The Potentials *of* Google Vision ᴀᴘɪ-based Networks *to* Study *Natively* Digital Images

Janna Joceli Omena
Center for Advanced Internet Studies (ᴄᴀɪs)

Elena Pilipets
Universität Klagenfurt

Beatrice Gobbo
Politecnico di Milano

Jason Chao
Universität Siegen

ABSTRACT

In this article, we present the potentials of Google Vision API-based networks for studying online images, covering three important modalities as part of a critical visual methodology: the content of the image itself, its specific 'audiencing' through web references (or image metadata), and the sites of image circulation. First, we conceptually and technically define different networks built upon computer vision features: *image-label, image-web entities*, and *image-domain*. Second, we present a research protocol diagram that illustrates how to build networks of images and respective descriptions or sites of circulation. Third, we discuss the potentialities of computer vision networks as a research device, stressing their data-relational (trans)formations and interpretative specifics. Three different case studies will be introduced as examples. In conclusion, we argue that such a visual methodology requires critical technical practices accounting for the multiple layers of technical mediation involved.

Keywords
Image analysis
Networks
Digital methods
Visual methodologies
Computer vision

AUTHORS

Janna Joceli Omena—Master in Contemporary Culture and New Technologies, Universidade NOVA de Lisboa. Research fellow at the Center for Advanced Internet Studies (CAIS). She is a member of iNOVA Media Lab and the Public Data Lab. Her research focuses on digital methods, digital network studies, and technicity-of-the-mediums in support of social and medium research. She is the editor of *Métodos Digitais: Teoria-Prática-Crítica* (ICNOVA, 2019) and the coordinator of the SMART Data Sprint. Some of her latest publications are: 'Digital Methods for Hashtag Engagement Research' (with E.T. Rabello and A.G. Mintz; *Social Media + Society*, Vol. 6, Issue 3) and 'Call into the Platform!' (with A. Granado; *Icono14*, Vol. 18, Issue 1).

Elena Pilipets—Doctor in Media Studies, Universität Klagenfurt. Postdoc researcher at the Department of Media and Communications, Universität Klagenfurt, and SMART (Social Media Research Techniques) researcher with iNOVA Media Lab, Universidade NOVA de Lisboa. Her teaching and research interests are related to media cultural studies, internet research, and digital methods. She is the current holder of the 2020 Carinthian Award for Young Social Sciences and Humanities Scholars. Her recent publications include 'Nipples, Memes, and Algorithmic Failure: NSFW Critique of Tumblr Censorship' (with S. Paasonen; *New Media & Society*, 1st Publ., 2020) and 'Digitale Medien und Methoden. Über den methodischen Umgang mit visuellen Plattforminhalten und Internet-Memes' (*Open-Media-Studies-Blog der Zeitschrift für Medienwissenschaft*).

Beatrice Gobbo—Master in Communication Design, Politecnico di Milano. Ph.D. student in Design at Politecnico di Milano. She is a member of the Density-Design Lab, a research group focused on information visualization and information design. Her current research is focused on the role of communication design and information visualization in the field of explainable artificial intelligence. Some of her recent publications include 'Research Protocol Diagrams as Didactic Tools to Act Critically in Dataset Design Processes' (with M. Mauri, M. A. Briones, and G. Colombo; in *INTED2020 Proceedings*) and 'Explaining AI through Critical Reflection Artifacts. On the Role of Communication Design within XAI' (in T. Reis, M.X. Bornschlegl, M. Angelini, and M.L. Hemmje, Eds.; *Advanced Visual Interfaces. Supporting Artificial Intelligence and Big Data Applications*; Springer, 2020).

Jason Chao—Master in Big Data and Digital Futures, University of Warwick. Master in Human Rights Law, University of London. He is a researcher and PhD candidate at the Universität Siegen. His current research is focused on the development of digital methods tools for the study of sensor media. He has backgrounds in software development and human rights advocacy. Some research software recently developed by him include: *AppTraffic*, to study the network traffic of mobile applications; and *Offline Image Query and Extraction Tool* (with J. J. Omena).

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

The Potentials of Google Vision API-based Networks to Study *Natively* Digital Images

Janna Joceli Omena
Center for Advanced Internet Studies (CAIS)
Bochum, Germany
janna.joceli.omena@cais.nrw

Elena Pilipets
Universität Klagenfurt
Department of Media and Communications
Klagenfurt, Austria
Elena.Pilipets@aau.at

Beatrice Gobbo
Politecnico di Milano
DensityDesign Lab
Milan, Italy
beatrice.gobbo@polimi.it

Jason Chao
Universität Siegen
Collaborative Research Centre "Media of Cooperation"
Siegen, Germany
teng.chao@uni-siegen.de

