

# Interpreting Climate Images on the Internet: Mixing Algorithmic and Interpretive Views to Enable an Intercultural Comparison (ANCI)

---

Birgit Schneider, Thomas Nocke, Paul Heinicker, Janna Kienbaum

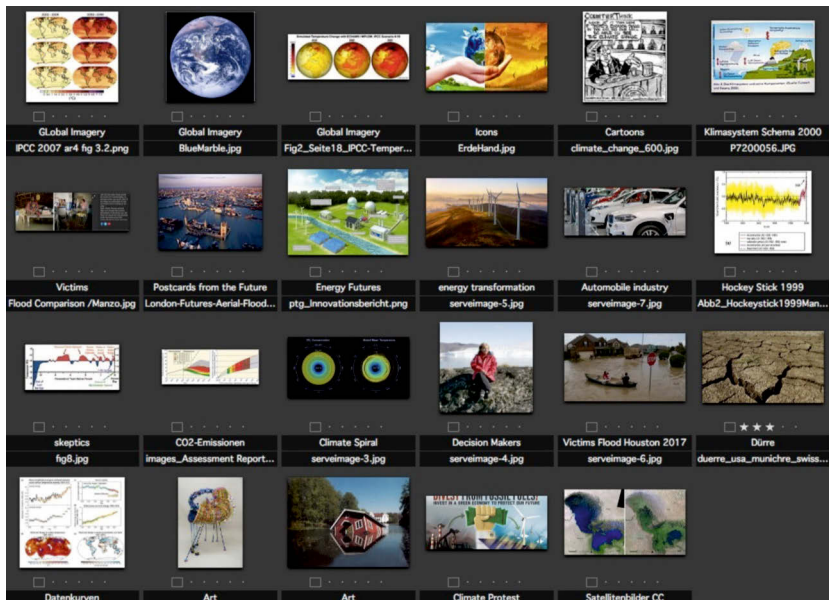
**Abstract** *Countless climate images are in circulation on the internet, such as burning globes, polar bears and photos of global climate impacts. These images are networked and generate a specific view of climate change. Our case study engages in an intercultural image comparison based on Google Image queries. As an interdisciplinary team of experts drawn from art history, media studies, interface design and computer graphics, our goal was to use a combination of qualitative and quantitative image analyses to explore the predominant discourses of digitised visual climate communication on the web. To this end, we automated the analysis of different formal features of climate images (such as colour values, density and composition) with the aid of computer-driven methods (such as computer vision and machine learning) to build a corpus of thousands of images. Our focus was on image similarities, a concept shared by both image theory and computer analysis. In this chapter, we elucidate the outcome of our research on a conceptual and technical basis. The core issue addressed here is the manner in which art-historical methods (such as iconography and the concept of visual framing) are transformed when using computer-generated methods of computer vision and machine learning to analyse image similarities. This chapter focuses on our various insights while also reflecting on the general question of networked images on a methodological level. Ultimately, we were able to identify the promising potential but also the key limits of algorithmic image recognition and sorting when using machine learning to study images on the internet.*

## Introduction

The research project *anci*—an acronym for analysing networked climate images—addresses the topic of climate change and its visual impact in communication. Ever since the popularisation of the World Wide Web, climate images have increasingly circulated online. Some motifs became deeply rooted in cultural web memory and now influence the world view of global warming in the realm between politics, science, art and popular culture. For the purpose of our research, we compared and

analysed these images and search for similarities with the help of ‘catch images’ (analogous to catch words), a term coined by the cultural scholar Aby Warburg (1866–1929).<sup>1</sup> The term refers to images that dominate the current global climate picture language in scientific, special or political contexts, such as photographs of the globe as a ‘Blue Marble’ and temperature curves of global warming in a hockey stick shape (Fig. 1).

Fig. 1: Selection of catch images. Excerpt from the Expression Media Pro database ‘Klimatathek’ by Birgit Schneider. Images in this database are sorted by means of hand-tagged keywords.



By using a method of arranging image reproductions on large panels in order to understand their relatedness, Warburg sketched out a key method for comprehending networked images in the early twentieth century.<sup>2</sup> Today, in the age of image databases, his approach provides a great model for digital image research. While

1 “The interest here is primarily in the transfer of the image into the sphere of everyday political life as well as the iconography of these everyday images, which are apostrophized as public images or—to use a fitting word from Aby Warburg—as striking images.” Diers, *Schlagbilder: Zur politischen Ikonographie der Gegenwart*, 7 / Diers, *Catchimages: On the Political Iconography of the Present*, 7 (translated using DeepL).

2 See Aby Warburg, *Der Bilderatlas Mnemosyne / The image atlas Mnemosyne* (translated using DeepL) (Berlin: Akademie Verlag, 2008).

analysing networked images, similarity becomes systematically and methodically important as a concept, because it serves as a way to group images. Already in the late nineteenth century, the systematic comparison of images shaped a medial dispositive in academic art instruction.<sup>3</sup> Today, it is increasingly being used as an “intellectual operation”<sup>4</sup> in computer-controlled calculation procedures, such as computer vision and machine learning. We used Warburg’s methodology as a starting point for our digital image studies.

## Climate images sorted by Google Images

We applied our mixed-methods approach to the general question of climate communication, as our interest was limited to mainstream image communication in cultural spaces. We therefore deliberately did not examine social media; instead, we were interested in the largest possible, extensively used and highly formalised media spaces.

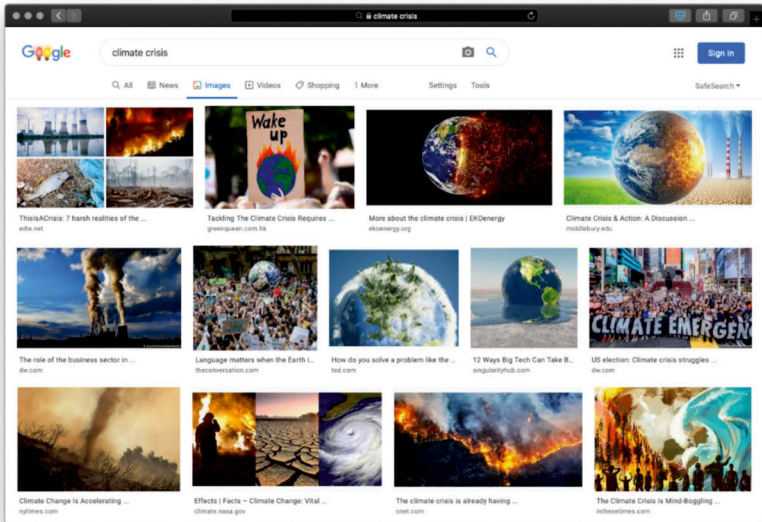
In contrast to other cultural data presented in this book, climate images on the internet are ‘native’ digital sources that are interconnected via HTTP protocol and have metadata. If you use Google to search for climate images, you will encounter a fairly standardised visual language for representing climate change—one that is dominated by globes, smoking chimneys and impacted landscapes (Fig. 2). At the same time, if you look at larger amounts of searched images, there are considerable differences in the preferred types of images (e.g., photos, maps, curves, pictograms or cartoons) and also in the framing of climate change.

