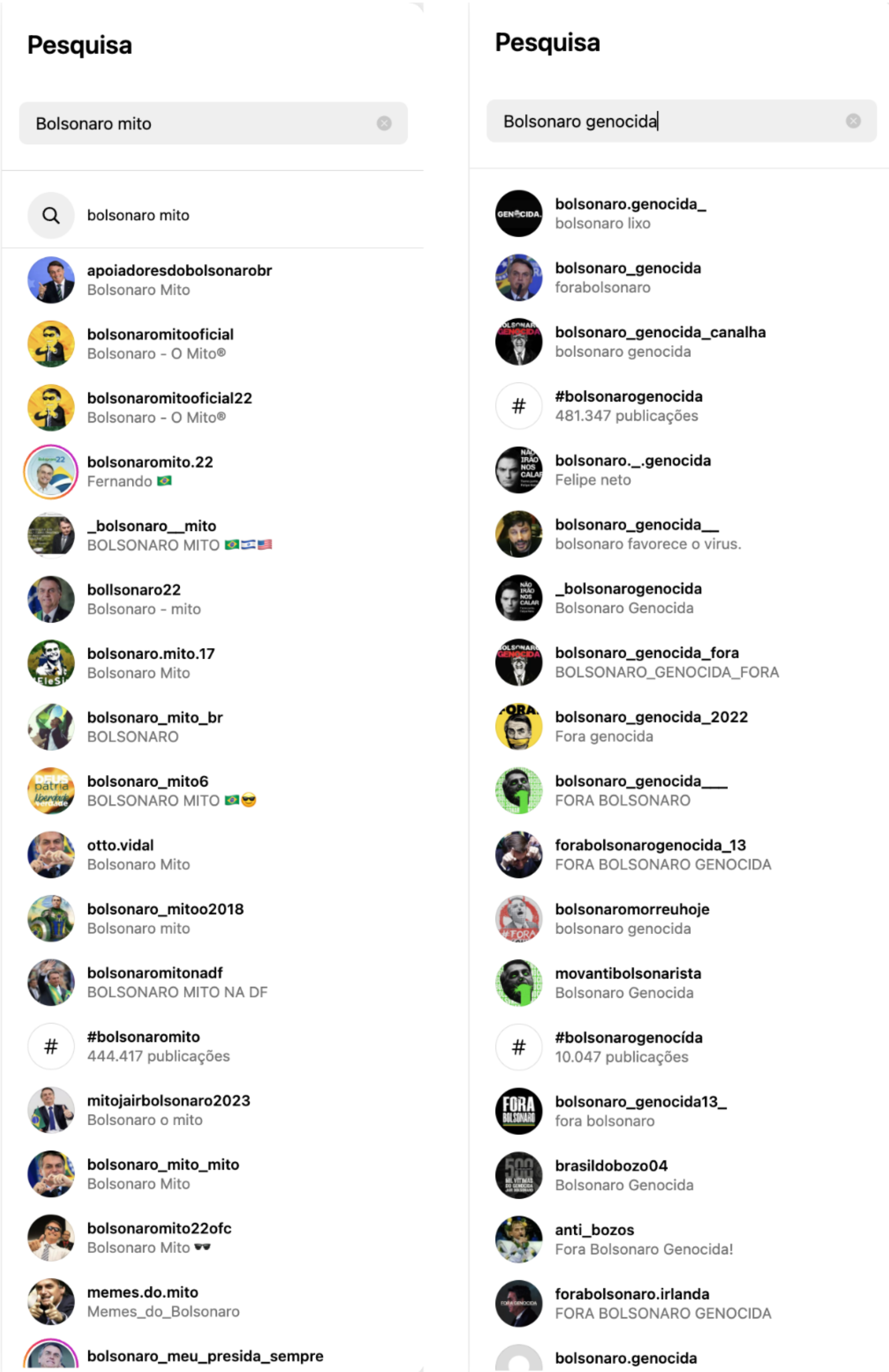

Figure 5. Example of contextual keywords in the project Profiling Bolsobot Networks (screen capture: 26 April 2023)