INTRODUCTION

Discussing the potentials of digital methods for studying visual material, Gillian Rose (2016) highlights the embeddedness of images within digital infrastructures of image 'audiencing'. Machines form part of this audience (Hoelzl & Marie, 2015; Paglen, 2014; Rettberg, 2020). Organizing images in replicable, searchable, and spreadable formations, natively digital interactions present both an iterative process and a distributed accomplishment (Hochman, 2014; Steyerl, 2009). The image no longer operates solely as an iconic object of representation but becomes reconfigured as a calculable surface within an ensemble of machine-learning techniques, ranking cultures, recommendation algorithms, and platform environments through which it is operationalized (Azar et al., 2021; MacKenzie & Munster, 2019; Rubinstein & Sluis, 2013). The dynamics that render images analyzable are site-specific and informed by various digital records (e.g., hashtags, timestamps, links) (Gerlitz & Rieder, 2018). Consequently, natively digital images are always networked through platform-specific and web-native data, producing enormous amounts of visual, textual, and numerical information 'at scale' (Manovich, 2020; Parikka & Dvořák, 2021).

This article enters the study of visual material using digital methods (Rogers, 2013), a networked approach to visual research (Colombo, 2018; Niederer & Colombo, 2019) and computer vision. So far, methods combining automated visual analysis with platform metadata have demonstrated the potentials of quali-quantitative approaches for understanding the interplay between the

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

narrative patterns of images and techniques of content distribution within and across different platform cultures of use (Niederer & Colombo, 2019). Examples include semantic mapping of cities (Ricci et al., 2017), studies of image circulation (d'Andréa & Mintz, 2019), institutional communication (Omena & Granado, 2020), and affective affordances of hashtag publics (Geboers & Van De Wiele, 2020). The multiple features that can be automatically extracted from an image through computer vision —such as objects, faces, backgrounds— raise new questions of interpretability along with the need to shift between different scales of observation, reference, and comparison (Manovich, 2020).

In this article, we further develop this analytical trajectory by repurposing Google Vision features as part of a critical visual methodology.[1] We discuss (1) the content of the image itself, (2) its specific 'audiencing' through references obtained from the web, and (3) the sites of image circulation (Rose, 2016). To this end, we conceptually and technically define three different networks: The *image-label* network enables the study of content and visual composition through literal annotations of the image itself; the *image-web entity* network positions the images in relation to web content, using Google's ranking systems and Knowledge Graph as a source to deliver web annotations; the *image-domain* network groups images in connection to web pages or uniform resource locators (URLs) that host fully (or partially) matching images. While acknowledging uncertainty in defining *how* Google Vision API assigns its different values, we discuss computer vision networks in their potential for situating natively digital images through their data-relational (trans)formations.

CHARACTERIZING COMPUTER VISION NETWORKS

We refer to *computer vision networks* as a proposal inspired by the work of Donato Ricci, Gabriele Colombo, Axel Meunier, and Agata Brilli (2017) and further developed by Colombo (2018). Mapping the symbolic and material elements of the Parisian urban nature debate, the authors merged a collection of Twitter images linked by the keyword 'nature' into a network using a web-based vision API (IMAGGA'S API). This work has inspired other scholars (Geboers & Van De Wiele, 2020; Mintz et al., 2019; Omena et al., 2020; Silva et al., 2018) in exploring a particular characteristic of computer vision: the labeling of visual content according to predefined tags/labels provided by machine learning models. This feature converges into what now is referred to as *image-label* networks, while other features are still understudied, such as detecting web entities or web pages with full matching images that respectively afford what we are calling computer vision *image-web entities* and *image-domain* networks. These latter introduce novel articulations for the interpretation of online images, including their potential value for the studies of image circulation and image contextualization. In the following pages, we

[1] Gillian Rose (2016) proposes four modalities for interpreting visual materials: sites of production, the image itself, its circulation, and its 'audiencing'.

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

conceptually and technically inform what computer vision networks are and what they are made of while introducing an open-source research software created to support Google Vision API.

What networks are we talking about?

We refer to networks as a space of connections captured *in* and *by* digital records (see Venturini et al., 2015, 2019), yet responsive to "the technological processes inherent to the web environment and APIs *in which* and *through which* online communication are structured, captured and merged with other records" (Omena, in press; see Gerlitz & Rieder, 2018). This latter situation refers to platform *grammatization,* which more specifically in this regard, reflects how digitally native images *participate* in and *are mediated by* technological infrastructures, while embedded into cultures of use. It is crucial to distinguish between networks directly granted by data collection methods (using APIs, scraping, or crawling) and those built on top of web data and the use of a range of software (see Omena & Amaral, 2019). The former is *designed via ready-made* files whereas the latter is *constructed*. For example, when directly granted by APIs, the studied network results from a good query design and the use of a web-based application for extracting data (e.g., YouTube Data Tools) from an API (e.g., YouTube Data API). The output file (e.g., graph dataset format, short for .GDF) feeds software (e.g., Gephi) to visualize and analyze the network (e.g., YouTube networks of channels' connections). However, *computer vision networks* must be built; therefore, a longer list of practical tasks is required since data collection tools are unable to deliver *ready-made* files of image networks that one can easily download and visualize. The networks we are talking about are made of natively digital images that are reconstructed through computer vision and by the choices that the researcher has to make along the process of curating, building, visualizing, and analyzing the network itself. We then understand *computer vision networks* as an ensemble of computational mediums, data, methods, research, and technical practices orchestrated by the researcher(s). In one way or another these networks demand some mastery of technical knowledge and practices (imposing challenges from the methodological standpoint) while they introduce new forms and formats of understanding society and our everyday actions (see Venturini et al., 2019).

