

Visual media analysis for Instagram and other online platforms

Big Data & Society
January–June: 1–23
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517211022370
journals.sagepub.com/home/bds



Richard Rogers

Abstract

Instagram is currently the social media platform most associated with online images (and their analysis), but images from other platforms also can be collected and grouped, arrayed by similarity, stacked, matched, stained, labelled, depicted as network, placed side by side and otherwise analytically displayed. In the following, the initial focus is on Instagram, together with certain schools of thought such as Instagramism and Instagrammatics for its aesthetic and visual cultural study. Building on those two approaches, it subsequently focuses on other web and social media platforms, such as Google Image Search, Twitter, Facebook and 4chan. It provides demonstrations of how querying techniques create online image collections, and how these sets are analytically grouped through arrangements collectively referred to as metapictures.

Keywords

Social media, visual media, visualisation, metapicture, digital methods, image analysis

Introduction: Analysing ‘visual media’ online

As indicated by the title, the analysis is of ‘visual media’ online, a term that could be more specific such as ‘visual social media’ (Leaver et al., 2020), ‘digital visual media’ (Dean, 2019), ‘digital visual artefacts’ (Leszczynski, 2018) or even ‘digital images, digitally analysed’ (Rose, 2016). An even more straightforward designation could be ‘social media images’ (Pearce et al., 2020). Each narrows the purview in a distinctive manner. The first concerns social media platforms driven by images (as Instagram, Tumblr, 4chan and perhaps even Flickr), the second emphasises image formats as memes, the third the inclusion of visual objects as emojis and gifs and the fourth digital images, which covers the digitised together with the internet-native. Social media images, the last category, would widen the platforms under study beyond ‘visual social media’ to any with images. In the following, the online visual media discussed includes all the above categories or emphases and is restricted to images online, rather than video, though YouTube’s thumbnails are addressed.

But the point of departure is not just the objects of study and their key terminological specificities. It is

also the outputs of their study, i.e. their analytical and interpretive arrangement. Akin to the manner in which digital research methods often fuse the tool and the method, in what has been referred to as ‘programmed method’ (Borra and Rieder, 2014), here analysis and arrangement are melded, the outcome of which could be referred to as a metapicture, in Mitchell’s terminology. The collections of images are framed in a manner of display that enables critical reflection on them (Gronstad and Vagnes, 2006; Mitchell, 1994).

This metapicturing, then, seeks to nestle itself between qualitative visual analysis and interpretation (from, say, art history) (Ferne, 1995) and quantitative knowledge visualisation where one deftly chooses the visualisation type to fit the data set (Börner and Polley, 2014). That is, rather than an emphasis on individual images, the metapicturing applies the analytical reading

Media Studies, University of Amsterdam, Amsterdam, Netherlands

Corresponding author:

Richard Rogers, University of Amsterdam, Turfdraagsterpad 9,
Amsterdam 1012 XT, Netherlands.
Email: R.A.Rogers@uva.nl



to images *en group* (Colombo, 2018). It also retains the images in the metapicture rather than rendering them as data points and visualising them as abstractions.

In what follows are approaches and techniques for deriving and arranging groups of images as metapictures that reflect upon them. The grouping arrangements (with built-in critical reflection) are further specified as image reuse, image trends, image vernaculars, dominant image, image presence, image quality, image staining (or tarnishing), image circulation, image engagement, image associations, image removal and feed competition. Each is taken in turn, together with how they meet the needs of critical questions concerning online reputation, style spaces, societal discourse, cultural conflict, versions of events, content moderation, polarisation and alignment as well as others.

Finally, by way of conclusion, there is also an emphasis on making the metapicture active in the research undertaking. By making it active, I mean the consideration of software outputs and visual arrangements as starting points rather than culminations of investigations (though one may arrive at a metapicture strategy anytime throughout a project). Envisaging the output or metapicture with the built-in critical reflection thereby becomes a part of project formulation. But it also takes it a step further by inviting the formulation of questions that include the reflective arrangement of image groups, a theme to which I return in the conclusion.

Studying Instagram

Instagram launched in 2010 with the retro Polaroid icon, square picture format and well-named filters such as Lark (good for nature shots). As an object of study it has evolved over the years, with its association with selfies (Walker Rettberg, 2014), a staging aesthetic (Manovich, 2016) as well as a new kind of traveller outside of the tourist and traveller, the digital nomad (Bozzi, 2020). It is also connected with micro-celebrities and lifestyle influencers (Senft, 2013), who engage in ‘visibility labour’ (Abidin, 2016) through regularly posting, at once potentially earning themselves revenue as well as engaging in what one scholar has called the ‘war of eyeballs’ (Abidin, 2014).

These internet celebrities (large and small) build followings, or fan bases (Marwick and Boyd, 2011) through active engagement on the platform by following others and liking and commenting on their posts. Some of this ‘relational labour’ (Baym, 2015) may be automated; among the tools deployed by heavy users to build relationships in the form of followings are software apps that automatically like posts of others who follow them, in a form of interactivity once referred to as ‘interpassivity’, where the machine acts on your

behalf (Zizek, 1998). It also outsources a type of ‘affective labour’ (Hardt and Negri, 2000), the pleasant-seeming, outward display of a job one would rather not do.

The authenticity of the follower and like counts is an issue for especially Instagram (and to a lesser extent other platforms), given how follower factories (a term in the same family as click farm) offer ‘fame for sale’ (Cresci et al., 2016), where one can purchase followers, likes as well as comments (Lindquist, 2018). These can be of varying quality (and duration), given how some are detected by the platform itself and purged, resulting in infamous cases of sudden, telling drops in the follower counts of celebrities, athletes and politicians but also so-called pseudo-influencers (Castro, 2021; Confessore et al., 2018) (Figure 1). In keeping with the question of whether (and when) one can trust social media data and engagement analysis (Paquet-Clouston et al., 2017), such findings have been discussed in terms of ‘manipulated reputation’ (Aggarwal, 2016) and the ‘engineered self’ (Van Den Hoogen, 2019).

The detection of ‘fake followers’ is itself an academic (and marketing industry) undertaking, with online tools employing ‘signals’ (e.g. lack of profile picture;

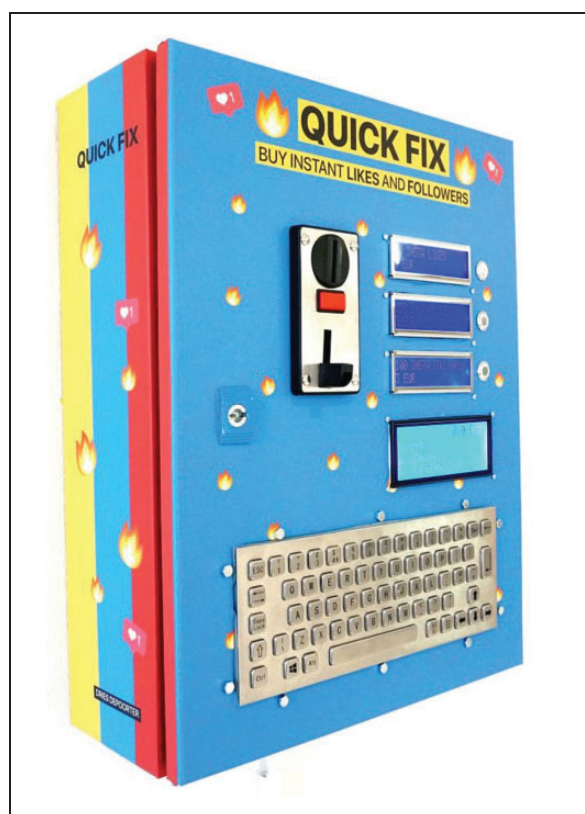


Figure 1. ‘Quick fix: Machine selling likes and followers’, interactive installation by Dries Depoorter. Source: Depoorter (2019).

multiple followers with same creation date) and ratios (low like-to-follower ratio) to aid in the hunt (HypeAuditor, 2021), along with more elaborate techniques in the computational literature (Sen et al., 2018) (see Figure 2). It is often reminded that there is a percentage of inauthentic accounts that one has in one's follower account anyhow, especially influencers', who are 'bombarded with fake followers' (Purba et al., 2020: 2763) and also are used (as well as recommended) as seed accounts for new users building a nascent profile.

Most recently, in the study of the platform, Instagram, like other social media, has been found to have a misinformation problem, first in its artful usage by Russian disinformation operatives and their unknowing co-conspirators during and especially after the 2016 U.S. presidential election campaigning (Howard et al., 2019), and also by purveyors of dubious content about COVID-19, vaccines and other contemporary topics (Cinelli et al., 2020; Colombo and De Gaetano, 2020; Massey et al., 2020). Whilst more attention in this regard has been paid to Facebook and Twitter, Instagram also has renewed promises of vigilance in their content review and moderation, an issue that came before them some years earlier with images associated with pro-eating disorder as well as other content removal controversies (Chancellor et al., 2016; Gerrard, 2020). Faced with content moderation, adherents in these communities would circumvent platform curbing through complex lexical and other

manoeuvring (e.g. a hashtag switch from #thighgap to #thyghgapp). Other work has focused on users whose accounts were unexpectedly or illegitimately suspended. They have described inflexible account policies and clunky review pathways (Myers West, 2018). If cut off by the platform, account reinstatement has proven arduous.

Instagramism and style spaces

As it is an image-driven platform, Instagram (but also 4chan, which is treated briefly below), is often associated with image analysis, particularly selfies; however, much scholars remind us that those pictures comprise a relatively miniscule percentage of the images uploaded to the platform (Caliandro and Graham, 2020). It is said to have its own aesthetic, or evolving aesthetics (Leaver et al., 2020), and as such can be approached as a 'style space' (Manovich, 2011a).

With the notion, Instagramism, Lev Manovich (2016) seeks to capture the movement to create and record beautiful forms and appearances in a way that enables building (and maintaining) a following, evidenced by overall follower count and likes per post. He describes Instagram styling as design photography. It could be similar in composition to that in *Kinfolk*, the international 'slow lifestyle' magazine (Manovich, 2017) (see Figure 3). Many Instagram users are said to expect staging. Indeed, the practice of sharing a photo after meticulously setting the scene contrasts with the

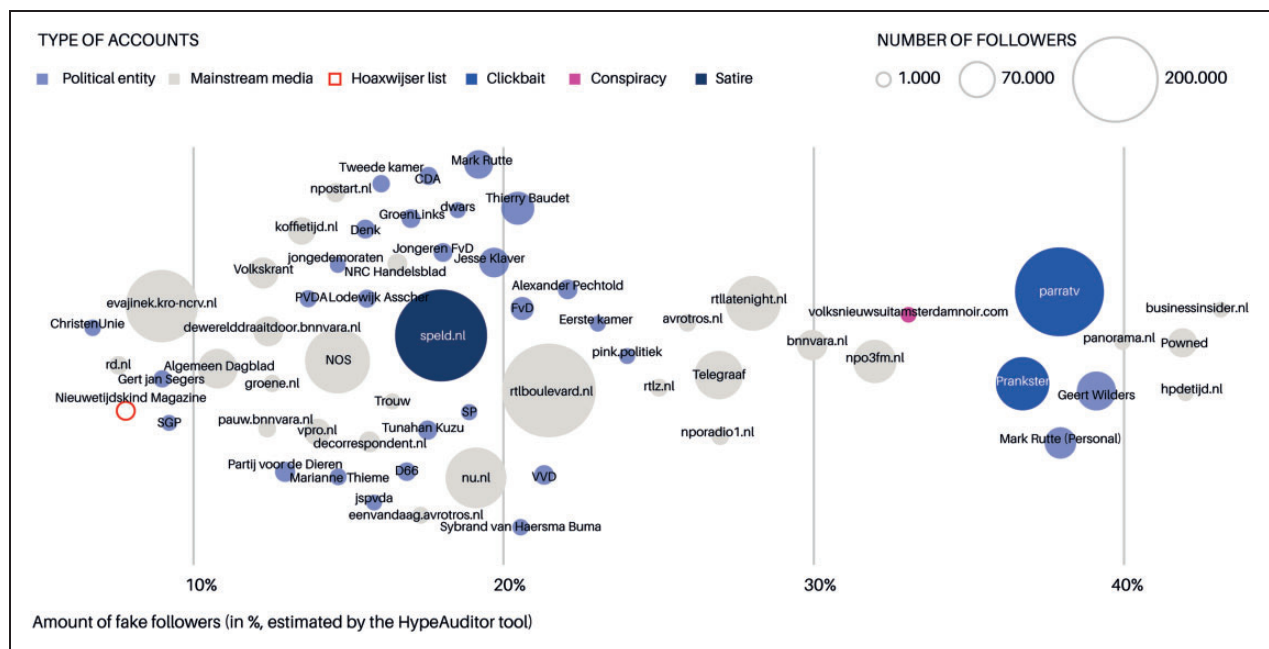


Figure 2. Comparison of percentages of fake followers on Instagram accounts held by political figures, media organisations and others, according to percentages of 'mass' and 'suspicious' followers, 2019. Source: Colombo and De Gaetano (2020).