In the context of the reiterative process of interpretation, digital visual research offers possibilities beyond what is visible in data visualisation interface software. In the Profiling Bolsobot Networks project, image metadata was consulted, i.e. using spreadsheet software and the image folder. Also, the Instagram interface was visited, and a Google Slides shared document was created for recording collective annotation and screen captures. Image-based analysis encompassed dominant image clusters grouped by colour similarity, image repetition, the presence or absence of profile avatars and suspicious human clusters. Figure 6 illustrates each of these analytical units showcasing the potential of software-oriented research and digital methods. Colour similarity revealed political and ideological representations that are either shared or reappropriated by pro- and anti- Bolsonaro supporters, e.g. the Brazilian flag, turn right or left symbols and the use of Anti-Facist logos. In comparison, the avatar[8] and human clusters[9] point to unobtrusive bots and fake accounts attempting to pass as individuals. The latter are known to be a costly service in the bot market, precisely because the accounts resemble ordinary people. The former, the so-called ghost accounts, in comparison, refers to an unsophisticated typology of bots that operate in unobtrusive modes, such as turning on the private feature on Instagram and playing a role as central and bridging actors in bot following networks (see Omena et al. 2024).

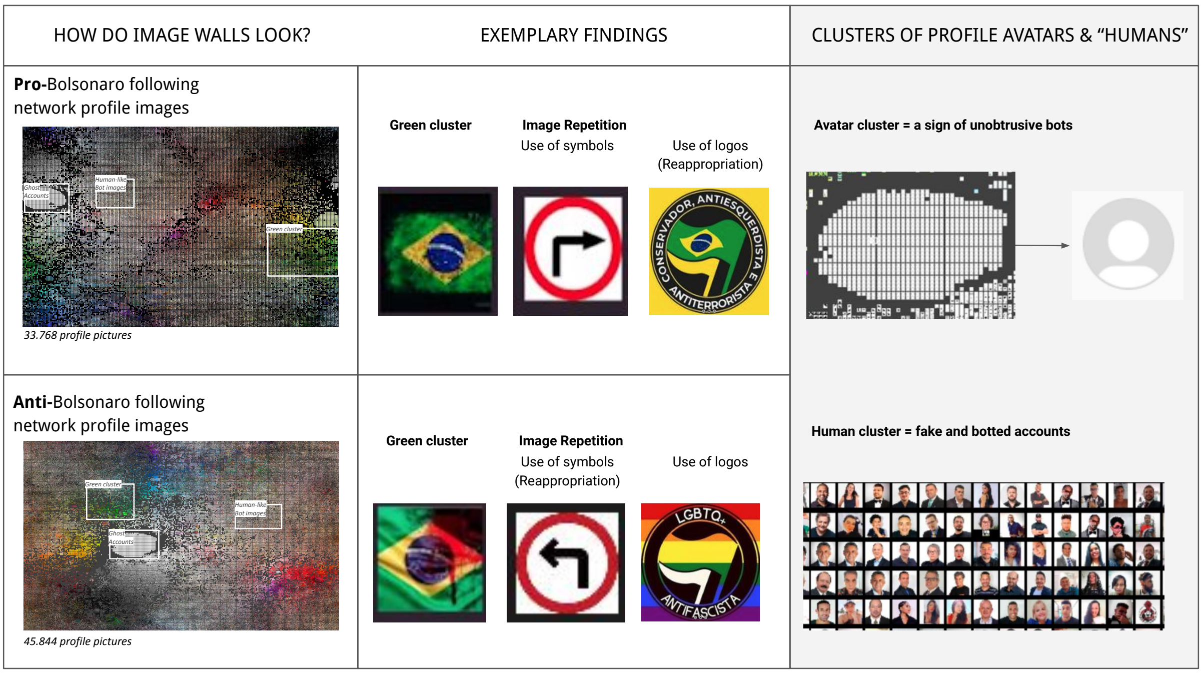

Figure 6. Profiling Bolsobot Networks project and image collection-based analysis

The Profiling Bolsobot Networks project offered a descriptive insight into the process of rich data curation, iterative interpretation, and multi-level inquiries. These practices are guided by a comprehensive understanding of the topic under investigation (i.e. antagonistic political debates in Brazil) and, equally importantly, demonstrate how close engagement with software opens up new avenues and modes of thinking with image datasets through quali-quanti visual methods.