Understanding computer vision features

Computer vision is the computer's capacity to recognize visual features through algorithmic techniques, using these learnings to identify and classify objects and scenes (Szeliski, 2021). The computer vision community creates, trains, and tests machine learning models and neural networks to classify images, but the use of this technology is diverse and accommodates different purposes. To the market-

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

place or government security services, it offers services such as content moderation (e.g., by recognizing offensive or unwanted images or detecting racist and adult content), predictive analytics, or the controversial ability of face recognition. In social research, computer vision has been useful to diagnose gender bias and for the studies of visual misinformation or political ideology of images (see Garimella & Eckles, 2020; Schwemmer et al., 2020; Xi et al., 2020). In this paper, we introduce the potentials of Google Vision API-based networks to study natively digital images. Next, we introduce Google Vision features in more detail, justifying the API's usage and clarifying the role of networks as a research device.

The API was launched in May 2017, offering several features of which we suggest using *label detection* and *web entities* and *pages detection*. In general, *labels* and *web entities* provide textual descriptions to images, while *web pages* describe where images are found. In detail, *label detection* provides both general and specific labels with confidence score and topicality rating afforded by machine learning models. Its level of specificity has proven to be the most detailed, when it comes to image classification, and in comparison to other APIs, namely IBM Watson and Microsoft Azure (Mintz et al., 2019; Silva et al., 2020). Google Vision also possesses features not found in other computer vision APIs: the detection of *web entities* and *pages*. This provides web content to an image, using the power of Google Image Search to "detect topical entities such as news, events, or celebrities within the image" (Google Cloud, 2017) and to find identical or similar images on a list of URLs or web pages.[2] Page URL, page title, matching image URL are sources that supply web pages with full or partial matching images, while Google's Knowledge Graph provides the information about or what defines a *web entity* (a person, a place, or a thing). The Knowledge Graph is "a system that understands facts and information about entities from materials shared across the web, as well as from open source and licensed databases", gathering over 500 billion facts about five billion entities (Sullivan, 2020). This tells us that web entities provide *references to an image obtained from the web environment*, which goes beyond the content of the image itself. We assume that the detection of web entities and pages may include Google's ranking systems, license data, public sources, and factual information received directly from content owners.

How can researchers access Google Vision features and practically use them? Over the years, Script files have been a crucial technical element to invoke computer vision APIs. These tools impose some technical challenges for social researchers, particularly for non-developers or coders. Against this background, we introduce Memespector Graphical User Interface (GUI) (Chao, 2021), a research software tool developed by one of the authors, which serves well the study of large image collections through the lens of computer vision. Memespector GUI builds on Bernhard Rieder's original *Memespector* project (later expanded to

2 Our experience in finding little explanatory documentation about Google Vision reflects a problem already flagged by Timnit Gebru and colleagues: the lack of proper documentation about data provenance and the creation and use of datasets in the machine learning community (Gebru et al., 2020).

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

a Python script by André Mintz), by lowering the technical barriers to researchers *easily invoking* Google Vision API (as we will demonstrate later).

Combining Google Vision features and images as networks

In this section, we explain how Google Vision features are rearranged together with images as networks, presenting their potentiality for visual methodologies. Figure 1 shows a picture of the mosquito (*Aedes aegypti*) that transmits the Zika virus and the outputs of Google Vision to detect *labels*, *web entities*, and *pages*. When looking at the image containing the mosquito, the *label* feature renders back a literal description such as 'insect' and 'invertebrate', but also less reliable attempts like 'orb weaver spider' and 'organism'. While *labels* do not allow users to go beyond the content that can be read through images, *web entities*, instead, are capable of describing the same image by recognizing it, for instance, as a 'mosquito-to-borne disease, the Zika virus'. These *web entities* are responsive to data pulled from pages where the mosquito image was found (Robinson, 2017) but also to the existing entities of Google Knowledge Graph. Therefore, web entities are textual descriptions rooted in web content and provided by the lens of Google Image Search's ranking systems (Google Cloud, 2017; Robinson, 2017). Image description gains different meanings depending on what technological grammar is used.