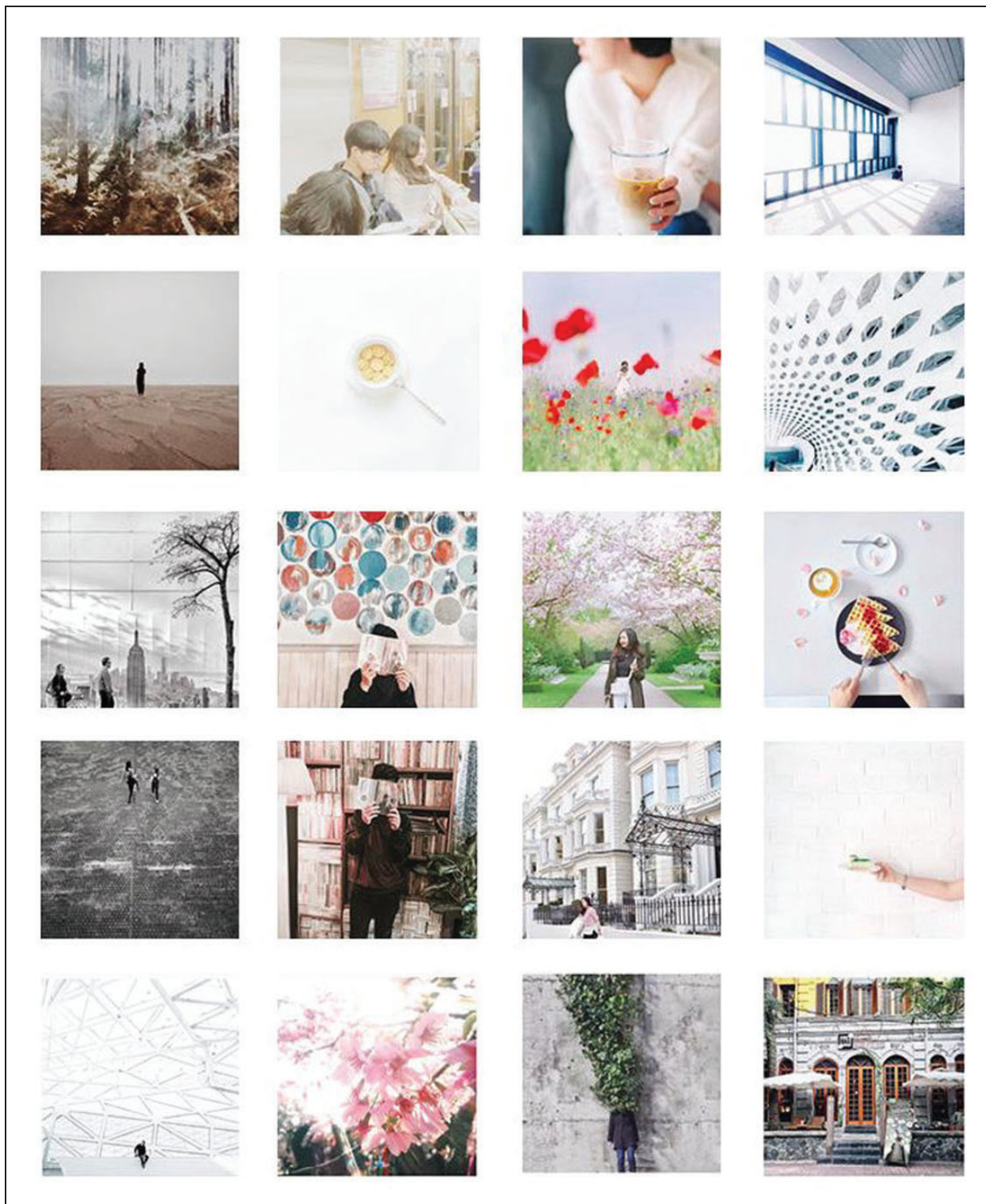


Figure 3. Instagramism. Design photography by Instagram users. Source: Manovich (2017).

pointing, shooting and sharing performed on other platforms, including messaging apps, where the aim is the maintenance of ‘ambient intimacy’ (Burgess and Baym, 2020).

Both selfie culture as well as Instagramism also could be said to be another instance of the internet hitting the streets, in at least two senses. The one concerns ‘place-making’ (Pink, 2008), that is, how shops,

restaurants, museums and other locales are Insta-styled. They may create a space where Instagram-type photos can be taken, but they also may create a mise-en-scène that could be termed Insta-photogenic (Budge, 2020a). Platform-place-making is a form of ‘gathering power’ (Casey, 1996) that privileges streets, buildings, bridges and squares by becoming Insta-destinations at once heavily hashtagged and geotagged (Boy and Uitermark, 2017). A second sense is what is sometimes termed participatory or ‘social photography’ (Budge, 2020b). Instagram-ready picture booths or styled settings are constructed for events, in order for platform users to have their pictures taken, eager to be remembered as a part of it (in front of the backdrop and associated with a hashtag). They also can snap their own selfies, individual or group ones, and subsequently post them with event hashtags, and other individuals tagged, to not just belong but also to be viewed and ‘liked’, thereby building platform capital.

Image similarity – Reuniting reused images

Studying style spaces could be performed by image grouping by similarity, or by using software that compares and groups similar images. As in other projects that employ it, such as network analyses, the grouping is a step in a more elaborate analytical procedure, rather than the endpoint. For example, ImagePlot (Software Studies Initiative, 2011) and ImageSorter (Visual Computing, 2018) group images according to their formal properties. Ingesting a collection (or folder) of images, the software orders them most often by similarity, particularly hue or saturation (see Figure 4). One use of such image grouping is to study the reuse of the same or similar images. The analysis rests on how the software bunches together similarity and sameness rather than sameness only. In other words, the cluster will amass both high-res as well as ‘poor images’ (Steyerl, 2009) of the same as well as



Figure 4. Image grouping. Originals, copies, original copies and poor images group together. Source: Geboers et al. (2020).

visually similar ones. The poor image is one that is of lower quality than the original or the copy, because it is the product of copy and paste, reworking for meme production, downsizing for a thumbnail or a preview, or otherwise processed downward, so to speak.

Another image type of interest in this regard is the ‘original copy’, a term for digital imagery that blurs the distinction between the original and the copy (Fehrmann et al., 2010). It is described as secondary images, or off-shoots, that attain the status of the original. These, too, would be grouped together because their properties compare well with the original and any copies, however poor. Thus, software could be said to reunite the original with its reuses, such as the copy, the original copy as well as the poor images.

Image trends and vernaculars

The software output of arrayed images may be the starting point of an analysis into what the image groups communicate per subject matter or issue, or how the medium may shape the form the message assumes. As a case in point, the images associated with #climatechange on Instagram or Twitter, for instance, can provide a sense of how the framing of the issue is evolving (in a form of social or environmental communication research) (see also Figure 5). With the entrance of youthful activists and influencers such as Greta Thunberg and Extinction Rebellion, are we seeing more urgency expressed in the imagery (burning planet)?

We also may examine how platforms depict an issue in a particular manner (Niederer, 2018; Pearce et al., 2019b). Overall, when sorted by popular posts tagged with #climatechange, Instagram may portray more individual ‘small actions’ (‘teens plant trees’), whereas Twitter may have more of the charts and figures of science, or the memes of politicians (French Premier Macron’s ‘Make our planet great again’). Depending on the platform used to access the issue, the question of who acts, and how it is visually displayed, is answered differently.

Studying how platforms, or communities within them, have a particular style of presentation and cultures ‘native’ to them is also known as platform vernacular research (Gibbs et al., 2015). In a particular issue space such as climate change, the one will prefer a certain format (the landscape), whereas the other routinely deploys another (the chart). To visualise such image languages or vernaculars, there is the image stack technique, which falls under the larger category of composite images (Colombo, 2018). In this visualisation practice, the order of the images is retained, and each image’s opacity is lowered equally (e.g. with 10 images, each is dimmed by 10%). The retention of the order of the layers grants more visibility to the higher ranked images (see Figure 6). Ranking here also can be by engagement score.

Platform vernacular research, it should be pointed out, ought not ignore the absence of a typical user of a social media platform (Gerlitz and Rieder, 2013), let alone one with over one billion registered users, as

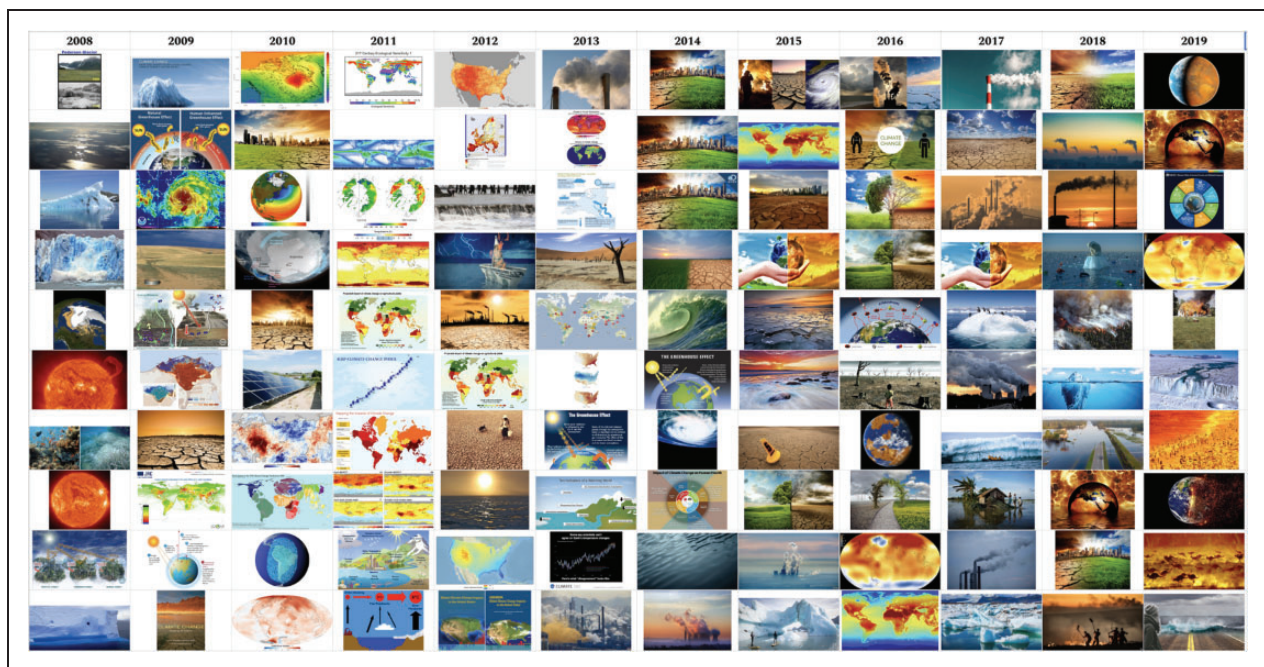


Figure 5. ‘Climate change’ query results from Google Image Search, overtime, 2008–2019. Source: Pearce et al. (2019a).