[8] It is a default profile image used in Instagram for profiles that have not uploaded a profile picture.
[9] Headshot pictures using the same white background, probably coming from stock image webpages, AI generative tools or more craft montages.

https://doi.org/10.33621/jdsr.v6i2.214

5. Principles for extra-hard visual research

Building upon methodological reflections from the case studies, we present critical considerations and recommendations for digital social research, proposing a manifesto for extra-hard data and quali-quanti visual methods (QQVM). This manifesto serves as an invitation and explanation for designing and implementing digital methods research in a manner that goes beyond treating visual data as an isolated entity. Instead, it emphasises the importance of considering the environment from which the data originates and, aligning with Omena's methods theory (see Omena et al. 2024), the additional layers of meaning that arise through software extraction, analysis, and visualisation, resulting in what we call extra-hard visual data.

Data curation and collection aligned with and responsive to the online space being investigated. Extra-hard visual data suggests the avoidance of disconnected practices, such as selecting keywords without thorough navigations of the platform environment that mediates the case. This includes asking empirical questions about the use of hashtags, keywords, usernames, or other digital objects for data scraping, crawling, or API calls. Researchers may need to dedicate time to observing, searching, and engaging with the online space before the completion of the full research design. This includes understanding the prevalent usage culture and unique technological grammar of the platform. Additionally, technical knowledge of the extraction software is necessary, as crucial decisions regarding data collection parameters will be imprinted in the scraped data.

Considerations of the diverse platform (sub)spaces, user dynamics, and cultures informing digital datasets (in- or post- research design). While diverse platform affordances and user dynamics may suggest that platform comparisons on the basis of platform-specifically designed research do not necessarily produce authentically comparative data, rich post-design comparisons may allow for a review of how platform spaces and diverse national contexts have impacted the data and results obtained. As such, extra-hard data may be achieved through individualised research designs and subsequent cross-project analysis and triangulation. This also allows for a partial mitigation of time and funding restrictions that are particularly out of reach for early-career and otherwise underfunded or unsupported researchers.

Exploratory, iterative, and multi-level analysis of visualisation software. Visualising networks (e.g., using Gephi) or image collections (e.g., with ImageSorter) requires the recognition that visualisation software is not an end in itself but serves as a starting point to explore the relational nature of online data. Conducting multi-level inquiries requires both a solid technical understanding of the chosen tools and an awareness of the steps and decisions that contribute to the final visualisation. For example, examining images grouped by colours and identifying image repetition requires questioning how the datasets, platforms, and data collection decisions contribute to this pattern.

Contextualising statistical or metadata. Ideally, mined platform data will be interrogated in light of the wider context of a given protest case towards triangulating the activity of digital social movements. This may include on-the-ground engagement with protesters through interviews or other methods that engage directly with users and provide depth and context to statistical data obtained remotely.

6. Discussion & conclusion

6.1 Easy, hard, & extra-hard data

This paper has presented evidence from the field and current research on visual politics and protest to show the presence of what we have consequently termed *extra-hard data*. We demonstrated that extra-hard approaches are a possibility afforded by new digital and internet research methods, as well as a prerequisite for reading visual data. We argued that easy data access has distracted scholars from the potentialities of the computational turn through which digital methods have provided avenues for data contextualisation and situatedness towards more nuanced, robust, rich, comparative, and versatile visual research that better contextualises politics and protest. The notion (and even presence) of extra-hard data suggests that the collection and interpretation of easy or hard data are matters of choice and opportunity, but not in itself an impossibility of any specific method including the use of computational tools.

This paper additionally suggests that research software and data alone do not conduct or establish research; they rely on researchers to unpack their potential. By asserting that "extra-hard" data already exists, we argue that limiting such data to traditional humanities-based textual analysis constrains its broader possibilities. For instance, while topic modelling can provide a surface-level analysis, it cannot replace the profound hermeneutical work required for in-depth understanding. Similarly, software alone cannot establish meaningful research in and by itself. This understanding stems from practical and technical considerations, including becoming familiar with research software, their methods and data visualisation, the data's origin and the context in which it is collected, and the empirical interpretation of the resulting outputs from this process.