Labels are always accompanied by confidence score and ranked by topicality rating, informing the probability of the textual descriptions assigned to an image. As we see in Figure 1, the description of the mosquito image starts with generic labels arranged by topics that contain the highest confidence score, e.g., 'insect' (0.9698934) and 'mosquito' (1,9772), moving to detailed classification such as 'arthropod' (0.800119) and 'infection' (0.4909). The factual or more contextualized descriptions provided by *web entities* allow different readings of the same visual content and can complement each other.

Now, we look at the same image but paying attention to its sites of appearance by using *web page detection*, which provides a list of web pages or URLs in which fully (or partially) matched images are found. The feature offers a dynamic perspective on how images flow across platforms (URL) and which actors —link (sub) domains— engage with such visuality. For instance, the mosquito image has appeared in scientific journals (collections.plos.org), health clinics (tratamentoendometriose.com.br), news media (correiodoestado.com.br), social media and blogs (i.pinimg.com; 1.bp.blogspot.com).

In this understanding, we now present three types of networks built upon the outputs of Google Vision API for the detection of labels, web entities, and pages, corresponding to (1) *image-label*, (2) *image-web entities*, and (3) *image-domain*. These networks have two types of nodes (bipartite graphs): one being always the image, whereas the other stands for the vision API chosen feature.
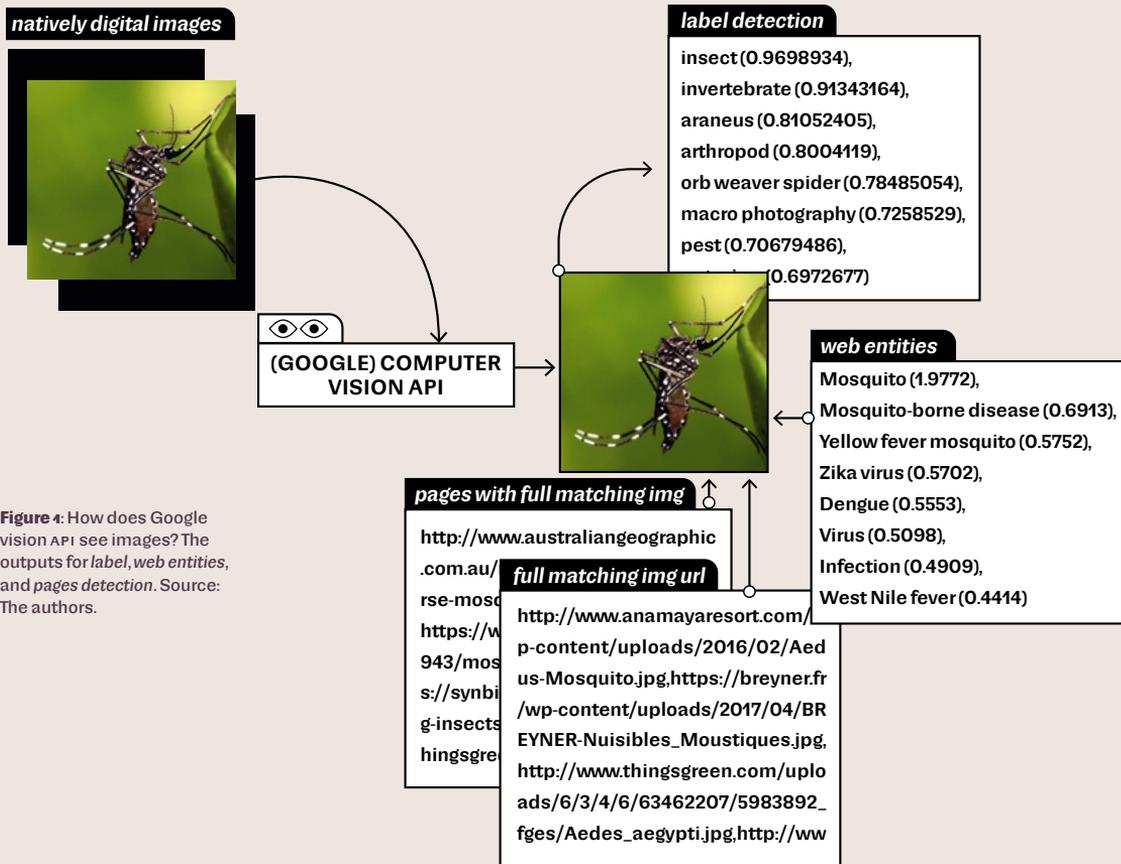
DISEÑA 19
AUG 2021
ARTICLE.4

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

natively digital images

label detection

insect (0.9698934),
invertebrate (0.91343164),
araneus (0.81052405),
arthropod (0.8004119),
orb weaver spider (0.78485054),
macro photography (0.7258529),
pest (0.70679486),
(0.6972677)