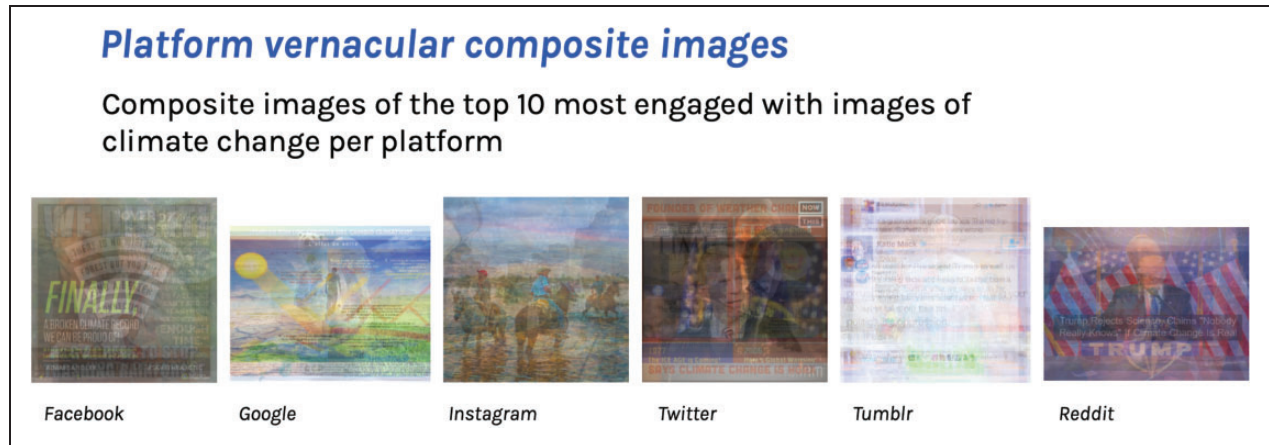


Figure 6. Image composites comparing top 10 most engaged-with climate change images across platforms, and technique. Source: Colombo (2018) and Pearce et al. (2019b).

Instagram. There are myriad uses of platform, be they documenting everyday life, styling like an influencer, campaigning or posting pictures of kittens (Caliandro and Graham, 2020); diversity of use lies within countries and cultures (Leaver et al., 2020). But, as discussed below, when one queries Instagram, demarcates a substantive space and compares it to the same space on another platform, meaningful distinctions for media and social research may emerge. For example, one platform may perform content moderation particularly well or poorly, in one language space versus another, when the topic explored is the relationship between 5 G and COVID-19, a subject of popular conspiracy theory (EU DisinfoLab, 2020). One may have more divisive content, others healthier dialogue (Niederer and Colombo, 2021).

The dominant image

Whether for social or media research, there are basic approaches to studying sets of images with software. When one orders images by formal property and outputs an array of them in the form of an image wall or cluster map (for example), initially one will note one or more dominant images (see Figure 5). The dominant image may be studied in terms of a format (such as meme) or as a message (slogan or counter-slogan). It should be noted that the technique also shows the opposite of the dominant. It yields marginal and orphan images. One may also place arrayed image collections from competing hashtags side by side – #blacklivesmatter and #bluelivesmatter or #protectthevote and #stealthevote – so as to enable comparative dominant and marginal image analysis.

There is an additional application, exemplified in work performed on the use of the climate-related hashtag, #parisagreement, before and after U.S. President

Trump's announcement in June 2017 that the U.S. would withdraw from it (see Figure 7). Here, the software allows one to compare the quantities of images, and thus activity, one month before and after in a form of 'direct visualisation' of all the images (compared to translating the images to data points and visualising as a data representation, as in Figure 2) (Manovich, 2011b; Niederer and Colombo, 2019). It allows for both a distant reading (Moretti, 2013) in the quantity comparison as well as a close reading as one can zoom in and take note of the dominance of 'Make Our Planet Great Again', the placard-like image (and cut-out hashtag) that launched the campaign by French President, Emmanuel Macron, on 1 June 2017, in reaction to the Trump announcement. With the U.S. abdication, it calls on France (and Europe) to lead the way. Here, close reading refers both to the method as well as the account one can make from it (Schur, 1998).

But the software also can order the images chronologically. If, in a cultural analytics approach (Manovich, 2020), one studies the digitised covers of a magazine over time, or the works of an artist, one can pinpoint instances of change in order to enable a fledgling periodisation (Manovich, 2016). A large collection of Instagram images, chronologically visualised, could point to larger style trends, such as the rise (and potentially fall) of a particular aesthetic associated with the platform. Is Kinfolk waning in favour of cottage core, the internet aesthetic associated with an idealised rural life? Or are particular formats becoming dominant (e.g. sudden memeification of an issue space)? In both examples of image analysis with software – grouping by formal property and ordering by date – the metapicture is the prompt for interpretation. The interpretative work is often practically undertaken through forms of annotation: labelling the clusters and the transition points and periods, respectively.

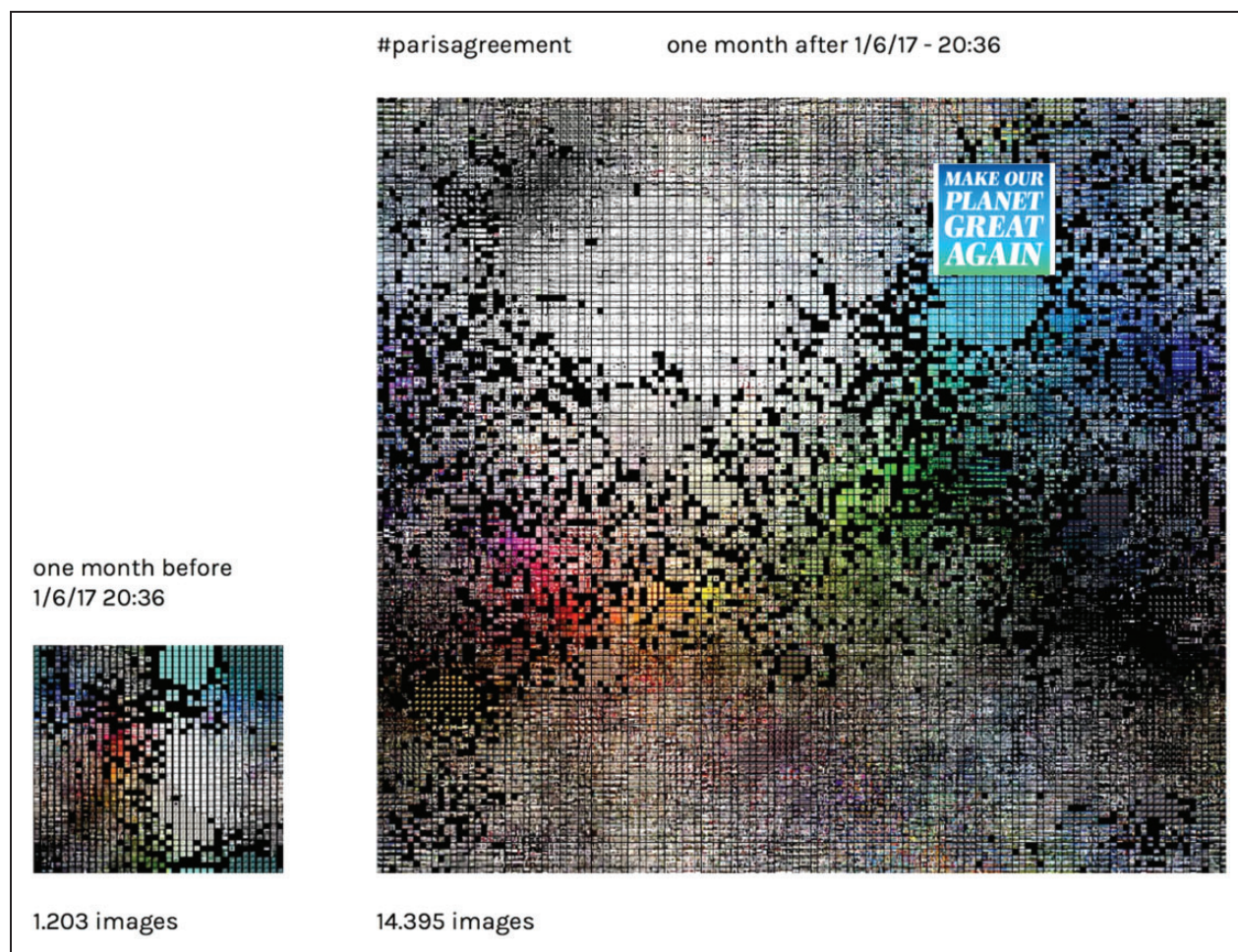


Figure 7. Dominant image analysis with ImageSorter. French President Macron’s meme, ‘Make our planet great again’ as dominant image on Instagram, both one month before and one month after U.S. President’s Trump announcement that the U.S. would withdraw from the Paris climate agreement. Source: Niederer and Colombo (2019).

Instagrammatics – Querying for image sets

Whereas Instagramism could be construed as media style and trend research, Instagrammatics is closer to internet as well as social research (Highfield and Leaver, 2016). It utilises queries of the Instagram Application Programming Interface (API) (when it was accessible), Instagram scrapers, CrowdTangle or other social listening software, making collections of posts for further interpretive study. Queries can be made of hashtags, user accounts and geo-coordinates, including combinations thereof. Hashtag-based queries include influential single hashtags (#blacklivesmatter, #metoo, #stealthevote), sets of related hashtags or competing hashtags (either single or sets of related ones such as #lovewins and #jesuswins – see Figure 8). One may create co-hashtag maps and label the clusters, showing the subcultures or discourses in a particular movement

or issue space, such as ‘settler superiority’ in Canada’s indigenous territories (Karsgaard and MacDonald, 2020). Here, Instagram becomes a platform to study the size of publics and counterpublics, together with their discursive imagery and hashtags.