---

3 See Heinrich Dilly, “Lichtbildprojektion—Prothese der Kunstbetrachtung” / “Photo projection—prosthesis of art perception” (translated using Deepl), in: Horst Bredekamp et al. (ed.): *Kunstwissenschaftliche Untersuchungen des Ulmer Vereins, Verband für Kunst- und Kulturwissenschaften* (Gießen: Anabas, 1975), 153–172.

4 Felix Thürlemann, “Bild gegen Bild: Für eine Theorie des Vergleichenden Sehens” / “Image against Image: For a Theory of Comparative Seeing” (translated using Deepl), in: Aleida Assmann, Ulrich Gaier, Gisela Trommsdorff (eds.): *Zwischen Literatur und Anthropologie—Diskurse, Medien, Performanzen* (Tübingen: Narr, 2005), 176.

Fig. 2: Research query via Google Image Search using the keyword 'climate crisis' taken on 2nd December 2020, Berlin, Germany, screenshot.







Since the 1990s, the automated indexing of websites has provided some of the main interfaces with which we access the World Wide Web. Alphabet Inc.'s Google is considered to be the dominant search engine in many parts of the world; it is also the most visited website ever, with several billion searches being clicked every day.<sup>5</sup> Despite acknowledging that Google's web service does not provide a neutral representation of hierarchies on the World Wide Web, we used the image query of Google Images as the basis of our image corpus. With the aid of the *Tor* browser and its VPN tunnelling capabilities, we technically imitated search queries from different cultural regions relating to climate issues. In total, we extracted roughly 16,000 images using eight keywords (e.g., climate change, climate change disaster, global warming) throughout seven cultural regions.<sup>6</sup> These images were accessed and edited via an online interface based on the Tabulator framework (Fig. 3).<sup>7</sup>

5 Statista, "Number of explicit core search queries powered by search engines in the United States as of April 2021".

6 Australia, USA, Brazil, Germany, Kenya, United Arab Emirates, Bangladesh. Instead of countries, we speak of locales in this regard. These locales serve as a cultural spatial unit. A locale combines the idea of a specific language with a region or country.

7 <https://tabulator.info/>.

Fig. 3: Web interface of our database of extracted images from Google Images according to specific keywords. Screenshot, 2020.

Image	Last modified	Image deter...	Width	Height	Country	Language	Search term
	2018/04/06	wheeled vehicle	283	400	Germany	German	climate change
	2018/06/01	cross	620	207	Germany	German	climate change
	2017/10/18	earth	912	513	Germany	German	climate change
	2018/01/23	geodesic dome	1200	800	Germany	German	climate change

## Methodology

In our research design, we applied specific combinations of methodological approaches. We oriented ourselves towards the mixed-methods designs put forward by the empirical social scientist John Creswell.<sup>8</sup> Two central meta-methods were of crucial importance in our research project: on the one hand, the method of *image comparison and image similarity* (qualitative) and, on the other, the method of *data visualisation*.

For our qualitative analysis of images, we combined different methods and disciplinary perspectives. On the level of *Bildwissenschaft* (the term for this German field

8 We oriented ourselves to Sequential Exploratory Design, the Sequential Explanatory Design and the Embedded or Nested Design. Cf. Creswell; Clark, *Designing and Conducting Mixed Methods Research*. For our image study on cross-cultural image comparison with Google, we decided to use Creswell's 'embedded design'. Although quantitative and qualitative approaches are still strictly separated here, these can be considered to be more in dialogue due to the simultaneity and dependence of the design. The quantitative perspective dominates the qualitative perspective, although interventions within the ongoing methodological approach can make it more flexible.

of image research can be translated as visual or image studies) and art history, we used image comparison in iconographic and iconological analysis methods,<sup>9</sup> but we also drew on rhetoric-semiotic<sup>10</sup> approaches. For the analysis of the distribution and migration of images, we turned to media-science approaches and collective image discourses.<sup>11</sup> For our cross-cultural image comparison with Google Images, we drew on the framing method<sup>12</sup> as used in communication science in tandem with the art-scientific theory of iconology as put forward by Erwin Panofsky.<sup>13</sup>

Frames generate their power through a symbolic charge. Images, in particular, make it possible to quickly and easily assimilate content that—in comparison to text—tends to lead latently to specific constructions of meaning: “Images are powerful framing tools because they are less intrusive than words and as such require less cognitive load. Therefore, peripheral rather than central processing may be activated and audiences may be more likely to accept the visual frame without question.”<sup>14</sup> The visual frames can be derived from the image contents in combination

- 
- 9 See Panofsky, *Sinn und Deutung in der bildenden Kunst / Meaning and interpretation in the visual arts* (translated using DeepL) (Köln: Dumont, 1978 [1955]).
- 10 See Lynda Walsh, “The Visual Rhetoric of Climate Change”, in: *Wires Climate Change*, Volume 6, Issue 4, 361–368, 2015; See Lynda Walsh, “Tricks, Hockey Sticks, and the Myth of Natural Inscription: How the Visual Rhetoric of Climategate Conflated Climate with Character”, in: Thomas Nocke, Birgit Schneider (eds.): *Image Politics of Climate Change Visualizations, Imaginations, Documentations* (Bielefeld: Transcript, 2014), 81–105; See also Birgit Schneider, *Klimabilder: Eine Genealogie globaler Bildpolitiken von Klima und Klimawandel / Climate Images: A Genealogy of Global Image Politics of Climate and Climate Change* (translated using DeepL) (Berlin: Matthes & Seitz, 2018).
- 11 See Jürgen Link, “Literaturanalyse als Interdiskursanalyse: Am Beispiel des Ursprungs literarischer Symbolik in der Kollektivsymbolik” / “Analysis as Interdiscourse Analysis: Using the Example of the Origin of Literary Symbolism in Collective Symbolism” (translated using DeepL). In: Jürgen Fohrmann, Harro Müller (eds.): *Methoden diskursanalytischer Ansätze* (Frankfurt a.M.: Suhrkamp, 1988).
- 12 See Robert Entman, “Framing: Toward Clarification of a Fractured Paradigm”, in: Mark R. Levy (ed.): *Journal of Communication* (Volume 43, Issue 4, 1993); Jörg Matthes, *Framing* (Baden-Baden: Nomos, 2014); Elke Crittman, “Visual Frames—Framing Visuals. Zum Zusammenhang von Diskurs, Frame und Bild in den Medien am Beispiel des Klimawandeldiskurses” / “Visual Frames – Framing Visuals. On the Connection between Discourse, Frame and Image in the Media Using the Example of Climate Change Discourse” (translated using DeepL), in: Stephanie Geise, Katharina Lobinger (eds.): *Visual Framing: Perspektiven und Herausforderungen der visuellen Kommunikationsforschung* (Köln: Halem, 2015), 95–116; Stephanie Geise, Katharina Lobinger, *Visual Framing: Perspektiven und Herausforderungen der Visuellen Kommunikationsforschung / Visual Framing: Perspectives and Challenges in Visual Communication Research* (translated using DeepL) (Köln: Halem, 2015); Daniela V. Dimitrova, Lulu Rodriguez, “The levels of visual framing”, in: *Journal of Visual Literacy*, (30:1), DOI: 10.1080/23796529.2011.11674684, 48–65.
- 13 See Panofsky, *Sinn und Deutung in der bildenden Kunst / Meaning and interpretation in the visual arts* (translated using DeepL).
- 14 Dimitrova; Rodriguez, The levels of visual framing, 50.