6.2 Binary perspectives reconsidered

All this is not to say that the easy-hard discussion is a hard binary or, with extra-hard data, a ternary. If anything, the methodological potentialities presented here provide room to question such distinctions. In using the term extra-hard data, we suggest that there is room for developing data practices in digital social research that move away from methodological habits that regard single aspects of common methodological binaries (e.g. easy and hard, qualitative and quantitative, or visual and textual) and speak to the relational nature of data and computational media in their own terms (particularly the case with visual data). In many ways, this is a matter of bridging research mindsets that strictly apply methodological binaries. Binary perspectives do not apply in the very same way in digital social research, a point illustrated in the quali-quanti approach. Nevertheless, qualitative internet research and digital methods are often implied to be at odds. While qualitative research tends to be connotated as small-scale but rich and contextualised, software-based approaches carry the connotation of being descriptive and decontextualised. Even so, extant research (including the cases presented here) suggests that these complement each other, for example in that software-based approaches situate qualitative data and the latter provides texture to the former.

There is, in that regard, a need to rethink how the scientific community views software-based methods, and to critically interrogate visual methods and its tools as an isolated technical process detached from qualitative inquiry. If anything, it may be the binaries, singular choices, and resulting hierarchies that create a preponderance of certain types of research and threats of digital bias in the first place. The notion of easy data may potentially even lead to disengaged attitudes towards computational mediums that are required to design and implement digital social research and the data collected through these due to the implied flat, reductionist, or instrumental data uses, i.e. a disregard of computational mediums as carriers of meaning and the technical-relational nature of online data. These are, in part, of course, tied to the choice and relevance of available tools, platform governance, access restrictions, funding, and other factors that may produce easy data approaches, but they also imply at times fuzzy understandings and applications of methodological processes as well as misunderstood or overlooked efforts in software-based research. In light of these considerations, this paper argues in favour of methodological approaches that break these conventions.

Beyond these provocations, it is our hope that future research will increasingly employ hybridity, triangulation, intersectionality, rich cross-project comparisons, and efforts at contextualisation. Even so, we understand that such efforts are underpinned by underlying inequalities and power differentials. Extra-hard data often remains the domain of privileged individuals, institutions, and regions, above all in relation to access to data, networks, skills, funding, security, and other forms of privilege. In relation to digital methods specifically, these differences in equity become particularly significant due to the field's specific requirements for software skills, costly training workshops, and the related necessity for funding. This complexity is exacerbated by issues around tool transparency, platforms' 'research gatekeeping', invisible social dynamics, changing platform features, platform governance, and issues in tool usage conventions. This may well mean that 1) academic institutions and platform providers will need to offer the conditions that allow extra-hard data to become a feasible option (through training, funding, and access), and that 2) scholars conducting research with extra-hard data need to expand their practices of making digital methods worksheets, methods recipes, and coding notebooks publicly available towards democratising access to such methods.

Acknowledgments

S.M.Ö. is supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 101027274. We collectively wish to thank the ECREA Visual Cultures Section for providing the pre-conference on "Visual Politics & Protest" as a platform for collaboration. J.J.O would like to express their gratitude to the Public Data Lab collaborators of the Profiling Bolsobot Networks project, which received financial support from the Center for Advanced Internet Studies (CAIS). The authors additionally thank the collaborators of the case study projects: Patricia Prieto-Blanco, Xuanxuan Tan, and Norita Mdege for VisFem; Thais Lobo, Giulia Tucci, Mengying Li, Maria Luiza Paschoal, Marius Liedtke and Ilya Lavrov for profiling Bolsobot visual cultures; Ololade Faniyi, Emily Edwards, Sarah Ford, Debipreeta Rahut and Taufiq Al Makmun who are current and past members of her research lab titled "Tooling Around: Situated Data Analytics, Digital Humanities, Digital Archiving and Data Feminism" and involved in the major methods work for the two publications referred to.