The queries also can be user accounts, e.g. influencers. When queried together with hashtags, the research could concern the extent to which influencers use responsibly their symbolic power and voice, both generally and in a particular hashtag space (Niederer and Colombo, 2021). The queries could be geo-coordinates. The sourcing of a set of geotagged posts, together with a date range, may result in a collection of images of an event, such as the storming of the U.S. Capitol building. Geotagged posts from Washington, DC on 6 January 2021 (the day of the storming) would be a collection of significant historical interest. Making such a collection close to the event, and subsequently making another one with the same query design

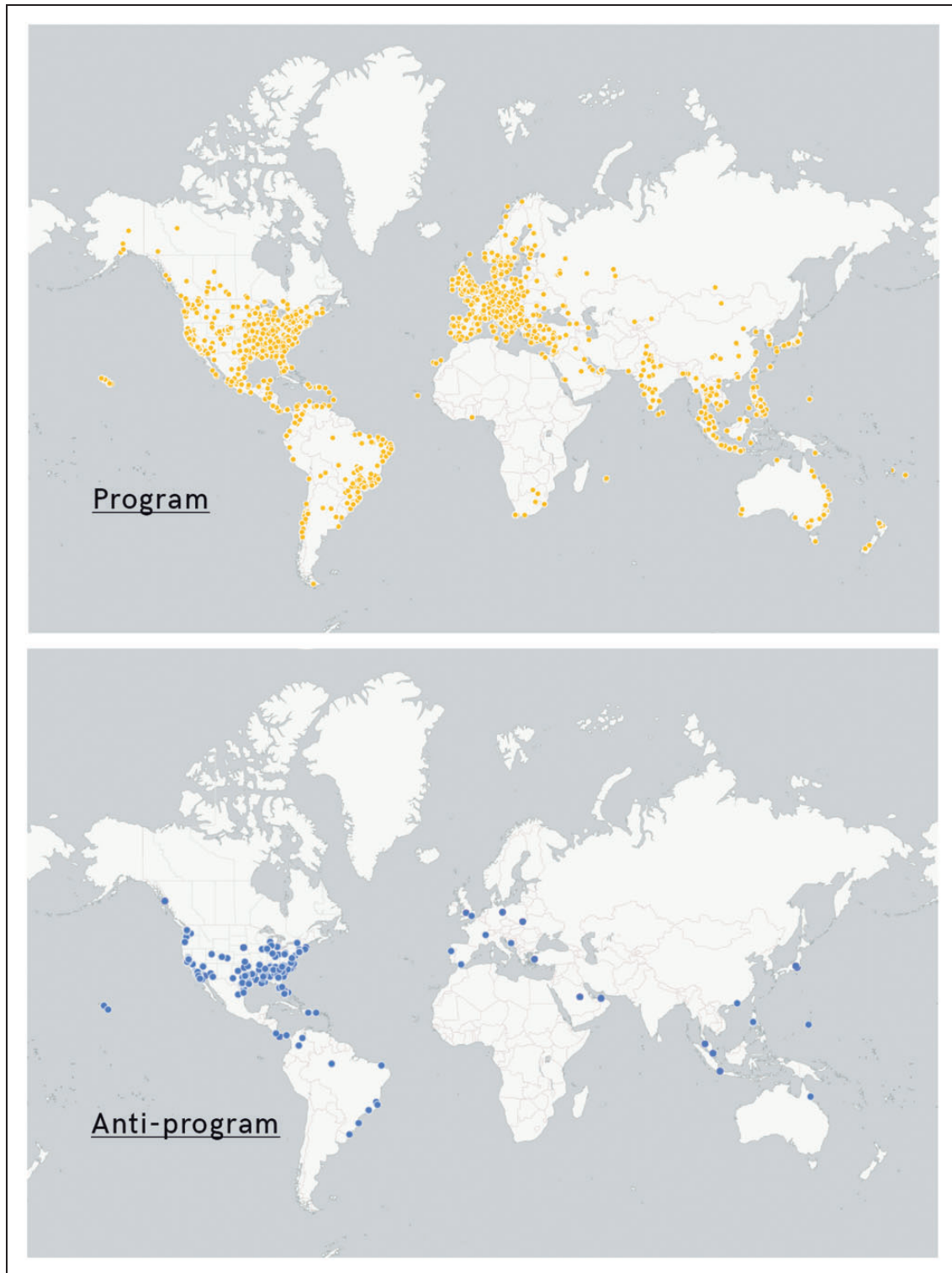


Figure 8. Geo-locating posts tagged #lovewins or #celebratepride (program) and #loveloses or #jesuswins (anti-program) after the U.S. Supreme Court ruling on same-sex marriage, 2015. Source: Rogers (2018).

sometime later, would enable the study of content moderation, the cleaning up of the platform and of evidence. Combining geo-coordinates with hashtags, such as in the well-known #selfiecity project, opens

further avenues of analysis, for it allows one to geo-locate particular moods and sentiments (Tifentale and Manovich, 2015). Where is #stealthevote most prominently located? Is it just as urban as rural?

Metapictures as visual media analysis

I would like to turn to a set of metapictures that, like Instagrammatics, result from visual media analysis with digital research methods (Rogers, 2019) but make image sets from other platforms, including Wikipedia, Google Image and Web Search, Twitter, Facebook as well as YouTube. One is a technique that creates an image grid that compares the presence or absence of images in articles about the same subject matter but written up in different languages on Wikipedia. How to interpret ‘missing’ images? Another examines the top images for a Google Image Search of Gezi Park (Istanbul) protests, extracting the embedded metadata in the pictures taken, in order to determine the price of the cameras that took them and the presence of citizen journalists in the so-called ‘pop-up news ecology’. A third employs a ‘staining’ technique that shades (and, in a sense, tarnishes) search engine results that are construed as preferred placement or misinformation. The others concern the posts (with images) that most animate a movement or group on Facebook, the emojis that are used in tandem with particular war images, the videos (expressed as thumbnail

images) referenced by 4chan users that have been removed by YouTube and the contrasting feeds of U. S.-based conservative and progressive Facebook users.

Image presence

In the study of cultural (and national) points of view on Wikipedia, Emina Sendijarevic and I deployed cross-cultural (or cross-lingual) image analysis to compare the ‘same’ articles in different Wikipedia language versions (2012). Certain images are dear to one article to make a cultural case, whilst conspicuously absent in another (see Figure 9). The absence may explain or provide context to emphasis, such as the image of a gravestone of a 13-year-old boy, prominent in the Bosnian article on the ‘Srebrenica genocide’, and for years absent in the Serbian one entitled ‘Srebrenica massacre’. (At the time of writing it is a part of a ‘picture gallery’ on the bottom of the Serbian article, rather than incorporated into the article itself.) The image analysis, outputting a grid, or rows of images from each article next to one another, shows either their order (and thus prominence) or their matches (and



Figure 9. Image grid analysis comparing images in the ‘same’ Wikipedia article on the Srebrenica massacre/genocide across different language Wikipedias. Image insert is of the grave of a 13-year-old boy, present in the Bosnian and not present in the Serbian article. Source: Rogers and Sendijarevic (2012) and Rogers (2013).

thus distinctiveness). It can include the templates, too, such as if an article has an issue with its ‘neutral point of view’ (in this case the Serbian) or has attained the status of ‘featured article’ (as has the Bosnian).

The cross-cultural image analysis may be paired with other analytical techniques such as comparing references from two or more articles or placing side by side (translated or original-language) tables of contents, where one notes for example that the Bosnian article on the Srebrenica genocide closes with a section on revisionism and the Serbian-language Srebrenica massacre piece with the controversy of calling the events a genocide. The cultural specificity of the differing accounts of the events continues decades later, seemingly hardening in stance rather than dovetailing. Stance solidification over time may also be depicted by highlighting the lack of change. For example, one accentuates (through shading) telling section headers or sentences that have long endured.

Image quality and camera grade recognition

In data journalism research, a technique in the open source intelligence tradition may be utilised for a form of ‘source criticism’, where one is interested in the origins of the source, in our case whether a picture was taken by a professional grade camera (and thereby presumably by a professional photojournalist). With EXIF data one can discern the camera type, date the pictures and examine the extent to which it has been edited or ‘touched up’. One can look up the prices of the cameras, too.

This form of visual media analysis may be applied in the examination of the apparent significance of citizen journalism and the ‘pop-up news ecology’ (Wall and el Zahed, 2015) at Gezi Park in Istanbul as demonstrators gathered in 2013 to protest the planned construction of a shopping mall in the urban green space (Ozduzen and McGarry, 2020). The environmental concerns evolved into multi-issue, political demonstrations, met with water cannons and other security apparatus, where many images were taken by the protesters themselves and subsequently circulated. It often has been argued that amateur or citizen documentation of events is significant in how they are reported and ultimately remembered (Robinson, 2009). They may challenge dominant and official accounts (Robinson and DeShano, 2011); they may supply alternative (evidentiary) material (Bruns and Highfield, 2012).

In an analysis of Google Image results for the query ‘Gezi Park’, we found that particularly iconic images (particularly the ‘lady in the red dress’ peppered sprayed by a policeman) remained at the top for well

over a month (see Figure 10). Thus, searching Google Images for events would not be a way to follow the goings-on, as they rather cement depictions rather than track them, at least at the time.

In the event, the iconic photograph of a lady in a red dress was taken by a professional photojournalist. Other top-ranked images were also shot by photojournalists or stringers.

Particularly poignant findings can be made on the basis of the metadata available in the images made iconic by Google in the sense of granting them their presence at the top of the results over time. Examining the data embedded in the top images, or EXIF data, we found that the photos with the most staying power were taken by professional grade cameras and edited by costly software, rather than from smartphones or cheap consumer devices without filters (Allan, 2015; De Amicis et al., 2013). We thereby relied on both ‘mechanically captured metadata’ (e.g. camera make, exposure) and that derived from user activity (software editing) (Hochman, 2014; Rubinstein and Sluis, 2008). Looking up the brands and models, we graphed the price tags of the cameras that took the top images, showing the continued significance of professional journalists and their documentation of the events rather than (mainly) upstart citizen journalists, at least in the event documentation returned by Google Image Search.

Image staining (or tarnishing)

A third example of visual media analysis is in the realm of search engine returns critique, otherwise known as algorithmic auditing (Sandvig et al., 2016). How to make an account to describe and expose the trouble with the top returns for an engine query? In the legal, commercial realm, accounts are often made of how search engines return their own properties (or those that bought space) at the top of the results, once known and scandalised as preferred or paid placement (Vaidhyanathan, 2012). Google notoriously returned its own e-commerce site over those of its competitors, landing itself in antitrust lawsuits (Romm, 2020; Scott, 2017). It is also accused of populating search engine returns with its own services, YouTube videos as well as information (or knowledge graph) panel, prior to serving other ‘organic’ results, as they are known in industry parlance. For example, search for any place and one is returned ‘Google places’. In the event, in 2019 Google reached a milestone in that more 50% of Google searches resulted in zero organic clicks (Fishkin, 2019).

Autocomplete (when the engine suggests a query) has been criticised for its ill-mannered suggestions (Baker and Potts, 2013). Autosuggested completions



Figure 10. Top images in Google Image Search for ‘Gezi Park’, 26 May–3 July 2013. Source: De Amicis et al. (2013).

for Jews, for example, prompted the engine company to intervene, removing offensive outputs (Cadwalladr, 2016). Google ads have been found to be racist, raising the question of how algorithms work, and whether (and how) their training material should be reconsidered (Sweeney, 2013). Google images have been found to have issues of representation when comparing the outputs of professional versus unprofessional hairstyles for work as well as pregnancy versus unwanted pregnancy (Bogers et al., 2020; Noble, 2018). In both cases, the research strategy of employing ‘counter-queries’ (e.g. ‘professional’/‘unprofessional’) serves to root out what is considered the ideal.

There are metapictures of preferred placement or rankings of one type of result compared to others that make use of annotating or ‘staining’ certain results and leaving others alone (see Figure 11). In one example, a journalistic piece emulates scrolling through engine results and by shading Google properties demonstrates that the first 35% (quite a long scroll) are Google’s own (Jeffries and Yin, 2020). (The article is entitled: ‘Google’s top search result? Surprise! It’s Google’.) The scrutiny of engine results by classifying them one way or another (through the lens of discrimination, misinformation, political leaning and other manners) can be visualised by colouring cells in a spreadsheet, as in misinformation is reddened, whilst other information is left blank (Torres and Rogers, 2020). One notes both the placement and amount at a glance. The ranking of a website for the same query over time may be graphed, for example in the Issue Dramaturg project that portrayed the ‘drama’ of search engine space as a website, routinely returned at the top of the results, one day vanishes from the first 1000 returns, likely because of an algorithmic ‘update’ (Rogers, 2013). Changes over time to rankings are also visualised with RankFlow, where one compares how a number of websites or videos in YouTube, for example, wax and wane in the search engine returns (Rieder et al., 2018).