with their mode of representation (e.g., shooting angle of a photograph): “Visuals, like text, can operate as framing devices insofar as they make use of various rhetorical tools—metaphors, depictions, symbols—that purport to capture the essence of an issue or event graphically.”<sup>15</sup>

We chose the framing approach so as to systematically do justice to the vast amount of Google images. Here we combined deductive and inductive approaches. The framing categories used in communication studies and developed by communication scholar Robert Entman allowed us to deductively filter and analyse climate images by means of algorithmic image analysis, that is, by emphasising a ‘problem’, ‘cause’, ‘solution’ or ‘moral aspect’. Also, we were able to inductively derive frames from the image corpus based on the algorithmically generated image clusters.

Another method which helped us analyse algorithmically generated image clusters was *data visualisation*. We understand data visualisation as a mediator between computer science and art history or *Bildwissenschaft* (image studies). It makes the inherent complexity of data and abstract processes visible and thus tangible for the different perspectives and content analysis. Indeed, data visualisation methods were key to our analysis on many levels. They helped us when visualising statistical results using diagrams, such as tables or interfaces, but also when clustering corpora as image plots (structured overviews by different features). And, finally, they made it possible for us to illustrate the processes of machine learning itself. The resulting data visualisation generated by our research combined all previous algorithms into one image.

The entire process—beginning with image data collection (web scraping) and ending with visualisation—could be portrayed in the form of a pipeline (Fig. 4). In the first step, we compared the extracted images for similarities by means of a pre-trained machine-learning model. The result was a high-dimensional data set that could not be understood at first sight, because it was too complex. For this reason, we reduced it to a dimension level that was perceptible to humans by using another machine-learning algorithm called *t-SNE*.<sup>16</sup> We then visualised the reduced data set and grouped it using a clustering algorithm called *k-means*.<sup>17</sup> The potential offered by the visualisation comes in the form of an interface between humans and mechanical processes, thus providing crucial access for all the described processes (Fig. 5). The challenge lies in working with such highly abstracted images that can certainly mask the underlying processes to some extent.

---

15 Dimitrova; Rodriguez, The levels of visual framing, 51.

16 See Geoffrey Hinton and Laurens van der Maaten, Visualizing Data using t-SNE (2008).

17 See J. MacQueen, “Some methods for classification and analysis of multivariate observations”, in: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, (1, 1967), 281–297.

Fig. 4: The tools and algorithms that were used and the data that were visualised. Image: *anci* 2019.

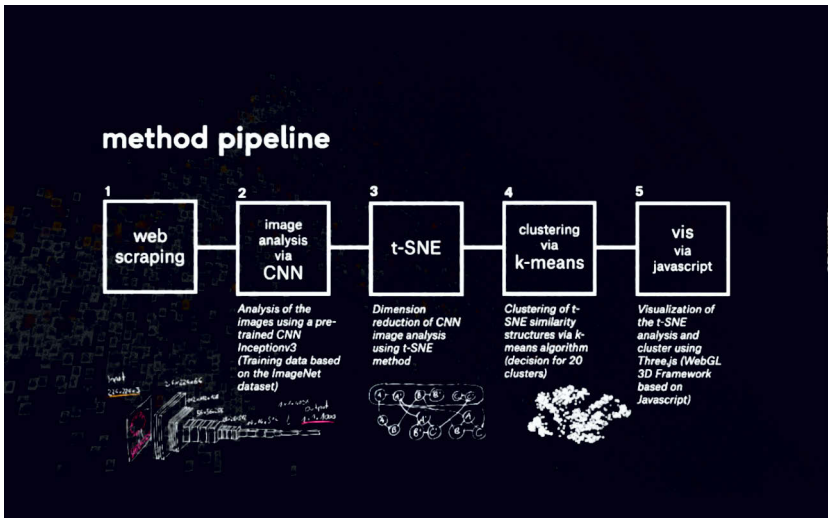
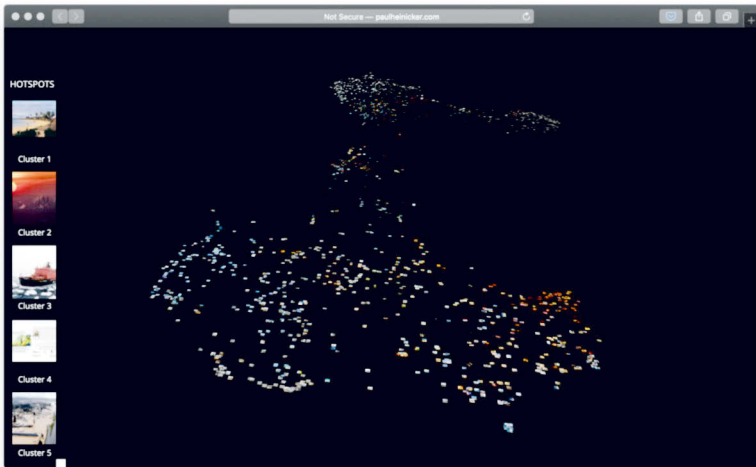


Fig. 5: Screenshot of the interface built on the basis of PixPlot by Yale Digital Humanities Lab. Screenshot, 2020.<sup>18</sup>



18 See Yale DH Lab, 'PixPlot', accessed 19 July, 2021, <https://dhlab.yale.edu/projects/pixplot/>.



## Key motifs and image types

We were able to identify a number of dominant and/or salient motifs and image genres in the scanned images. These motifs were expressed either as technically generated image clusters or as qualitative image groups although differences in the various locales became obvious, too. The key motifs are as follows:

*Text-image documents, including*

- Text-only documents*
- Diagrams*
- Covers of books and brochures,*
- PowerPoint slides with text,*
- Posters with slogans about climate protection*

*Maps*

*Infographics and charts*

*Occasional cartoons or cartoon-style graphics*

*People, including*

- Conferences, group meetings (most of which could be linked to a specific political context)*
- Scenes of protest and demonstrations*

*Depictions of the earth as a globe*

*Contrasting depictions/diptychs: green landscape versus parched ground*

*Forest fires*

*Ice landscape, e.g., polar sea with ice floes*

*Floods/flooding*

*Drought/dry soil*

*CO<sub>2</sub> emissions, refineries*

*Polar bears*

*Forest/trees/plants*

*Agriculture/farming/cultivation*

Our content review of the image types allowed us to determine the central image types and/or genres within the Google search results. For the most part, three types appeared, sorted by average weighting:

- 1. Photographs, divisible into images depicting (a) the environment, nature, landscape, (b) people (conferences, groups, politics versus protest/demonstrations), (c) the earth, and (d) the polar bear
- 2. Text-image documents, including text-only documents (documents, diagrams, covers of books and brochures, slides with text, activist posters with slogans, maps, cartoons etc.)