References

Akrich, M., & Latour, B. (1992). A summary of a convenient vocabulary for the semiotics of human and nonhuman assemblies. In W. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp.259–264). Cambridge: MIT Press.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. Third International AAAI Conference on Weblogs and Social Media, 361–362. https://doi.org/10.1136/qshc.2004.010033

Baun, P. S., Schlüter, M., et al. (2022). Let's Play War: Inside 4chan's intergroup rivalry, contingent community formation, and fandomized war reporting. Retrieved from https://wiki.digitalmethods.net/Dmi/SummerSchool2022chugginguhg

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action. Digital media and the personalisation of contentious politics. *Information, Communication & Society*, 15(5), 739–768. https://doi.org/10.1080/1369118X.2012.670661

Berry, D. (2011). The computational turn: Thinking about the digital humanities. *Culture Machine*, 12, 1–22. https://culturemachine.net/wp-content/uploads/2019/01/10-Computational-Turn-440-893-1-PB.pdf

Burgess, J., Angus, D., Carah, N., Andrejevic, M., Hawker, K., Lewis, K., … Li, L. (2021, November 6). Critical simulation as hybrid digital method for exploring the data operations and vernacular cultures of visual social media platforms. https://doi.org/10.31235/osf.io/2cwsu

Burgess, J., & Bruns, A. (2015). "Easy data, hard data: The politics and pragmatics of Twitter research after the computational turn". In G. Langlois, J. Redden, & G. Elmer (Eds.), *Compromised data: From social media to Big data* (pp. 93–111). Bloomsbury Publishing.

Burgess, J., & Matamoros-Fernández, A. (2016). Mapping sociocultural controversies across digital media platforms: one week of #gamergate on Twitter, YouTube, and Tumblr. *Communication Research and Practice*, 2(1), 79-96. https://doi.org/10.1080/22041451.2016.1155338

Burgos-Thorsen, S. & Munk, A. K. (2023). Opening alternative data imaginaries in urban studies: Unfolding COVID place attachments through Instagram photos and computational visual methods. *Cities*, 141, 1-21. https://doi.org/10.1016/j.cities.2023.10447

Chao, J. (2021). Memespector-GUI: Graphical User Interface Client for Computer Vision APIs (Version 0.2.5) [Computer software]. https://doi.org/10.5281/zenodo.7704877

Colombo, G., Bounegru, L., & Gray, J. (2023). Visual Models for Social Media Image Analysis: Groupings, Engagement, Trends, and Rankings. *International Journal Of Communication*, 17, 28. https://ijoc.org/index.php/ijoc/article/view/18971

Colombo, G. (2019). Studying Digital Images in Groups: The Folder of Images. In L. Rampino & I. Mariani (Eds.), Advancements in Design Research: 11 PhD Theses on Design as We Do in Polimi (pp. 185–195). Franco Angeli Open Access.

D'Andrea, C., & Mintz, A. (2019). Studying the live cross-platform circulation of images with computer vision API: An experiment based on a sports media event. *International Journal of Communication*, 13. https://ijoc.org/index.php/ijoc/article/view/10423

de Zeeuw, D., & Tuters, M. (2020). The Internet is Serious Business: On the Deep Vernacular Web and its Discontents. *Cultural Politics*, 16(2), 214-32. https://doi.org/10.1215/17432197-8233406

Geboers, M., Van De Wiele, A., & Chad, T. (2020a). Regimes of visibility and the affective affordances of Twitter. *International Journal of Cultural Studies*, 23(5), 745–765. https://doi.org/10.1177/1367877920923676

Geboers, M., Stolero, N., Scuttari, A., Van Vliet, L., & Ridley, A. (2020b). Why Buttons Matter: Repurposing Facebook's Reactions for Analysis of the Social Visual. *International Journal of Communication*, 14(22). https://ijoc.org/index.php/ijoc/article/view/11657

Gibbs, M.; Meese, M., Arnold, M., Nansen, B., & Carter, M. (2015). #Funeral and Instagram: Death, Social Media, and Platform Vernacular. *Information Communication and Society*, 18(3), 255-268. https://doi.org/10.1080/1369118X.2014.987152

Gray J., Bounegru L., Milan S., Ciuccarelli P. (2016). Ways of seeing data: Toward a critical literacy for data visualizations as research objects and research devices. In Kubitschko S., Kaun A. (Eds.), *Innovative methods in media and communication research* (pp. 227–251). Palgrave Macmillan.