Image circulation

Apart from critiquing their occasional, discriminatory labelling practices, lack of ethnic diversity and the scraped origins of the training sets behind them (Crawford and Paglen, 2019; Sindors, 2020), computer vision, a catch-all term for the automated recognition and classification of images, may be repurposed for critical social and media research, mainly into image resonance and circulation (d’Andrea and Mintz, 2019) but also as exploratory work into the imagery associated with an event or issue (see Figure 12).

A project studying the adoption of the lean-in photo collection by Getty images is one example (Aiello and

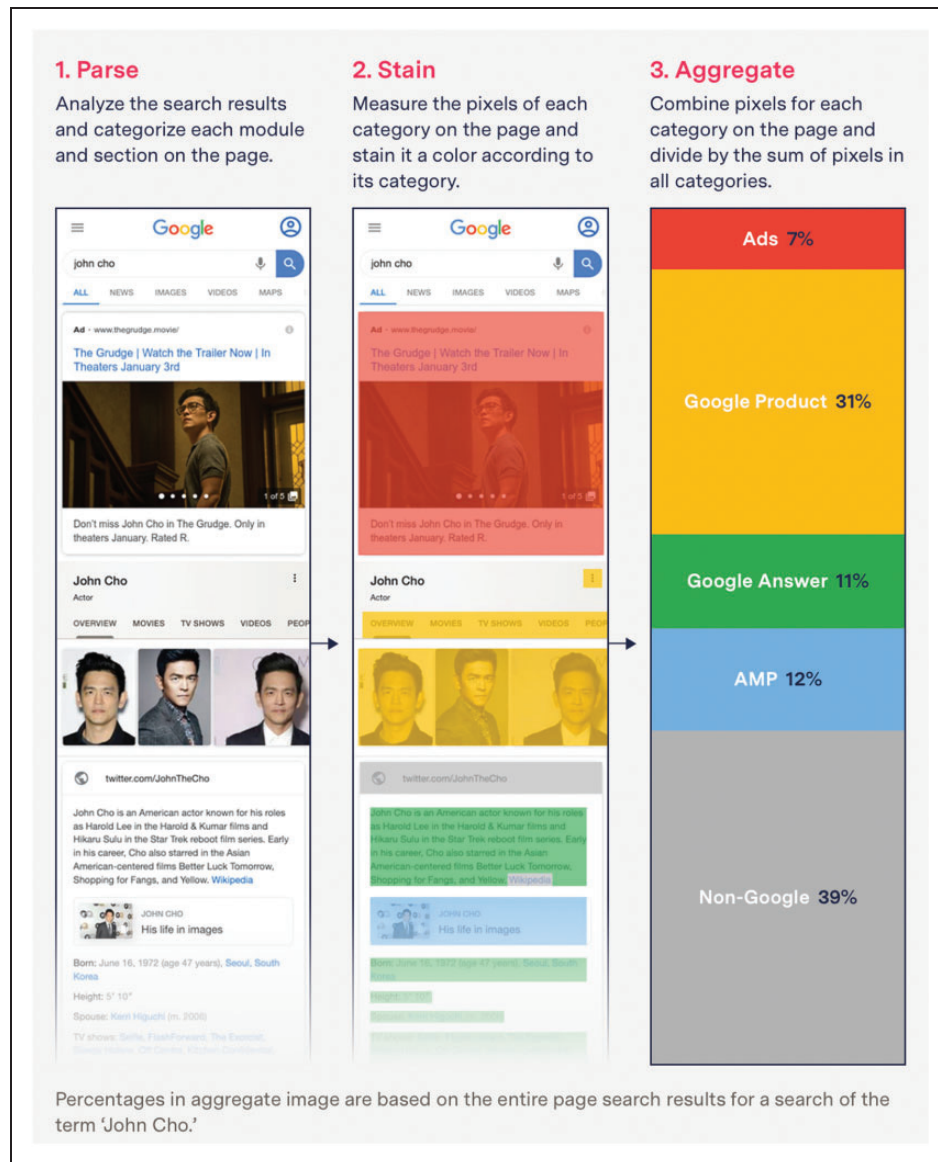


Figure 11. Image staining technique. Percentage of an entire page of Google results that contain Google properties. Source: Jeffries and Yin (2020).

Parry, 2020). The stock photo set has been heralded as ground-breaking, for it portrays women in powerful ways rather than as soft, sexualised and motherly (Miller, 2014). But how are these empowering stock images actually put to use in the media? By employing reverse image search (either via a computer vision API or in an image search engine), the research may locate instances of image usage, and enquire into whether they indeed break new ground. The images may be empowering, but their usage, we found, did not match evolution of Getty women's image usage (from 2007 to 2017) as headlined in a *New York Times* article as 'from sex object to gritty woman' (Miller, 2017). The researchers found that they rather reinforced pre-lean-

in collection themes. For example, it was found that portrayals of women in scientific or tech environments were found in stories about the challenges of breaking into the fields rather than ones simply about that work. Working women of colour were used mainly in publications with such a target group, rather than not. The most circulated image from the collection was of a young, white woman with long hair. Thus, the collection, at least in its actual usage, did not appear to meet its goals.

Computer vision techniques offer more than reverse image search for the study of resonance and circulation of particular imagery. They also label images, both in terms of content as well as format. When studying

CrowdTangle's coverage, however, has been questioned, given that misinformation (that would constitute election interference) may be circulated in posts not indexed by the system such as by individual accounts and private groups (Tech Transparency Project, 2020).

Image engagement

Studying engagement on platforms, through CrowdTangle or other marketing tools such as BuzzSumo, should not be equated with popularity measures, at least according to Facebook's head of news feed, John Hegeman (Newton, 2020). In an exchange on Twitter with Kevin Roose, the *New York Times* tech journalist, he made the distinction between engagement (measured by interactions) and popularity or reach (impressions), which is a common marketing distinction. The latter data point, however, is available only internally at Facebook, indicating the value of the data. The context of the exchange concerns how far-right conservative sources dominate the content with the most engagement on Facebook (Roose, 2020). In fact, the journalist has been making a daily list, posting it from the Twitter account, @FacebooksTopTen. Facebook's internal list of most popular (measured by 'reach'), contrariwise, were far less politically charged, thereby downplaying the journalist's claim and Facebook's culpability (see Figure 13).

Showing engagement by ranked list is one method of enabling interpretation (of who's on top and who's on the list), and another is a tree map, where visual media posts are resized according to engagement score. It is a visual analysis technique whereby one curates a list of pages or accounts (such as alt-right or less extreme 'alt-lite' actors) and subsequently determines which posts have received the highest engagement scores, resizing them by score. In the example of research conducted with the U.K. Home Office, it was found that anti-Islam (or counter-jihadist) posts particularly animated the alt-lite, as visualised, though it should be remarked that other, even more extreme posts may well have been removed (through content moderation) prior to the analysis (see Figure 14) (Alt-Right Open Intelligence Initiative, 2017).

Until the advent of what's referred to as 'platform lockdown', Facebook's Pages API enabled research into how well posts on Facebook are interacted with on a set of pages, such as social movement or protest group (Bruns et al., 2018). One still can undertake that work through the use of the FacePager software or manually. Facebook, however, is nudging researchers away from analysing posts on a predefined set of pages



Figure 13. Facebook pages with most engagement (interaction counts, on left) and greatest reach (view counts, on right). List tweeted by Kevin Roose (*New York Times*), with response by John Hegeman (Facebook).

to investigating top-performing posts or web URLs per keyword query.

Image-emoji associations

The comparative study of image formats, more generally, is one operationalisation of platform vernacular analysis. Others would be how hashtags, reactions or emojis are used together with images. On Facebook, for example, sad and angry reactions tend to be chosen together, as was found when analysing posts concerning the Syrian war (Geboers et al., 2020). In order to make findings such as these, one analytical technique is the production and interpretation of a co-occurrence network map, where images and another digital object or artefact (in this case, reactions) appear on the same map (see Figure 15).

Bunches of images together with hashtags, reactions or emojis come together in a graphing technique that seeks to optimise clustering. It relies on the 'networkedness' of the images by virtue of being linked by the same hashtag and/or reaction (Niederer, 2018).

Typically, in Gephi, the network analysis software, a force atlas algorithm is chosen followed by community detection; such a combination encourages distinct clusters to appear. Once so rendered, the second step in the

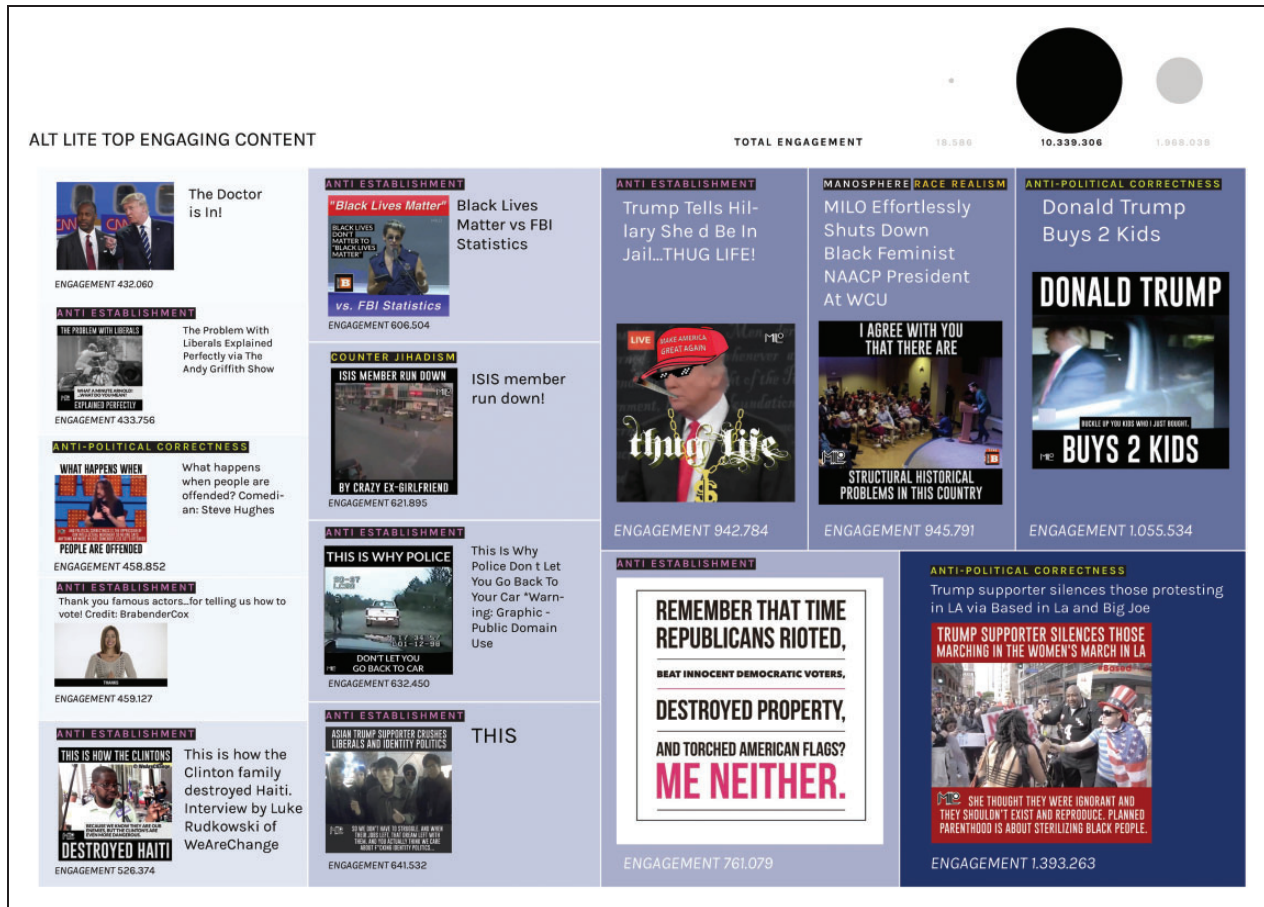


Figure 14. Most-engaged with Facebook posts of set of alt-lite pages, arrayed as tree map, 2016.

visual media network analysis is to label the clusters by inspecting their contents. For reactions and emojis, one would ask, which sentiments cluster with which images? For hashtag-image networks, the question concerns both how the images may be interpreted through their linkage to hashtags, and vice versa.