- 3. Highly artificial photomontages (representations of contrast of planet earth and landscapes).

## Intercultural framing

The t-SNE algorithm serves as a method with which to frame images. The automated image analysis and clustering process of the above-mentioned key motifs, search images and image types helped us to derive statements about visual frames. First is the *frame of vulnerability*, which includes the consequences and impacts of climate change as a problem definition and is the largest image group in all country queries. It includes photographs of icy landscapes (polar sea/glacier/ice floes), drought, flood, heat/forest fire and the polar bear cluster.

Equally dominant and prevalent alongside the frame of vulnerability is the *frame of cause representation*. This was demonstrated in all image queries involving the subject of CO<sub>2</sub> emissions (refineries/chimneys with smoke). Photographs of people are also concisely represented; and they can be assigned to the visual *frame of actors*. They symbolise the frame that emphasises approaches to solutions: on the one hand, they tend to be photographs of politicians and international climate conferences, and on the other, they are photographs showing scenes of protest and demonstrations. One particularly interesting result of the Google images was the photographs we were able to assign to the *frame of moral evaluation*. Most significant here were the images relating to the *frame of global anxieties* as embodied in artificially assembled photographs representing the Earth as either the iconic Blue Marble or as a globe or as a burning planet.

Our comparison of the image clusters or groups (via the screenshot table, see Fig. 7) made it possible to determine country-specific distributions as a tendency. The following are our key findings in this realm for the year 2020:

- The t-SNE visualisations of the Google images from Bangladesh and Kenya show the largest number of images featuring people. In particular, they contain scenes of politics, negotiations, conferences and/or meetings. Based on random sampling, the motifs tended to be locally assignable, such as one large set of images representing Prime Minister Sheikh Hasina Wajed of Bangladesh. Surprisingly, very-few-to-no-images featuring politicians could be found in the t-SNE visualisations of the American, Australian and German search queries.
- The Kenyan search query contained the largest proportion of depictions of people engaged in cultivation, followed by Bangladesh and the United Arab Emirates. In an international comparison, Kenyan Google images have the largest proportion of subjects relating to agriculture; and in comparison to the other search queries, this image theme is only marginally represented.

- Also surprising is the marginality of protest images. Images of predominantly young climate protection activists stand for the global use of this image type. Various photos are used 'globally', that is, the same motif appears in different countries. For example, a photo of protesting students and climate activists in New York's Times Square can be found in Google images from Bangladesh and Australia.
- One cluster of polar bears and one of ice landscapes appear throughout every search group. Only in the Kenyan search query is such a cluster missing. Surprisingly, on the other hand, the Kenyan t-SNE visualisation did not show any image cluster relating to the theme of drought.
- With regard to the vulnerability frame, we made an interesting finding about forest images. Internationally, the forest as a motif is predominantly shown to be burning, most clearly in the Google images from the USA, Australia, the United Arab Emirates and Brazil. Only the German Google images showed a group of images that included a technically generated hotspot image depicting the forest as a dried up patch. It would appear that the forest fire in the form of dramatic and high-contrast photographs (black-orange) is superior as a motif of vulnerability to that of drought.
- In general, the theme of heat dominates. The burning globe becomes a denotation of global warming in contrast to the 'Blue Marble'. The t-SNE visualisations, especially those from the USA, Australia and the United Arab Emirates followed by Brazil show a very high proportion of this image motif. In the t-SNE for Kenya and Bangladesh, this image group could not be detected at all.
- A country-specific statement is perhaps identifiable in the group with motifs of flooding. The largest image cluster was located in the search query of Bangladesh, followed by those of the USA and Australia. In particular, photographs featuring people standing in flood waters are found here. Flood scenes from a bird's-eye view, on the other hand, exhibit a global character. For example, photographs showing flooding in Texas after Hurricane Harvey in 2017 are found in Google images from Bangladesh.

## Method reflection

The research theme of networked climate images is already embedded in quantitative rules of logic, for example digital image files published and distributed via hypertext transfer protocols on the internet. For this reason, our interests here are more concerned with the degree to which we can have qualitative approaches in digital infrastructures. We want to know how quantitative and qualitative approaches are entangled and how they are differentiated.

Just as our qualitative research is based on established conceptual frameworks, our quantitative analysis is similarly dependent on externally prepared data sources, programming interfaces (APIs) and pre-written algorithms. This is due to the high degree of technical complexity involved in algorithms for image and context analysis and data aggregation on the internet. It means that in most cases it is possible to work solely within a constructed technical framework. This framework, however, is often characterised by services provided by technology and software companies the ‘modes of operation’ of whose algorithms are difficult to understand.

The *PageRank* algorithm, for example, which orders the search results on Google Images, is responsible for the type and diversity of climate images that we analyse. Around 200 factors influence this algorithm, some of which are known but the most relevant of them are kept well hidden.<sup>19</sup> This is why such a technical framework is often referred to as a ‘black box’. Still, it is possible to take a glimpse into the generative system and thus into the construction of the data by using open-source solutions and self-programmed algorithms and the exercise should be part of the analysis.

The qualitative data of our research is based on an analysis of climate images by humanistic means of interpretation. However, the distinction between qualitative and quantitative data suggests a categorically different treatment. Our research experience instead shows that both sides need to be approached with particular scepticism. This is because both share the belief of being artificially constructed and convey—in addition to their values—underlying assumptions and prerequisites. Nevertheless, it is important to note that qualitative data is closer to an anthropocentric understanding of insights whereas quantitative data follows a more structural logic.

In the course of our mixed-methods research, it became clear that it was almost impossible to handle exclusively qualitative data. Instead, qualitative and quantitative approaches interfere with one another in manifold ways. A clear separation between purely qualitative and purely quantitative data could not be maintained. A strict split can be obtained solely on a heuristic level, but it simply doesn’t hold true in practice.