(GOOGLE) COMPUTER VISION API

web entities

Mosquito (1.9772),
Mosquito-borne disease (0.6913),
Yellow fever mosquito (0.5752),
Zika virus (0.5702),
Dengue (0.5553),
Virus (0.5098),
Infection (0.4909),
West Nile fever (0.4414)

pages with full matching img

http://www.australiangeographic
.com.au/
rse-mosc
https://w
943/mos
s://synbi
g-insects
hingsgre

full matching img url

http://www.anamayaresort.com/
p-content/uploads/2016/02/Aed
us-Mosquito.jpg,https://breyner.fr
/wp-content/uploads/2017/04/BR
EYNER-Nuisibles_Moustiques.jpg,
http://www.thingsgreen.com/uplo
ads/6/3/4/6/63462207/5983892_
fges/Aedes_aegypti.jpg,http://ww

Figure 4: How does Google vision API see images? The outputs for *label*, *web entities*, and *pages detection*. Source: The authors.

1. Networks of images and labels (image classification based on predefined/custom labels). Nodes are images and labels. Connections mean the occurrence of labels in relation to images.

2. Networks of images and web entities (image classification based on web content). Nodes are images and web entities. Connections mean the occurrence of web entities in relation to images.

3. Networks of images and sites of circulation across the web (referring to URLs or web pages in which images are found). Nodes are images and (sub) domains. Connections mean the occurrence of (sub) domains in relation to images.

Figure 2 illustrates how the three networks may look like (without the images), depicting the same collection of images (related to #microcephaly) that are interpreted differently according to Google Vision API technological grammar. A *static perspective* of the visual content follows the descriptive capacities of *labels* and *web entities*, while *web page detection* implies a *dynamic perspective* following the sites of image circulation. When reading a network of image descriptions,

DISEÑA 19
AUG 2021
ARTICLE.4

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

Figure 2: Three types of computer vision networks, three different perspectives for studying online images. Nodes are images (gray) and Google Vision features (orange). Source: The authors.

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

general labels followed by more specific ones drive the position of the images. The spatialization of the network relates to what is *in* the image (label detection) or how the image is described and contextualized *within* the web environment and *through* Google's lens (web entities detection). In the network, the position of the images highlights their literal or contextual descriptions unrelated to the idea of movement. In a network of image circulation (web page detection), node position is driven by the sites of image appearance; consequently, the interpretation of the images is not static but dynamic. These perspectives help the practice of reading computer vision networks, as we will demonstrate in the case studies.

THE RESEARCH PROTOCOL DIAGRAM

This section displays the methodological process for creating computer vision networks, emphasizing the transformations of images throughout this procedure. This digital methods approach requires researchers' engagement with technical practices (design and implement research with these methods), inviting researchers to be acquainted with computational mediums in conceptual, technical, and empirical terms (Omena, in press). This includes a critical look at the platform hosting the analyzed images and an understanding of the chosen computer vision API and research software as mediums of "expressing a will and a means to know" (Rieder & Röhle, 2018, p. 123).

The protocol diagram (Figure 3) exposes what it takes to build and interpret a computer vision network, considering its technical content that reflects situated online activity, *thick* layers of technical mediation, and researchers' interventions. The protocol is meant to be a self-explanatory and self-sustainable visual tool highlighting both the dataset design and the visualization process (Mauri et al., 2020; Niederer & Colombo, 2019), also providing a dissemination artifact that allows replicating the same method for building other networks upon the features of different computer vision APIs. We call attention (see Figure 4) to the role of the researcher(s) whom, as we argue, should account for the content of the technical mediums, while engaging with technical practices. The five steps of the protocol will be introduced below.

The process starts with *the art of querying* (Figure 4), which refers to the researcher's ability to make research questions through search queries (Rogers, 2019; Taibi et al., 2016), accounting for the formulation of specified (e.g., #femalepresentingnipples) or underspecified (e.g., #microcefalia; climate emergency) queries to curate a list of image URLs as a point of departure. The choice of keywords matters as much as the knowledge required when using extraction software or a Python script to collect images from digital platforms. Researchers should recognize that these tools cannot perform alone, but are instead conditioned by their choices, decisions, and technical knowledge.