Apart from deriving and labelling clusters, visual network analysis has as part of its instrumentarium the study of centrality, brokerage and distance. Which reactions or emojis are central to a particular image space. How far apart are hashtag-image clusters? Which ones link or bridge the subspaces? Here, one annotates and tells stories with network maps.

Image removal

Scholarly attention to 4chan, the image board associated with extreme speech and visual content as well as snark, waxed considerably with the election of Donald Trump as U.S. president in 2016, given the role of/pol/, the 'politically incorrect' board, in rallying Trump support through 'meme magic' (Ling et al., 2021; Tuters, 2019). 4chan is ephemeral, meaning that it must be

routinely scraped or archived for its meaningful study. One can envisage an image grid timeline that is additive, showing in chronological order (per board) what is accruing on the platform. Arraying them in such a manner could show memeification or the hardening (or softening) of content, as also mentioned above in the discussion of the dominant image. Grouping through similarity in an image wall could show the extent to which memes dominate (and which memes).

In an approach where one platform is used to study another, researchers harvested YouTube URLs shared on 4chan/pol/ and entered these into YouTube at two intervals in time, fetching the video preview image (and other data) (OILAB, 2019) (see Figure 16). The before-and-after image walls display the widespread (and seemingly simultaneous) removal of videos, though with some still available for viewing. Here, one is able to visualise the scale and timing of content removal, using a presence/absence technique. The videos that survived the sweep become of special interest to study where the platform draws the line, given that they are not quite extreme enough.

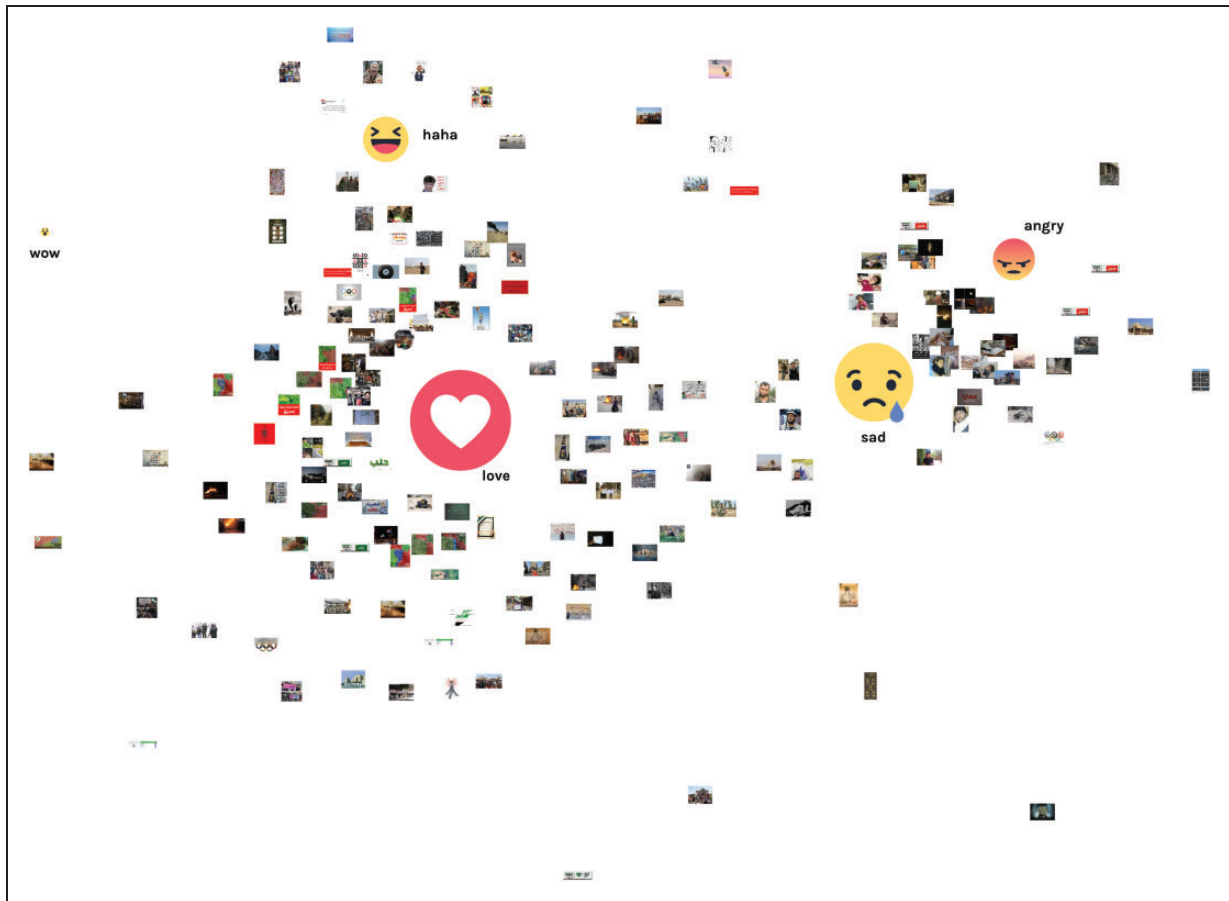


Figure 15. Image-reaction (emoji) bipartite graph from posts on Facebook pages concerning the Syrian war, 2017. Source: Scuttari et al. (2017).

Feed competition

The main purpose of putting source sets side by side is to consider alignment or agreement. Comparing Google Image search outputs in 2005 between ‘apartheid wall’ and ‘security fence’ show two distinctively different structures that had just been erected between Israel and the Palestinian Territories, the one concrete and graffitied, and the other lightweight and high-tech (Rogers, 2013). There was next to no agreement between the two sets of visual outputs when comparing the preferred official Israeli terminology compared to the Palestinian.

The comparative display of political media feeds has become prevalent in the past few years, given concern with filter bubbles (Pariser, 2011a) (see Figure 17), or homophilic recommendation systems (Chun, 2018), especially Facebook’s news feed. The ‘Blue Feed, Red Feed’ project by the *Wall Street Journal* (2016–2019) interactively displayed the stories published by conservative and progressive media sources on Facebook, refreshed hourly (Keegan, 2016).

The side-by-side presentation afforded a good view of perspectival divides in media. It was a simulation of stories the users would be expected to receive. In the Citizen Browser project by the data journalist group, The Markup, the technique is taken a step farther, in that they seek not just to simulate the divide in media but rather to demonstrate it by showing the viewers’ actual feeds, i.e. the ‘content social media companies choose to amplify to their users’ (Faife, 2021) (see Figure 18). Feed capture would be a step towards auditing the filter bubble, and the extent to which Facebook’s feed narrows or broadens horizons. Relying on nationwide panels of internet users who donate their feeds to the project, it was found that after the Capitol rampage progressive Facebook users were routinely fed mainstream media such as *CNN* and *NPR*, whilst conservative users sources considered rather less reliable (such as *The Daily Wire* and *Breitbart* most salient) (Lecher and Keegan, 2021). Reliability measures of sources may be gleaned from NewsGuard, among other media labelling organisations.



Figure 16. Content removal on YouTube. Comparison of extreme video availability before and after ‘purge’ on 6 June 2019. Source: OILAB (2019).

Visual media analysis

Above I discussed image reuse, image trends and vernaculars, dominant image, image presence, image quality, image staining (or tarnishing), image circulation, image engagement, image associations, image removal and feed competition as approaches for visual media analysis. Having collections of images as inputs, these

seek to output a kind of metapicture that retain and frame the images in a manner that invites critical reflection upon them. It enables close reading of images (through their retention) as well as more distant reading through their arrangement.

They also open lines of inquiry made implicit when the approach is combined with the metapicture technique. For example, do dominant images emerge when

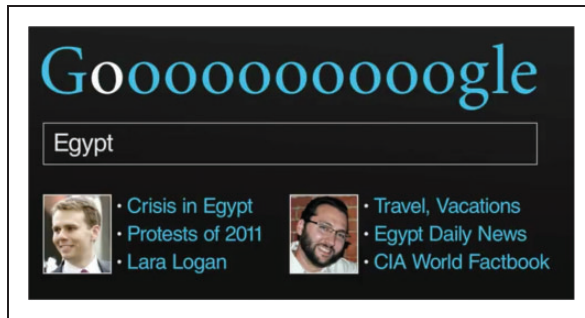


Figure 17. Filter bubble as differing personalised Google results. Source: Pariser (2011b).

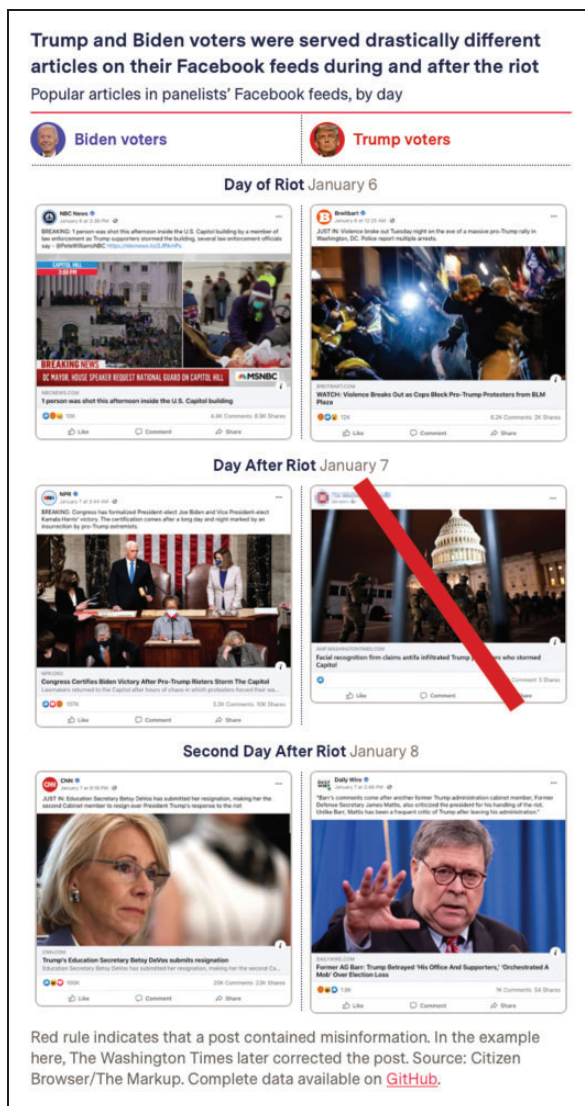


Figure 18. Feed competition. Side-by-side view of stories fed to feeds of liberal and conservative users of Facebook, showing competing accounts of events, January 2021. Source: Lecher and Keegan (2021).

grouping sets of them that share the same issue hashtag? When does a new style appear in a set of chronologically arrayed covers of a magazine? When placed side by side, to what extent do sets of pictures or media feeds show agreement? These formulations would exemplify how the visual media analysis outlined above seeks to make the metapicturing active in the research.