Computer-generated image analysis represents the result of various qualitative decision-making processes:

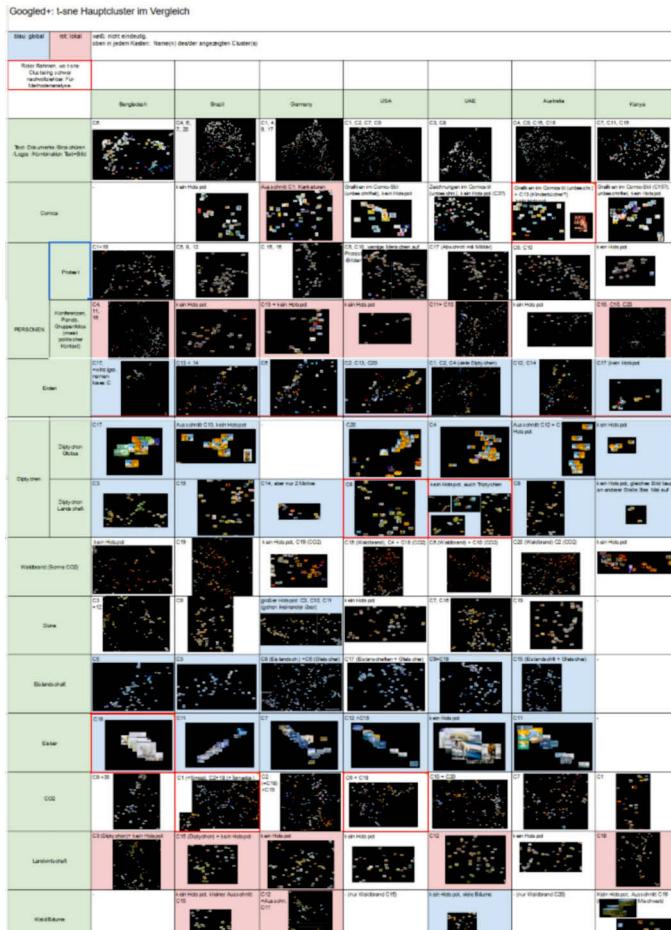
1. The selection of cultural regions to research climate images was undertaken via qualitatively predefined indices.
2. The Google Images corpus itself was based on predefined and partly hidden functionalities of the *PageRank* algorithm.

---

19 See Danny Sullivan, “Dear Bing, We Have 10,000 Ranking Signals to Your 1,000. Love, Google”, in: *Search Engine Land*, accessed July 19, 2021, <https://searchengineland.com/bing-10000-ranking-signals-google-55473>.

3. The extraction of images was based on a pre-existing framework and network technologies making use of the open-source software Tor to conceal the user's location.
4. Machine learning on inception (pre-trained similarity)
5. We clustered the calculated results that were visualised in a web interface so that groupings became visible to the human eye and could be qualitatively read and used to derive the key motifs, for example in our screenshot table (Fig. 7).

Fig. 7: Screenshot table with image groups of the t-SNE visualisations.



## Collaborating in an interdisciplinary team

In the process of our interdisciplinary cooperation, we became increasingly aware of the extent to which humanistic methods of art history and image analysis were themselves quantitative in parts. This is the case as they are focused on image comparison. Conversely, the quantitative method was also subjected to qualitative decision-making processes at various points in the analysis. This occurred in the conception of the algorithm, that is, in the conscious isolation of the image phenomenon from a complex context (experimental arrangement), and in the breaking down of visual phenomena as objects of investigation into formalisable elements (divide and conquer). Even in the adaptation of the data basis and/or the individual functions by, for example, changing the threshold values, the programmer made purely subjective distinctions in consultation with the qualitative question. And, last but not least, the evaluation of the generated results and the subsequent modification of the program in the sense of ‘translation’ of the image-study analyses also constitute a qualitative-subjective act. By working together, we were forced to reflect intensively on the inadequacy of our respective discipline-oriented views as producing insights that were either solely qualitative or solely quantitative. The group spent a lot of time working to achieve interdisciplinary understanding, which meant we had to explain and make transparent our procedure much more than in homogenous research contexts.

In addition to the process of actual image analysis, we extended our research to include various formats of cross-disciplinary exchange—theoretical as well as practical—by means of experimental formats, such as hackathons and workshops. We also promoted and opened up our research project to a wider audience and tried out various formats in addition to the ordinary research process, similar to labs. For example, we developed open-source tools and held workshops at digital methods institutions.<sup>20</sup>

What is typical for interdisciplinary teams was true for our team too. The members of the group did not work at the same institutes and thus needed to arrange meetings in locations across the campus or, as in our case, across three institutions.

## The status of machine learning

Machine learning as a contemporary trend of statistics-led computation was of specific interest to us. However, we were more interested in the actual mathematical models than in the notion of learning, which is predominantly used in superficial

---

20 See DMI, Climate image spaces and the new climate movements: A Digital Methods Summer school DMI19 project report.

discourses about machine learning that are less interested in its functioning and more in exploiting certain narratives.<sup>21</sup> The use of machine learning calls for a particular case. In our study, we were interested in the automation of a certain idea of image similarity and how it compared to the human understanding of similarity. This resonated with our research scope of testing and reflecting upon a variety of algorithmic approaches. In order to work, a machine-learning algorithm needs to be ‘trained’ or calibrated preferably with a very large dataset. The 16,000 images we extracted from the Google Images search are too few to be usable for a neural network. For resource reasons, we therefore decided to use a pre-trained neural network following Google’s Inception v3 architecture based on the image dataset from ImageNet with over 14 million images. This was a momentous decision considering that ImageNet incorporated taxonomies and categorisations that were not specified for the case of finding similarities in climate images.<sup>22</sup>

We found that machine learning was a challenging but productive method to study large image datasets. In particular, the combination of machine learning and data visualisation offers an alternative to graph-based layout methods. This combination also provides an alternative to purely keyword-based image research where clusters are created by language alone. Instead of looking at a single image artefact in detail or creating a distant view to read patterns of meta-data by using machine learning, we faced a level abstraction through automation. The resulting data dimensions were far beyond the image surfaces at stake in classical image studies and, therefore, a different form of addressing these image structures was needed. Another problem emerged: the choice of parameters determined the distribution of clusters and thus the content reading of the data visualisations. Ultimately, this approach was made productive for qualitative image analysis, as it was based on human interpretation.

## Human in the loop—the need for human interaction

When we look at visualisations of machine-learning results, such as the t-SNE visualisations in our international comparison of thousands of climate images, the following question arises: How are these images grouped?

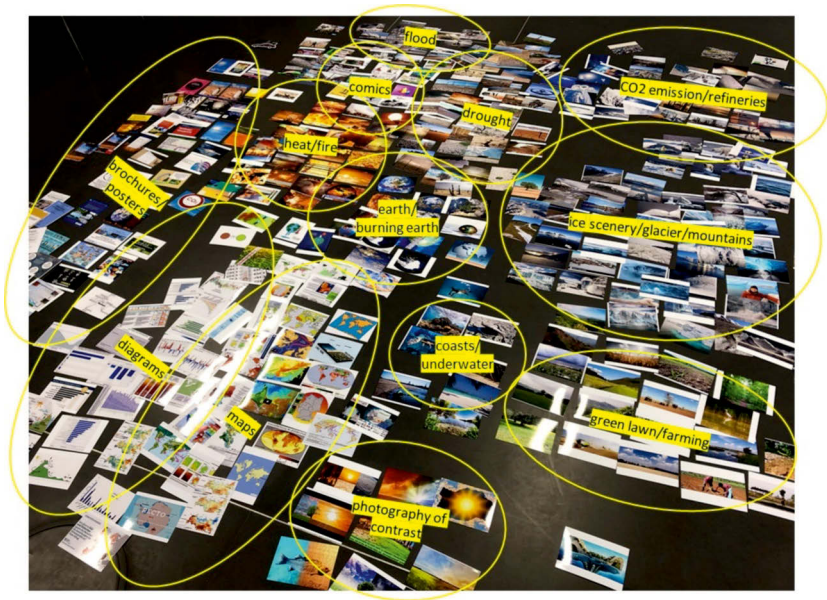
---

21 See Matteo Pasquinelli for similar thoughts. *How a Machine Learns and Fails: A Grammar of Error for Artificial Intelligence*, 2019, online: <https://kim.hfg-karlsruhe.de/how-a-machine-learns-and-fails>.