Hargittai, E. (2020). Potential biases in Big data: Omitted voices on social media. *Social Science Computer Review*, 38(1), 10–24. https://doi.org/10.1177/0894439318788322

Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., et al. (2020). Computational social science: Obstacles and opportunities. *Science*, 369,1060-1062. https://doi.org/10.1126/science.aaz8170

Karatzogianni, A. (2015). *Firebrand waves of digital activism 1994-2014: The rise and spread of hacktivism and cyberconflict*. Basingstoke: Springer.

Manovich, L. (2020). *Cultural analytics: Visualizing cultural patterns in the era of big data*. Cambridge, MA: MIT Press.

Manovich, L. (2007). Cultural Analytics: white paper (5/2007; latest update 11/2008). With contributions by Noah Wardrip-Fruin. http://lab.culturalanalytics.info/p/publications.html

Manovich, L., Gianchino, M., Chow, J. (2012) Image Montage v. 1.1. Available at: https://imagej.nih.gov/ij/plugins/image-montage/index.html

Marres, N. (2020). For a situational analytics: An interpretative methodology for the study of situations in computational settings. *Big Data & Society*, 7(2). https://doi.org/10.1177/2053951720949571

Marres, N., Colombo, G., Bounegru, L., Gray, J. W. Y., Gerlitz, C., & Tripp, J. (2023). Testing and Not Testing for Coronavirus on Twitter: Surfacing Testing Situations Across Scales With Interpretative Methods. *Social Media + Society*, 9(3). https://doi.org/10.1177/20563051231196538

Marres, N. (2017). *Digital sociology: The reinvention of social research*. Cambridge: Polity.

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., & Azzi, M. (2017). RAWGraphs: A Visualisation Platform to Create Open Outputs. In Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter (p. 28:1–28:5). New York, NY: ACM. https://doi.org/10.1145/3125571.3125585

Medialab Tools. Table2Net Available from https://medialab.github.io/table2net/

Nissenbaum, A., & Shifman, L. (2017). Internet Memes as Contested Cultural Capital: The Case of 4chan's /b/ Board. *New Media & Society*, 19(4), 483-501. https://doi.org/10.1177/1461444815609313

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. https://doi.org/10.7764/disena.14.40-67

Pearce, W., Özkula, S. M., Greene, A. K., Teeling, L., Bansard, J. S., Omena, J. J., & Rabello, E. T. (2020). Visual cross-platform analysis: Digital methods to research social media images. *Information, Communication & Society*, 23(2), 161-180. https://doi.org/10.1080/1369118X.2018.1486871

PhantomBuster [web application software]. https://phantombuster.com/

Philipps, A. (2012). Visual protest material as empirical data. *Visual Communication*, 11(1), 3–21. https://doi.org/10.1177/1470357211424675

Omena, J. J. (2021). Digital Methods and Technicity-of-the-Mediums. From Regimes of Functioning to Digital Research. Doctoral Thesis. Available at: https://run.unl.pt/handle/10362/127961

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

Omena, J. J., Lobo, T., Tucci, G., Bitencourt, E., de Keulenaar, E., Kerche, F., Chao, J., Liedtke, M., Li, M., Paschoal, M. L., & Lavrov, I. (2024). Quali-quanti visual methods and political bots: A cross-platform study of pro- & anti-bolsobots. *Journal of Digital Social Research*, 6(1), 50-73. https://doi.org/10.33621/jdsr.v6i1.215

Özkula, S. M. , Prieto-Blanco, P., Tan, X., & Mdege, N. (2024). Affordances and platformed visual misogyny: a call for feminist approaches in visual methods. *Feminist Media Studies*, 1-22. https://doi.org/10.1080/14680777.2024.2311355