The critical reflection of the images through their arrangement is born of digital research methods that curate or demarcate a set of them by making use of the affordances of the platforms (e.g. hashtags as substantive, grouping markers). While there are exceptions, such as adopting a user persona (in the case of the feed columns of progressive and conservative Facebook users) and extracting data from one platform to understand the workings of another (the YouTube deletion image walls made from URLs on 4chan), much of the demarcation work is undertaken through querying a platform, be it for hashtags or keywords (or both).

The queries of the platform result in sets that could be thought to organise substantive spaces. Subsequently, in the next step, those spaces' content engagement is measured, whether taken from the interface (front-end or user mode) or through the API or other means by which a developer would access the data (back-end or developer mode). Comparison of the two is occasionally of interest, given that a user's view (on Facebook, for example) could be considered less expansive than the API's, especially if one is interested in (politicised) newsfeed critique. In sum, the visual media analysis is the study of engagement and ordering in demarcated online spaces.

Finally, in striving to display the results of the work, the notion of a metapicture is invoked as a technique that retains and frames the images under study so as to enable critical reflection of them. As related in the Introduction, it straddles the hermeneutic and the data-driven at once by de-datafying the images (in the sense that they remain images) as well as by arranging them as an analytical interpretation, such as the counter-meme as dominant image, stained engine results as tarnished or 'news' feeds as out of alignment or politically polarised.

Acknowledgements

The author would like to thank Gabriele Colombo, Sabine Niederer and the editors of *Big Data & Society* for their valuable comments..

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The research is supported by SoBigData-PlusPlus, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 871042.

ORCID iD

Richard Rogers  <https://orcid.org/0000-0002-9897-6559>

Supplemental material

Supplemental material for this article is available online.

References

- Abidin C (2014) #In\$tagLam: Instagram as a repository of taste, a burgeoning marketplace, a war of eyeballs. In: Berry M and Schleser M (eds) *Mobile Media Making in an Age of Smartphones*. New York: Palgrave Pivot, pp.119–128.
- Abidin C (2016) Visibility labour: Engaging with influencers fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia* 161(1): 86–100.
- Aggarwal A (2016) Detecting and mitigating the effect of manipulated reputation on online social networks. In: *Proceedings of the 25th international conference companion on World Wide Web*, pp.293–297. New York: ACM.
- Aiello G and Parry K (2020) *Visual Communication: Understanding Images in Media Culture*. London: Sage.
- Allan S (2015) “iPhone-wielding amateurs”: The rise of citizen photojournalism. In: Atton C (ed.) *The Routledge Companion to Alternative and Community Media*. London: Routledge, pp.357–366.
- Alt-Right Open Intelligence Initiative (2017) Mapping the Alt-Right: The US alternative right across the Atlantic. Digital Methods Summer School, University of Amsterdam. Available at: <https://wiki.digitalmethods.net/Dmi/AltRightOpenIntelligenceInitiative> (accessed 25 May 2021).
- Baker P and Potts A (2013) “Why do white people have thin lips?” Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies* 10(2): 187–204.
- Baym NK (2015) Connect with your audience! The relational labor of connection. *The Communication Review* 18(1): 14–22.
- Bechmann A (2017) Keeping it real: From faces and features to social values in deep learning algorithms on social media images. In: *Proceedings of the 50th annual Hawaii international conference on system sciences*, pp.1793–1801. New York: IEEE Computer Society.
- Bogers L, Niederer S, Bardelli F, et al. (2020) Confronting bias in the online representation of pregnancy. *Convergence: The International Journal of Research into New Media Technologies* 26(5–6): 1037–1059.
- Börner K and Polley DE (2014) *Visual Insights: A Practical Guide to Making Sense of Data*. Cambridge, MA: MIT Press.
- Borra E and Rieder B (2014) Programmed method: Developing a toolset for capturing and analyzing tweets. *Aslib Journal of Information Management* 66(3): 262–278.
- Boy JD and Uitermark J (2017) Reassembling the city through Instagram. *Transactions of the Institute of British Geographers* 42(4): 612–624.
- Bozzi N (2020) #digitalnomads, #solotravellers, #remote-workers: A cultural critique of the traveling entrepreneur on Instagram. *Social Media + Society* April–June: 1–15.
- Bruns A, Bechmann A, Burgess J, et al. (2018) Facebook shuts the gate after the horse has bolted and hurts real research in the process. *Internet Policy Review*. 25 April. Available at: <https://policyreview.info/articles/news/facebook-shuts-gate-after-horse-has-bolted-and-hurts-real-research-process/786> (accessed 25 May 2021).
- Bruns A and Highfield T (2012) Blogs, Twitter, and breaking news: The Prodisage of citizen journalism. In: Lind R (ed.) *Prodisage Theory in a Digital World: The Intersection of Audiences and Production in Contemporary Theory*. New York: Peter Lang, pp.15–32.
- Budge K (2020a) Visually imagining place: Museum visitors, Instagram, and the city. *Journal of Urban Technology* 27(2): 61–79.
- Budge K (2020b) Remember me: Instagram, selfies and libraries. *Journal of the Australian Library and Information Association* 69(1): 3–16.
- Burgess J and Baym NK (2020) *Twitter: A Biography*. New York: New York University Press.
- Cadwalladr C (2016) Google, democracy and the truth about internet search. *The Guardian*, 4 December. Available at: <https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook> (accessed 25 May 2021).
- Caliandro A and Graham J (2020) Studying Instagram beyond selfies. *Social Media + Society* April–June: 1–7.
- Casey E (1996) How to get from space to place in a fairly short space of time. In: Feld S and Basso KH (eds) *Senses of Place*. Santa Fe, NM: School of American Research Press, pp.13–52.
- Castro JML (2021) Framing the impact of pseudo-influencers via communication ethics. In: Watkins B (ed.) *Research Perspectives on Social Media Influencers and Brand Communication*. Lanham, MD: Lexington Books, pp.145–160.
- Chancellor S, Pater JA, Clear TA, et al. (2016) #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pp.1201–1213. New York: ACM.
- Chun WHK (2018) Queering homophily. In: Apprich C, Chun WHK, Cramer F, et al. (eds) *Pattern Discrimination*. Lüneburg: Meson Press, pp.59–98.
- Cinelli M, Quattrociocchi W, Galeazzi A, et al. (2020) The COVID-19 social media infodemic. *Scientific Reports* 10: 16598.
- Colombo G (2018) *The design of composite images: Displaying digital visual content for social research*. PhD thesis, Politecnico di Milano.
- Colombo G and De Gaetano C (2020) Dutch political Instagram: Junk news, follower ecologies and artificial amplification. In: Rogers R and Niederer S (eds) *The*

- Politics of Social Media Manipulation*. Amsterdam: Amsterdam University Press, pp.147–168.
- Confessore N, Dance GJX, Harris R, et al. (2018) The follower factory. *The New York Times*, 27 January. Available at: <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html> (accessed 25 May 2021).
- Crawford K and Paglen T (2019) Excavating AI: The politics of training sets for machine learning. Available at: <https://excavating.ai> (accessed 25 May 2021).
- Cresci S, Di Pietrob R, Petrocchi M, et al. (2016) Fame for sale: Efficient detection of fake Twitter followers. *Decision Support Systems* 80: 56–71.
- CrowdTangle (2021) Agence France-Presses strategy for using CrowdTangle search and meme search to report misinformation, CrowdTangle Help. Available at: <https://help.crowdtangle.com/en/articles/4153905-agence-france-presse-s-strategy-for-using-crowdtangle-search-and-meme-search-to-report-misinformation> (accessed 25 May 2021).
- d'Andrea C and 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: 1825–1845.
- De Amicis G, De Gaetano C, Saulière S, et al. (2013) Gezi Park life: From trees to cops, Digital Methods Summer School 2013, 11 July. Available at: https://wiki.digitalmethods.net/pub/Dmi/GeziParkLife/PP_GreenpeaceProject_Final.pdf (accessed 25 May 2021).
- Dean J (2019) Sorted for memes and gifs: Visual media and everyday digital politics. *Political Studies Review* 17(3): 255–266.
- Depoorter D (2019) Quick fix: Machine selling likes and followers. Available at: <https://driesdepoorter.be/quickfix/> (accessed 25 May 2021).
- EU DisinfoLab (2020) *COVID-19 and 5G: A Case Study of Platforms Content Moderation of Conspiracy Theories*. Brussels: EU DisinfoLab.
- Faife C (2021) In Georgia, Facebook's changes brought back a partisan news feed. *The Markup*, 5 January. Available at: <https://themarkup.org/citizen-browser/2021/01/05/in-georgia-facebooks-changes-brought-back-a-partisan-news-feed> (accessed 25 May 2021).
- Fehrmann G, Linz E, Schumacher E, et al. (2010) Original copy: Secondary practices. In: Jäger L, Linz E and Schneider I (eds.) *Media, Culture, and Mediality: New Insights into the Current State of Research*, pp.77–86. Bielefeld: Transcript-Verlag.
- Fishkin R (2019) Less than half of Google searches now result in a click. Sparktoro Blog, 13 August. Available at: <https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/> (accessed 25 May 2021).
- Geboers M, Stoloro N, Scuttari A, et al. (2020) Why buttons matter: Repurposing Facebook's reactions for analysis of the social visual. *International Journal of Communication* 14: 1564–1585.
- Geboers M and Van De Wiele CT (2020) Machine vision and social media images: Why hashtags matter. *Social Media + Society* April–June: 1–15.
- Gerlitz C and Rieder B (2013) Mining one percent of twitter: Collections, baselines, sampling. *M/C Journal* 16(2): 620.
- Gerrard Y (2020) Social media content moderation: Six opportunities for feminist intervention. *Feminist Media Studies* 20(5): 748–751.
- Gibbs M, Meese J, Arnold M, et al. (2015) #Funeral and Instagram: Death, social media, and platform vernacular. *Information, Communication & Society* 18(3): 255–268.
- Gronstad A and Vagnes O (2006) An interview with W.J.T. Mitchell. *Image and Narrative* 15. Available at: http://www.imageandnarrative.be/inarchive/iconoclasm/gronstad_vagnes.htm (accessed 25 May 2021).
- Hardt M and Negri A (2000) *Empire*. Cambridge, MA: Harvard University Press.
- Highfield T and Leaver T (2016) Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice* 2(1): 47–62.
- Hochman N (2014) The social media image. *Big Data & Society* 1(2).
- Howard PN, Ganesh B, Liotsiou D, et al. (2019) The IRA, social media and political polarization in the United States, 2012–2018, Computational Propaganda Project, Oxford University. Available at: <https://blogs.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/12/The-IRA-Social-Media-and-Political-Polarization-in-the-United-States-2012-2018.pdf> (accessed 25 May 2021).
- HypeAuditor (2021) Instagram audit & fake follower check. HypeAuditor. Available at: <https://hypeauditor.com/free-tools/instagram-audit/> (accessed 25 May 2021).
- Jeffries A and Yin L (2020) Google's top search result? Surprise! It's Google. *The Markup*, 28 July. Available at: <https://themarkup.org/google-the-giant/2020/07/28/google-search-results-prioritize-google-products-over-competitors> (accessed 25 May 2021).
- Karsgaard C and MacDonald M (2020) Picturing the pipeline: Mapping settler colonialism on Instagram. *New Media & Society* 22(7): 1206–1226.
- Keegan J (2016) Blue feed, red feed: See liberal Facebook and conservative Facebook, side by side. *Wall Street Journal*, 18 May. Available at: <https://graphics.wsj.com/blue-feed-red-feed/> (accessed 25 May 2021).
- Leaver T, Highfield T and Abidin C (2020) *Instagram: Visual Social Media Cultures*. Cambridge: Polity.
- Lecher C and Keegan J (2021) Biden and Trump voters were exposed to radically different coverage of the Capitol riot on Facebook. *The Markup*, 14 January. Available at: <https://themarkup.org/citizen-browser/2021/01/14/biden-and-trump-voters-were-exposed-to-radically-different-coverage-of-the-capitol-riot-on-facebook> (accessed 25 May 2021).
- Leszczynski A (2018) Digital methods II: Digital-visual methods. *Progress in Human Geography* 43(6): 1143–1152.
- Lindquist J (2018) Illicit economies of the internet: Click farming in Indonesia and beyond. *Made in China Journal* October–December. Available at: <https://madeinchinajournal.com/2019/01/12/illicit-economies-of->