Figure 3: The methodological process for creating computer vision-based networks to study online images. Source: Adapted from Omena (in press), designed by Beatrice Gobbo.
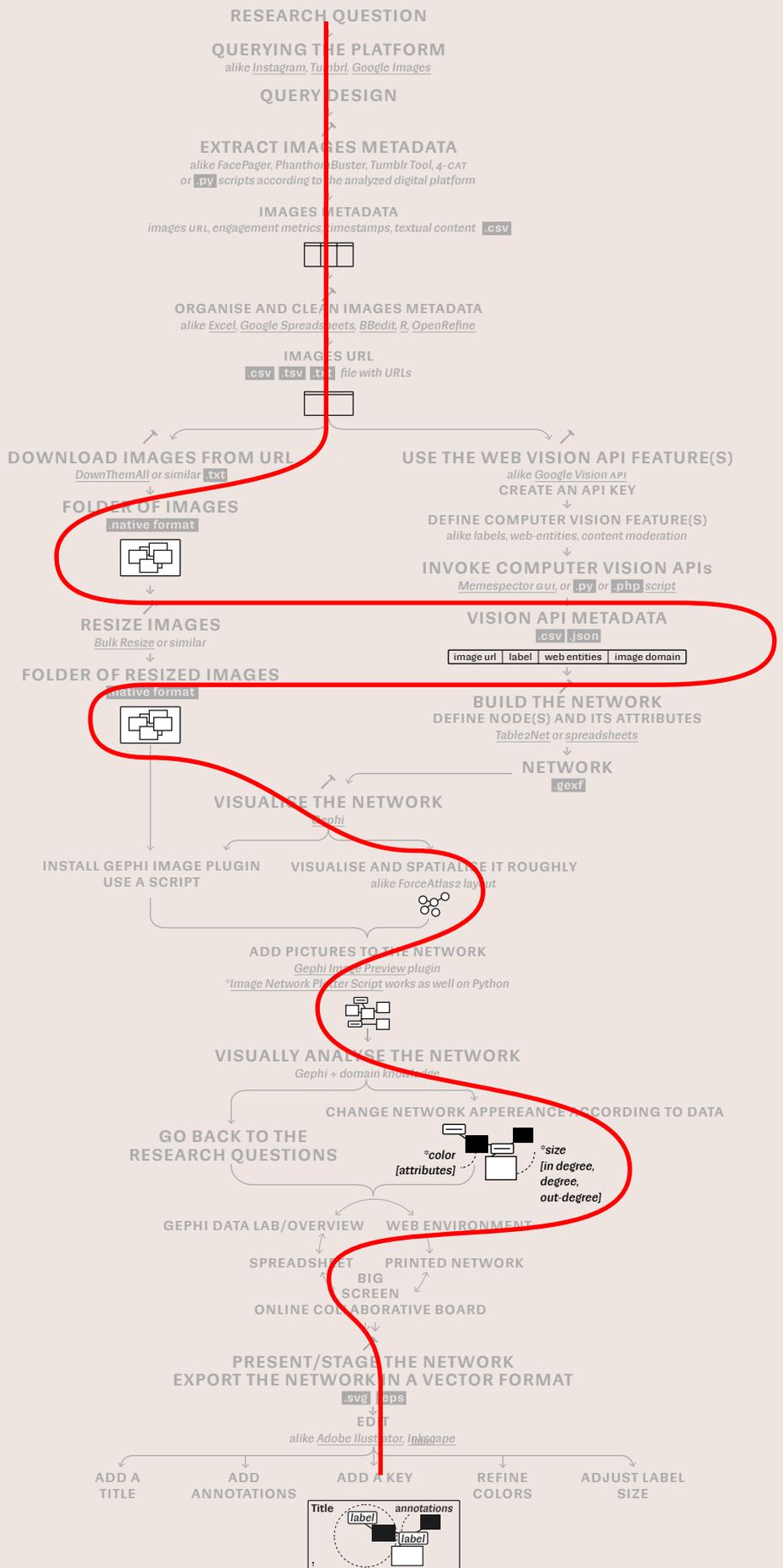
RESEARCH QUESTION

QUERYING THE PLATFORM
alike Instagram, Tumblr, Google Images

QUERY DESIGN

EXTRACT IMAGES METADATA
alike FacePager, Phanthom Buster, Tumblr Tool, 4-CAT
or `.py` *scripts according to the analyzed digital platform*

IMAGES METADATA
images URL, engagement metrics, timestamps, textual content `.csv`

ORGANISE AND CLEAN IMAGES METADATA
alike Excel, Google Spreadsheets, BBedit, R, OpenRefine

IMAGES URL
`.csv` `.tsv` `.txt` *file with URLs*

DOWNLOAD IMAGES FROM URL
DownThemAll or similar `.txt`

FOLDER OF IMAGES
`.native format`

RESIZE IMAGES
Bulk Resize or similar

FOLDER OF RESIZED IMAGES
`.native format`

USE THE WEB VISION API FEATURE(S)
alike Google Vision API
CREATE AN API KEY

DEFINE COMPUTER VISION FEATURE(S)
alike labels, web-entities, content moderation

INVOKE COMPUTER VISION APIs
Memespector GUI, or `.py` *or* `.php` *script*

VISION API METADATA
`.csv` `json`

image url	label	web entities	image domain

BUILD THE NETWORK
DEFINE NODE(S) AND ITS ATTRIBUTES
Table2Net or spreadsheets

NETWORK
`.gexf`

VISUALISE THE NETWORK
Gephi

INSTALL GEPHI IMAGE PLUGIN
USE A SCRIPT

VISUALISE AND SPATIALISE IT ROUGHLY
alike ForceAtlas2 layout

ADD PICTURES TO THE NETWORK
Gephi Image Preview plugin
*Image Network Plotter Script works as well on Python