22 In ImageNet, certain types of images are hardly categorized, which is also relevant for our research. The underrepresentation of art is considered in this article: Francis Hunger, “Why so many windows?”—*Wie die Bilddatensammlung ImageNet die automatisierte Bildererkennung historischer Bilder beeinflusst*, 2021. Zenodo. DOI: <https://doi.org/10.5281/zenodo.4742621>.

There is no automated method for the recognition of the clusters. We had no other criterion to determine the similarity of the images to one another—other than what we interpreted visually. In decision-making situations, such as boundary-finding and manual clustering, what came into play were our eyes (trained to see climate images), our theoretical knowledge and our ability to interpret. However, the human gaze does not completely reject data visualisations. Surprisingly, human annotations and algorithmic clusters seem to converge to a certain degree in the more discrete groupings (Fig. 8).

*Fig. 8: Manual image sorting of the printed out Google search corpus (Germany) with annotations. Photo taken by anci, 2020.*



This is significant because the process of creating data visualisations from the source images does not extract semantic information from the dataset. We assume that human annotations also take into account the prominent features of the images, such as shape, colours, contrast and composition when people evaluate the visualisations generated by t-SNE; and they do not merely consider the symbolic meaning as it is not available to the algorithm.

In this respect, the term ‘clustering’ can prove to be somewhat confusing in methodological approaches. Indeed, it was observed that, in technical terms, the



determination of a t-SNE cluster did not necessarily result from a great number of images, but from a high similarity ratio—most noticeably with duplicate images.<sup>23</sup>

Thus, it can be stated on a technical level that the higher the similarity factor of the images, the more likely it is to be a technically generated image cluster. On the level of human perception, we can say that the higher the frequency of motifs with a common similarity criterion, the more likely it is to be a ‘group’ of images. (The qualitative determination of it results, in particular, only from a size comparison of all images from all country queries.) The human gaze is thus required for the determination of dominant image groups, as it provides a correction to the algorithmic image sorting. For this reason, we distinguished in our study between the term ‘cluster’ as a technically generated set of images and the qualitative term ‘group’ to emphasise image frequency.

## The status of data visualisation

In our research, visualisation as a concept and practice is not bound to the idea of depicting the results of the research but rather as an essential element of almost every aspect of our mixed-methods approach. We conceptualised the method of visualisation as a crucial interface for communicating between quantitative and qualitative approaches, that is, between different ideas of how to structure information. On another note, visualisations were used to provide digital data, which would otherwise be incomprehensible to humans, with a subjective form to make them accessible for a qualitative analysis. In our research as an image-based, mixed-methods project in particular, it gave us scope for examining the status of imagery not only as the object of our research but also as a research modus. One interesting gain in knowledge resulting from the data-based image analysis can be found in the ontological image viewing. Indeed, from the point of view of image theory, an interesting translation of image understanding became clear in the mixed use of methods: the digitality of the images enabled constant and iterative changes between them as 1) a phenomenon, as 2) a dataset, as 3) a visualised table and, finally, as 4) a statistical graph.

---

23 Adrian MacKenzie and Anna Munster clustered images of computer vision research using t-sne in their AI-research to critically reflect on statistical computer vision. See: “Oscilloscopes, slide-rules, and nematodes. Toward heterogenetic perception in/of AI”, in: Natasha Lushetich, Iain Campbell (eds.): *Distributed Perception. Resonances and Axiologies* (London: Routledge, 2021), DOI: 10.4324/9781003157021-6.

## Discussion

The very idea of the digital as something new and fundamentally transformative—as in the myth of the digital—runs the risk of missing the particularities of what it actually means to work with digital structures.<sup>24</sup> Our focus on digital imagery sharpened our understanding of the algorithmic view. From a qualitative perspective, we increasingly repositioned our research questions and opened them up time and again in the process. The focus on technical methods shifted our emphasis from an interpretative view to a ‘structural view’, which resulted from algorithmic image recognition. Increased attention to algorithmically detected structures within the image areas and to machine-generated structures of image sorting and clustering became apparent. The analysis of algorithmic cluster structures as the technical ‘preliminary work’ of image sorting in accordance with the key similarity criteria was also an important component. For instance, the perception of colour contrasts and the dense and heterogeneous distribution of large numbers of images again provided us with ideas and inspiration for new content-related efforts.

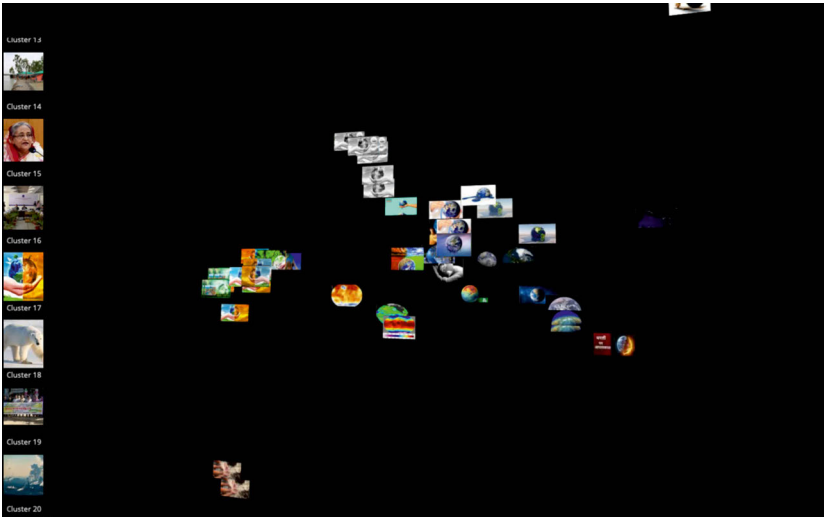
One particular conflict in our interdisciplinary cooperation came in the form of a shift of significance regarding the criterion of similarity (*tertium comparationis*) and the associated research questions. In cooperation with quantitative methods, the qualitatively determined similarity criterion shifted from its content on the phenomenological image level to the level of formalisable data and metadata, the analysis of which is dominated by principles of mathematical structural logic.