- the-internet-click-farming-in-indonesia-and-beyond/ (accessed 25 May 2021).
- Ling C, AbuHilal I, Blackburn J, et al. (2021) Dissecting the meme magic: Understanding indicators of virality in image memes. In: *Proceedings of 24th ACM conference on computer-supported cooperative work and social computing*, p. 81. New York: ACM.
- Manovich L (2011a) Style space: How to compare image sets and follow their evolution. Available at: http://manovich.net/content/04-projects/073-style-space/70_article_2011.pdf (accessed 25 May 2021).
- Manovich L (2011b) What is visualisation? *Visual Studies* 26(1): 36–49.
- Manovich L (2016) The science of culture? Social computing, digital humanities and cultural analytics. *Journal of Cultural Analytics* 1(1).
- Manovich L (2017) *Instagram and Contemporary Image*. Available at: <http://manovich.net/index.php/projects/instagram-and-contemporary-image> (accessed 25 May 2021).
- Manovich L (2020) *Cultural Analytics* Cambridge, MA: MIT Press.
- Marwick A and Boyd D (2011) To see and be seen: Celebrity practice on twitter. *Convergence: The International Journal of Research into New Media Technologies* 17(2): 139–158.
- Massey PM, Kearney MD, Hauer MK, et al. (2020) Dimensions of misinformation about the HPV vaccine on Instagram: Content and network analysis of social media characteristics. *Journal of Medical Internet Research* 22(12): e21451.
- Miller CC (2014) LeanIn.org and Getty aim to change women's portrayal in stock photos. *The New York Times*, 9 February. Available at: <https://www.nytimes.com/2014/02/10/business/leaninorg-and-getty-aim-to-change-womens-portrayal-in-stock-photos.html> (accessed 25 May 2021).
- Miller CC (2017) From sex object to gritty woman: The evolution of women in stock photos. *The New York Times*, 7 September. Available at: <https://www.nytimes.com/2017/09/07/upshot/from-sex-object-to-gritty-woman-the-evolution-of-women-in-stock-photos.html> (accessed 25 May 2021).
- Mitchell WJT (1994) *Picture Theory: Essays on Verbal and Visual Representation*. Chicago: University of Chicago Press.
- Moretti F (2013) *Distant Reading*. London: Verso.
- Myers West S (2018) Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society* 20(11): 4366–4383.
- Newton C (2020) Why no one knows which stories are the most popular on Facebook. *The Verge*, 22 July. Available at: <https://www.theverge.com/interface/2020/7/22/21332774/facebook-crowdtangle-kevin-roose-nyt-tweets-interactions-reach-engagement> (accessed 25 May 2021).
- Niederer S (2018) *Networked Images: Visual Methodologies for the Digital Age*. Amsterdam: Hogeschool van Amsterdam.
- Niederer S and Colombo G (2019) Visual methodologies for networked images: Designing visualizations for collaborative research, cross-platform analysis, and public participation. *Revista Diseña* 14: 40–67.
- Niederer S and Colombo G (2021) The earnest platform: Coverage of the US presidential candidates, COVID-19, and social issues on Instagram. In: Rogers R (ed.) *Mainstreaming the Fringe: How Misinformation Propagates on Social Media Platforms*. Amsterdam: Amsterdam University Press, forthcoming.
- Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- OILAB (2019) 4chan's YouTube: A fringe perspective on YouTube's great purge of 2019. *OILAB Blog*, 17 June. Available at: <https://oilab.eu/4chans-youtube-a-fringe-perspective-on-youtubes-great-purge-of-2019/> (accessed 25 May 2021).
- Ozduzen O and McGarry A (2020) Digital traces of “twitter revolutions”: Resistance, polarization, and surveillance via contested images and texts of occupy Gezi. *International Journal of Communication* 14: 2543–2563.
- Paquet-Clouston M, Bilodeau O and Décary-Héty D (2017) Can we trust social media data? Social network manipulation by an IoT botnet. In: *Proceedings of the 8th international conference on social media & society*, pp.1–9. New York: ACM.
- Pariser E (2011a) *The Filter Bubble*. New York: Penguin.
- Pariser E (2011b) Beware “online filter bubbles”. Ted Talk, March. Available at: https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles (accessed 25 May 2021).
- Pearce W, Niederer SD, Gaetano C, et al. (2019a) Changing visual vernaculars of climate. Digital Methods Summer School 2019, University of Amsterdam. Available at: <https://wiki.digitalmethods.net/Dmi/ChangingVisualVernacularsOfClimate> (accessed 25 May 2021).
- Pearce W, Niederer S, Özkula SM, et al. (2019b) The social media life of climate change: Platforms, publics, and future imaginaries. *WIREs Climate Change* 10: e569.
- Pearce W, Özkula SM, Greene AK, et al. (2020) Visual cross-platform analysis: Digital methods to research social media images. *Information, Communication & Society* 23(2): 161–180.
- Pink S (2008) An urban tour: the sensory sociality of ethnographic place-making. *Ethnography* 9(2): 175–196.
- Purba KR, Asirvatham D and Murugesan RK (2020) Classification of Instagram fake users using supervised machine learning algorithms. *International Journal of Electrical and Computer Engineering* 10(3): 2763–2772.
- Rieder B, Matamoros-Fernández A and Coromina Ó (2018) From ranking algorithms to ranking cultures: Investigating. *Convergence: The International Journal of Research into New Media Technologies* 24(1): 50–68.
- Robinson S (2009) “If you had been with us”: Mainstream press and citizen journalists jockey for authority over the collective memory of hurricane Katrina. *New Media & Society* 11(5): 795–814.
- Robinson S and DeShano C (2011) “Anyone can know”: Citizen journalism and the interpretive community of the mainstream press. *Journalism* 12(8): 963–982.
- Rogers R (2013) *Digital Methods*. Cambridge, MA: MIT Press.

- Rogers R (2018) Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication* 12: 450–472.
- Rogers R (2019) *Doing Digital Methods*. London: Sage.
- Rogers R and Sendijarevic E (2012) Neutral or national point of view? A comparison of Srebrenica articles across Wikipedia language versions. In: *Proceedings of Wikipedia Academy: Research and free knowledge*, 29 June–1 July, Berlin, Germany.
- Romm T (2020) Justice Department sues Google, alleging multiple violations of federal antitrust law. *Washington Post*, 20 October. Available at: <https://www.washingtonpost.com/technology/2020/10/20/google-antitrust-doj-law-suit/> (accessed 25 May 2021).
- Roose K (2020) What if Facebook is the real “silent majority”? *The New York Times*, 27 August. Available at: <https://www.nytimes.com/2020/08/27/technology/what-if-facebook-is-the-real-silent-majority.html> (accessed 25 May 2021).
- Rose G (2016) *Visual Methodologies*. London: Sage.
- Rubinstein D and Sluis K (2008) A life more photographic: Mapping the networked image. *Photographies* 1(1): 9–28.
- Sandvig C, Hamilton K, Karahalios K, et al. (2016) When the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication* 10: 4972–4990.
- Schur D (1998) *An Introduction to Close Reading*. Cambridge, MA: Harvard University.
- Scott M (2017) Google fined record \$2.7 billion in E.U. antitrust ruling. *The New York Times*, 27 June. Available at: <https://www.nytimes.com/2017/06/27/technology/eu-google-fine.html> (accessed 25 May 2021).
- Scuttari A, Stolerio N, Ridley A, et al. (2017) Emotional clicktivism: Facebook reactions and affective responses to visuals. Digital Methods Summer School 2017, University of Amsterdam. Available at: <https://digitalmethods.net/Dmi/EmotionalClicktivism> (accessed 25 May 2021).
- Sen I, Aggarwal A, Mian S, et al. (2018) Worth its weight in likes: Towards detecting fake likes on Instagram. In: *Proceedings of the 10th ACM conference on web science*, pp.205–209. New York: ACM.
- Senft TM (2013) Microcelebrity and the branded self. In: Hartley J, Burgess J and Bruns A (eds) *A Companion to New Media Dynamics*. London: Wiley, pp.346–354.
- Sinders C (2020) *Feminist Data Set*. USA: University of Denver.
- Software Studies Initiative (2011) ImagePlot software v. 1.1. Available at: <http://lab.softwarestudies.com/p/imageplot.html> (accessed 25 May 2021).
- Stepnik A, Martin A, Benedetti A, Karsgaard C, Yee Ting Ng C, et al. (2020) Black squares as (in)authentic behaviour: Displays of solidarity on Twitter, Instagram and Facebook. Digital Methods Summer School, University of Amsterdam. Available at: <https://wiki.digitalmethods.net/Dmi/SummerSchool2020BlackSquares> (accessed 25 May 2021).
- Steyerl H (2009) In defense of the poor image. *e-flux Journal* 10. Available at: <https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/> (accessed 25 May 2021).
- Sweeney L (2013) Discrimination in online ad delivery. *Communications of the ACM* 56(5): 44–54.
- Tech Transparency Project (2020) Facebook leans on states to spot voter interference. Tech Transparency Project, 23 September. Available at: www.techtransparencyproject.org/articles/facebook-leans-states-spot-voter-interference (accessed 25 May 2021).
- Tifentale A and Manovich L (2015) Selfiecity: Exploring photography and self-fashioning in social media. In: Berry DM and Dieter M (eds) *Postdigital Aesthetics: Art, Computation and Design*, pp. 109–122. Cham: Springer.
- Torres G and Rogers R (2020) Political news in search engines: Exploring Googles susceptibility to hyperpartisan sources during the Dutch elections. In: Rogers R and Niederer S (eds) *The Politics of Social Media Manipulation*. Amsterdam: Amsterdam University Press, pp.97–122.
- Tuters M (2019) LARPing & liberal tears: Irony, belief and idiocy in the deep vernacular web. In: Fielitz M and Thurston N (eds) *Post-Digital Cultures of the Far Right*. Bielefeld: Transcript, pp.37–48.
- Vaidhyanathan S (2012) *The Googlization of Everything*. Berkeley, CA: University of California Press.
- Van Den Hoogen J (2019) *The engineering of the self online*. MA Thesis, University of Utrecht.
- Visual Computing (2018) ImageSorter, HTW Berlin - University of Applied Sciences. Available at: <https://visual-computing.com/project/imagesorter/> (accessed 25 May 2021).
- Walker Rettberg J (2014) *Seeing Ourselves through Technology: How We Use Selfies, Blogs and Wearable Devices to See and Shape Ourselves*. Houndmills: Palgrave Macmillan.
- Wall M and El Zahed S (2015) Syrian citizen journalism: A pop-up news ecology in an authoritarian space. *Digital Journalism* 3(5): 720–736.
- Zizek S (1998) *The Interpassive Subject. Traverses*. Paris: Centre Georges Pompidou.