VISUALLY ANALYSE THE NETWORK
Gephi + domain knowledge

CHANGE NETWORK APPEREANCE ACCORDING TO DATA

GO BACK TO THE RESEARCH QUESTIONS

*color [attributes]

*size [in degree, degree, out-degree]

GEPHI DATA LAB/OVERVIEW

WEB ENVIRONMENT

SPREADSHEET

PRINTED NETWORK

BIG SCREEN

ONLINE COLLABORATIVE BOARD

PRESENT/STAGE THE NETWORK
EXPORT THE NETWORK IN A VECTOR FORMAT
`.svg` `.eps`

EDIT
alike Adobe Illustrator, Inkscape

ADD A TITLE

ADD ANNOTATIONS

ADD A KEY

REFINE COLORS

ADJUST LABEL SIZE

Title *annotations*
label
label

QUERY DESIGN AND DATASET DESIGN PROCESS

BUILDING THE NETWORK

VISUALISING

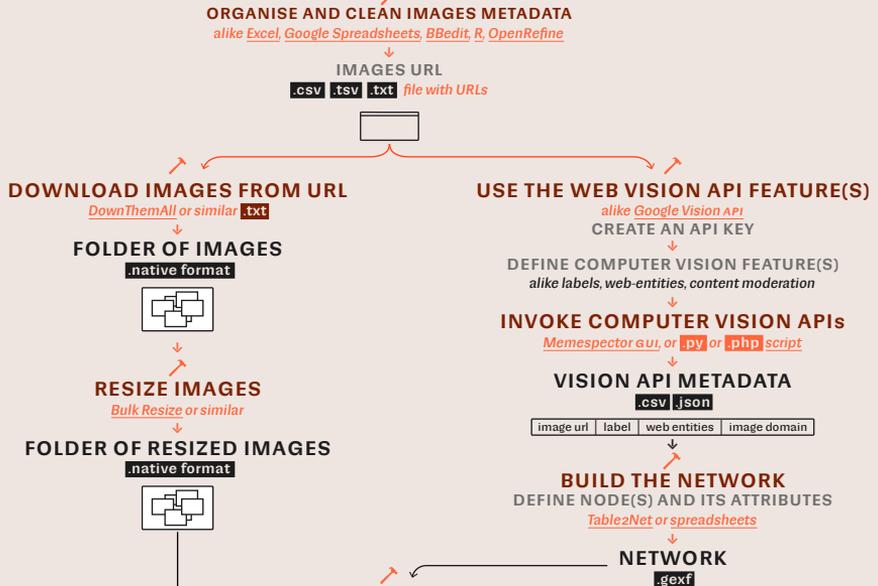**VISUAL NETWORK ANALYSIS**

VISUALISATION ARRANGEMENT

Figure 4: The art of querying platforms for creating an image dataset.
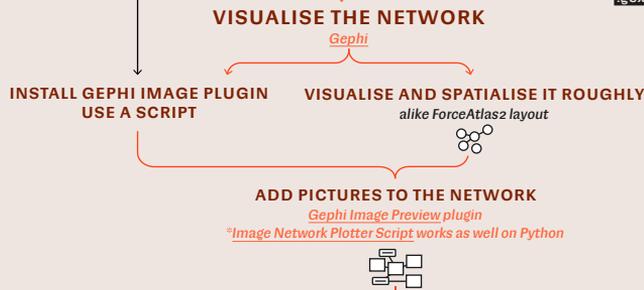
Figure 5: Building the network with the output files of the extraction software and the vision API.

Figure 6: Visualizing the network.
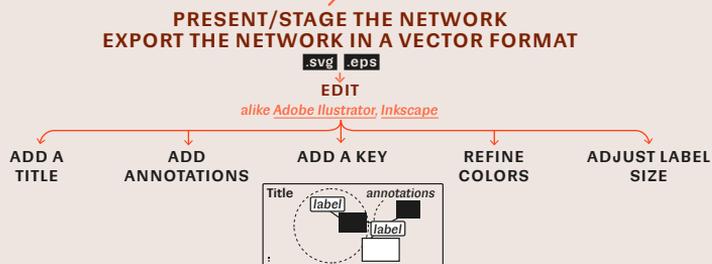
Figure 7: Visual network analysis and the navigational practice approach.

Figure 8: Presenting the network.

Figures 4–8: Adapted from Omena (in press) and designed by Beatrice Gobbo.