Özkula, S. M., Reilly, P. J., & Hayes, J. (2022). Easy data, same old platforms? A systematic review of digital activism methodologies. *Information, Communication & Society*, 26(7), 1470–1489. https://doi.org/10.1080/1369118X.2021.2013918

Puschmann, C. (2019). An end to the wild west of social media research: a response to Axel Bruns. *Information, Communication & Society*, 22(11), 1582-1589. https://doi.org/10.1080/1369118X.2019.1646300

Rabello, E. T., Gommeh, E., Benedetti, A., Valerio-Ureña, G., & Metze, T. (2022). Mapping online visuals of shale gas controversy: a digital methods approach. *Information, Communication & Society*, 25 (15), 2264–2281. https://doi.org/10.1080/1369118X.2021.1934064

Rasband, W.S. (1997-2018). ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/

Rieder, B. (2020). *Engines of Order: a mechanology of algorithmic techniques*. Amsterdam University Press.

Rieder, B. and Röhle, T. (2012) Digital Methods: Five Challenges. In D. M. Berry (Ed.), *Understanding Digital Humanities* (London: Palgrave Macmillan), pp. 67–84.

Ricci, D., Colombo, G., Meunier, A., & Brilli, A. (2017). Designing Digital Methods to Monitor and Inform Urban Policy: The Case of Paris and its Urban Nature Initiative. International Conference on Public Policy (ICPP3). https://hal.archives-ouvertes.fr/hal-01903809

Rogers, R. (1996). The future of STS on the web. *EASST Review*, 15(2), 25–27. https://www.researchgate.net/publication/239841669_The_Future_of_Science_and_Technology_Studies_on_the_Web

Rogers, R.. (2019). *Doing digital methods*. London: Sage.

Rogers, R. (2018). Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication*, 12, 450–472. https://ijoc.org/index.php/ijoc/article/view/6407/2248

Rogers, R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*, 8(1). https://doi.org/10.1177/20539517211022370

Rogers, R. & Giorgi, G. (2023) What is a meme, technically speaking? *Information, Communication & Society*, 27(1), 73–91. https://doi.org/10.1080/1369118X.2023.2174790

Rogers, R., & Lewthwaite, S. (2019). Teaching Digital Methods: Interview with Richard Rogers. Interviewer: S. Lewthwaite. *Diseña*, 14, 12–37. https://doi.org/10.7764/disena.14.12-37

Rose, G. (2016). *Visual Methodologies* (4th ed.). SAGE Publications.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W (2012). "NIH Image to ImageJ: 25 years of image analysis". *Nature Methods*, 9, 671-675.

Software Studies Initiative (2011). ImagePlot software v. 1.1. http://lab.softwarestudies.com/p/imageplot.html

Selfiecity. (2023). Retrieved 16 May 2023. https://selfiecity.net/#

Stiegler, B. (2012). Die Aufklärung in the Age of Philosophical Engineering. Computational Culture. https://doi.org/10.1017/CBO9781107415324.004

Timeline: 4535 Time Magazine Covers, 1. (2016). Timeline: 4535 Time Magazine Covers, 1923-2009. Retrieved 16 May 2023, from http://lab.culturalanalytics.info/2016/04/timeline-4535-time-magazine-covers-1923.html

Tromble, R. (2021). Where have all the data gone? A critical reflection on academic digital research in the post-API Age. *Social Media + Society*, 7(1), 1–8. https://doi.org/10.1177/2056305121988929

Venturini, T. Bounegru, T., Gray, J., & Rogers, R. (2018). A reality check (list) for digital methods. *New Media & Society*, 20(11), 4195-4217. https://doi.org/10.1177/1461444818769236

Venturini, T. (2024). Quali-quantitave methods. In A. Irwin & U. Felt (Eds.), *Encyclopedia of Science and Technology Studies* (forthcoming).

Venturini, T., Jacomy, M. and 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). https://doi.org/10.1177/20539517211018488

Visual Computing (2018) ImageSorter [Software]. HTW Berlin - University of Applied Sciences. https://visual-computing.com/project/imagesorter/

Xue, C. (2014). Image Preview Plugin. https://gephi.org/plugins/#/plugin/image-preview