This resulted in a change between the image as a phenomenon and the image as a technical dataset. By translating a qualitative question into a regulatory system, certain optimisation problems arose. The focus of our problem resolution was thus entirely on the adaptation of the inner formal logic of the images and thus obscured the view of content-related questions. In other words, causal questions often became questions of quantity and distribution. Differences also emerged in the use of vocabulary regarding the meaning of the term ‘cluster’ in our intercultural t-SNE-visualisations. The German translations of the word ‘cluster’ as ‘Anhäufung’ (accumulation) or ‘Gruppe’ (group) simply do not fit the technical meaning. With regard to machine-learning approaches, we recommend the thorough use of language so as to provide clarity and avoid confusion; however, we do not recommend referring to the visual clusters in t-SNE as anything other than technically and conceptually feasible (Fig. 9).

---

24 See Jan Distelmeyer, *Das flexible Kino. Ästhetik und Dispositiv der DVD & Blu-ray / The Flexible Cinema. Aesthetics and Dispositif of the DVD & Blu-ray* (translated using DeepL) (Berlin: Bertz & Fischer, 2012).

Fig. 9: Close-up screenshot from a 'globe' cluster in the t-SNE visualisation.



We need abstraction to grasp details, but details also maintain the models for abstraction. Qualitative methods deal with conceptual abstractions and derive their insights from interpretations of particular structural details as framed by the notion of 'close reading'. Similarly, quantitative approaches look on empirical models and ideas which are then realised as mathematical rule-sets and conventions. In computer science, the paradigm of 'divide and conquer', for example, describes a systemic idea of categorising problems into partial elements.<sup>25</sup> In terms of details, quantitative approaches, especially computer vision and machine-learning algorithms, are good and, if trained properly, prove better than humans at detecting patterns in (media) environments. Therefore, it did not seem like a conflict but rather like an undeniable and more complex relation within the methods themselves. Details can lead to oversimplified assumptions and abstractions can also embrace increased complexities. It is human cognition which always implies reduction since we rely on certain models and world views to adapt to changing realities. From here we can move to a reading of reductionism as a scientific search for causal-deterministic relationships. Our experience gained from a mixed-methods design showed us that a holistic perspective was more appropriate when trans-disciplinary problems were encountered, such as climate-change communication. Overall, the discussion around the idea of reduction was very important to our research, but it could not be properly framed with the two notions of detail and abstraction.

25 See [https://en.wikipedia.org/wiki/Divide-and-conquer\\_algorithm](https://en.wikipedia.org/wiki/Divide-and-conquer_algorithm).

Machine processes of image recognition enable new approaches to sorting images according to their formal-stylistic content and structures. The media- and data-critical analysis of the image recognition process of algorithms is decisive in this process. Art-historical classifications in accordance with the style of the images based on the connoisseur's gaze are supplemented by algorithmic glances and technical sorting logics. Both a (phenomenal) image analysis, such as Erwin Panofsky's iconographic-iconological three-phase model of visual levels of meaning, and an algorithmic analysis of the various dataset formats lead to the notion that pictoriality must be understood as the interplay of various image levels.<sup>26</sup> Above all, algorithmic image-sorting helps to manage the sheer mass of images and to find frames. The potential of mixed-methods image analysis can thus be found in its ability to open up a new possibility of analysing images. A more in-depth examination of technical and technological developments, such as machine learning, will lead to a more productive sharpening of the 'statistical gaze'. It does not mean that these technologies of algorithmic image sorting and analysis effectively extend or improve traditional image viewing; they do, however, enable other views that extend an interpretative perception based on conventions with analytical views that result from the media-critical analysis of algorithmic codes.

Uncertainty is a constitutive element of our quantitative and qualitative methods. In contrast to computational models, human models can handle the ambiguous and the undefined. Humanities scholars are aware of the danger, for example, of 'linguistic heteronomy' ('sprachliche Fremdbestimmung')<sup>27</sup> and of individual over-interpretation. The full impact of qualitative considerations reveals itself in the blurry realm of uncertainty. Quantitative approaches also require uncertainty as an essential condition. It tests their limits, thereby making an essential contribution to their productive re-iteration. Dealing with uncertainty is, therefore, indicative of essential mixed-methods ability.

Speaking from a general perspective, the discussion surrounding the success and failure of data-driven research will intensify in such a way that a still higher level of data positivism and "overestimation of statistical significance"<sup>28</sup> will find them-

---

26 See Harald Klinke, "Bildwissenschaft ohne Bildbegriff" / "Image science without image term" (translated using DeepL), in: Harald Klinke, Lars Stamm (eds.): *Bilder der Gegenwart: Aspekte und Perspektiven des digitalen Wandels* (Göttingen: Graphentis, 2013), 11–33.

27 Gottfried Boehm, "Ein Briefwechsel" / "An Exchange of Letters" (translated using DeepL), in: Marius Rimmel et al. (eds.): *Bildwissenschaft und Visual Culture* (München: Wilhelm Fink, 2014), 26.

28 Hanna Brinkmann; Laura Commare, "Why 'Anything goes' der Goldstandard sein sollte – Überlegungen zu Methodentradition und empirischen Forschungsansätzen in den Kunstwissenschaften" / "Why 'Anything Goes' Should Be the Gold Standard – Reflections on Methodological Tradition and Empirical Research Approaches in Art Science" (translated using DeepL), in: Verband österreichischer Kunsthistorikerinnen und Kunsthistoriker (VöKK) (ed.): *Newest*

selves reflected in an increasingly critical manner. With regard to the structure of interdisciplinary DH projects, we assume and hope that there will be an increasing balance of esteem between the humanities and computer-aided disciplines. Compared to the traditional humanities, there is still a focus of interest on computer science as the sole source of technical products and tools.

Nonetheless, the development of qualitative research questions and content is a crucial basis for the design of digital methods and tools. For precisely this reason, a balance should be struck if both approaches are to be considered on an equal footing. Furthermore, future DH endeavours should not be seen as completed, closed-circuit projects that run the risk of ending up in ‘DH graveyards’. The goal must be to continuously integrate both into general research and university teaching—the findings of digital imaging and the method-critical approaches from the interdisciplinary interlocking.

## Bibliography

- Paul Heinicker, Janna Kienbaum, Birgit Schneider, Thomas Nocke: *anci website: htps://anci.fh-potsdam.de/*, 2021.
- Boehm, Gottfried, and W.J.T. Mitchell. “Ein Briefwechsel”. In *Bildwissenschaft und Visual Culture*, edited by Marius Rimmele et al, 19–40. München: Wilhelm Fink, 2014.
- Brinkmann, Hanna, and Laura Commare: “Why ‘Anything goes’ der Goldstandard sein sollte – Überlegungen zu Methodentradiation und empirischen Forschungsansätzen in den Kunstwissenschaften”. In *Newest Art History – Wohin geht die jüngste Kunstgeschichte? Tagungsband zur 18. Tagung des Verbandes österreichischer Kunsthistorikerinnen und Kunsthistoriker*, edited by Verband österreichischer Kunsthistorikerinnen und Kunsthistoriker (VöKK), 169–172, 2017. [https://voekk.at/sites/default/files/downloads/tagungsbaende/Newest%20Art%20History\\_VoeKK-Tagungsband.pdf](https://voekk.at/sites/default/files/downloads/tagungsbaende/Newest%20Art%20History_VoeKK-Tagungsband.pdf).
- Diers, Michael. *Schlagbilder: Zur politischen Ikonographie der Gegenwart*. Frankfurt a. M.: Fischer, 1997.
- Dimitrova, Daniela V., and Lulu Rodriguez, Lulu. “The levels of visual framing”. In *Journal of Visual Literacy*, 30:1 (2013), 48–65, DOI: 10.1080/23796529.2011.11674684.
- Dilly, Heinrich. “Lichtbildprojektion – Prothese der Kunstbetrachtung”. In *Kunstwissenschaftliche Untersuchungen des Ulmer Vereins, Verband für Kunst- und Kulturwissenschaften*, edited by Horst Bredekamp et al., 153–172. Gießen: Anabas, 1975.
- Distelmeyer, Jan. *Das flexible Kino. Ästhetik und Dispositiv der DVD & Blu-ray*. Berlin: Bertz & Fischer, 2012.

- Christ, Katharina et al: DMI, Climate image spaces and the new climate movements: A Digital Methods Summer school DMI19 project report. <https://wiki.digitalmethods.net/Dmi/ClimateImageSpaces>, 2019.
- Entman, Robert. "Framing: Toward Clarification of a Fractured Paradigm". In *Journal of Communication* 43, 4, 1993. [https://is.muni.cz/el/1423/podzim2018/PO\\_L256/um/Entman\\_1993\\_FramingTowardclarificationOfAFracturedParadigm.pdf](https://is.muni.cz/el/1423/podzim2018/PO_L256/um/Entman_1993_FramingTowardclarificationOfAFracturedParadigm.pdf)
- Geise, Stephani, and Katharina Lobinger (eds.). *Visual Framing. Perspektiven und Herausforderungen der Visuellen Kommunikationsforschung*, Köln: Halem, 2013.
- Grittmann, Elke. "Visual Frames – Framing Visuals. Zum Zusammenhang von Diskurs, Frame und Bild in den Medien am Beispiel des Klimawandeldiskurses". In *Visual Framing: Perspektiven und Herausforderungen der visuellen Kommunikationsforschung*, edited by Stephanie Geise and Katharina Lobinger, 95–116. Köln: Halem, 2015.
- Hunger, Francis. "Why so many windows?" – *Wie die Bilddatensammlung ImageNet die automatisierte Bilderkennung historischer Bilder beeinflusst*, 2021. Zenodo. DOI: <https://doi.org/10.5281/zenodo.4742621>.
- Klinke, Harald. "Bildwissenschaft ohne Bildbegriff". In *Bilder der Gegenwart: Aspekte und Perspektiven des digitalen Wandels*, edited by Harald Klinke and Lars Stamm, 11–33. Göttingen: Graphentis, 2013.
- Latour, Bruno. *Science in Action. How to Follow Scientists and Engineers Through Society*, Cambridge: Harvard University Press, 1987.
- Link, Jürgen. "Literaturanalyse als Interdiskursanalyse: Am Beispiel des Ursprungs literarischer Symbolik in der Kollektivsymbolik". In *Methoden diskursanalytischer Ansätze*, edited by Jürgen Fohrmann and Harro Müller, 284–307. Frankfurt am Main: Suhrkamp, 1988.
- MacKenzie, Adrian, and Anna Munster. "Oscilloscopes, slide-rules, and nematodes. Toward heterogenous perception in/of AI". In *Distributed Perception. Resonances and Axiologies*, edited by Natasha Lushetich and Iain Campbell, 64–81. London: Routledge, 2021, DOI: 10.4324/9781003157021-6.
- MacQueen, James B. "Some methods for classification and analysis of multivariate observations". In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 1967, 281–297.
- Marr, David. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Cambridge, Mass.: MIT Press, 2010 [1982].
- Matthes, Jörg. *Framing*, Baden-Baden: Nomos, 2014.
- Panofsky, Erwin. *Sinn und Deutung in der bildenden Kunst* [1955]. Köln: DuMont, 1978.
- Pasquinelli, Matteo: How a Machine Learns and Fails: A Grammar of Error for Artificial Intelligence, 2019, online: <https://kim.hfg-karlsruhe.de/how-a-machine-learns-and-fails>.

- Schneider, Birgit. *Klimabilder: Eine Genealogie globaler Bildpolitiken von Klima und Klimawandel*. Berlin: Matthes & Seitz, 2018.
- Statista. "Number of explicit core search queries powered by search engines in the United States as of April 2021". Accessed July 19, 2021. <https://www.statista.com/statistics/265796/us-search-engines-ranked-by-number-of-core-searches/>.
- Sullivan, Danny. "Dear Bing, We Have 10,000 Ranking Signals to Your 1,000. Love, Google". In *Search Engine Land* Accessed July 19, 2021. <https://searchengineland.com/bing-10000-ranking-signals-google-55473>.
- Thürlemann, Felix. "Bild gegen Bild: Für eine Theorie des Vergleichenden Sehens". In *Zwischen Literatur und Anthropologie – Diskurse, Medien, Performanzen*, edited by Aleida Assmann, Ulrich Gaier and Gisela Trommsdorff, 163–174. Tübingen: Narr, 2005.
- Hinton, Geoffrey and Laurens Van der Maaten. *Visualizing Data using t-SNE*, 2008, accessible at <https://jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>
- Walsh, Lynda. "'Tricks', Hockey Sticks, and the Myth of Natural Inscription: How the Visual Rhetoric of Climategate Conflated Climate with Character". In *Image Politics of Climate Change Visualizations, Imaginations, Documentations*, edited by Thomas Nocke and Birgit Schneider, 81–105. Bielefeld: Transcript, 2014.
- Walsh, Lynda. "The Visual Rhetoric of Climate Change". In: *Wires Climate Change*, Volume 6, 4, (2015), 361–368.
- Warburg, Aby. *Der Bilderatlas Mnemosyne*. Berlin: Akademie Verlag, 2008.
- Warburg, Aby. "Einleitung zum Mnemosyne-Bildatlas [1929]". In: *Kulturwissenschaft. Eine Auswahl grundlegender Texte*, edited by Uwe Wirth, 137–145. Frankfurt am Main: Suhrkamp, 2008.
- Wikipedia: Divide-and-conquer-algorithm. Accessed December 2022 at [https://en.wikipedia.org/wiki/Divide-and-conquer\\_algorithm](https://en.wikipedia.org/wiki/Divide-and-conquer_algorithm).
- Yale DH Lab. "PixPlot". Accessed July 19, 2021. <https://dhlab.yale.edu/projects/pixplot/>.