RESEARCH QUESTION
↓
QUERYING THE PLATFORM
alike Instagram, Tumbrl, Google Images

QUERY DESIGN
↓

○ *researcher intervention*
○ *computational medium*
■ *software output*
■ *technical practice*

EXTRACT IMAGES METADATA
alike FacePager, PhanthomBuster, Tumblr Tool, 4-CAT
or `.py` *scripts according to the analyzed digital platform*

IMAGES METADATA
images URL, engagement metrics, timestamps, textual content `.csv`

ORGANISE AND CLEAN IMAGES METADATA
alike Excel, Google Spreadsheets, BBedit, R, OpenRefine
↓
IMAGES URL
`.csv` `.tsv` `.txt` *file with URLs*

DOWNLOAD IMAGES FROM URL
DownThemAll or similar `.txt`

FOLDER OF IMAGES
`.native format`

RESIZE IMAGES
Bulk Resize or similar
↓
FOLDER OF RESIZED IMAGES
`.native format`

USE THE WEB VISION API FEATURE(S)
alike Google Vision API
CREATE AN API KEY
↓
DEFINE COMPUTER VISION FEATURE(S)
alike labels, web-entities, content moderation

INVOKE COMPUTER VISION APIs
Memespector GUI, or `.py` *or* `.php` *script*

VISION API METADATA
`.csv` `.json`

| image url | label | web entities | image domain |

BUILD THE NETWORK
DEFINE NODE(S) AND ITS ATTRIBUTES
Table2Net or spreadsheets
↓
NETWORK
`.gexf`

VISUALISE THE NETWORK
Gephi

INSTALL GEPHI IMAGE PLUGIN USE A SCRIPT

VISUALISE AND SPATIALISE IT ROUGHLY
alike ForceAtlas2 layout

ADD PICTURES TO THE NETWORK
Gephi Image Preview plugin
*Image Network Plotter Script works as well on Python

VISUALLY ANALYSE THE NETWORK
Gephi + domain knowledge

GO BACK TO THE RESEARCH QUESTIONS

CHANGE NETWORK APPEARENCE ACCORDING TO DATA
*color [attributes]
*size [in degree, degree, out-degree]

GEPHI DATA LAB/OVERVIEW
SPREADSHEET

WEB ENVIRONMENT
PRINTED NETWORK

BIG SCREEN
ONLINE COLLABORATIVE BOARD

PRESENT/STAGE THE NETWORK
EXPORT THE NETWORK IN A VECTOR FORMAT
`.svg` `.eps`

EDIT
alike Adobe Ilustrator, Inkscape

ADD A TITLE   **ADD ANNOTATIONS**   **ADD A KEY**   **REFINE COLORS**   **ADJUST LABEL SIZE**

Title | label | annotations | label

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

3   Using the browser plug-in DownThemAll! or the Firefox Adds-on Download All Images.

4   The Offline Image Query and Extraction Tool can help in this task.

5   For instance, using https://bulkresizephotos.com/pt
6   A tutorial to obtain a credential key file from Google Vision is available: https://github.com/jason-chao/memespector-gui/blob/master/doc/GetKeyFromGoogleCloud.md

Once having the output file(s) of the extraction software in hand, it is crucial *to see and verify what image metadata has to offer* (e.g., engagement metrics, timestamp, username, URLs) and *what can (later) serve as valuable analytical information*. It is also advisable to download the images[3] as soon as possible due to the short lifespan of image URLs (Figure 5). Here, the mask used to name the image files, e.g., to opt for *name*.*ext* (image id and .JPG), facilitates both the making of the network and the analysis. Researchers can opt to explore specific collections of images within a dataset by selecting images according to associated account profiles, link domains, engagement metrics, published time (date, month, or year), or computer vision outputs (e.g., labels or web entities).[4] Additionally, when navigating the network visualization, the image filename (e.g., 0a1781ec1998eedf38e9742acd185beb870e089e.jpg) is also useful to filter and analyze ego networks.

In the process of dataset design, the images must be organized in a file containing the selected images' URLs or a folder of images locally stored. The latter option is the most suitable for social media images, due to the short life of URLS. The images must be resized after downloaded.[5] To gain access to a computer vision API, researchers need to sign up and get authentication keys[6] (Figure 5). Next, the folder of images located on the computer, or a text file containing image location (advisable for large datasets), is finally inserted into Memespector-GUI (Chao, 2021) to communicate and request information from the vision API. Alternatively, Python or PHP script files can also be used to invoke vision APIs. The researcher must be attentive to the script configuration file in use (config file), knowing how its operation meets the vision API features in order to take advantage of both technical mediums.

With the outputs of the extraction software and the vision API, we reach the point of building and visualizing the network (Figure 6). Here, images acquire a new layer of meaning provided by the vision API, e.g., *labels*, *web entities*. We now can *see and verify what the vision API metadata has to offer*. When making a network, crucial decisions have to be made; researchers can assign or create node attributes, but for doing that they must spend some time exploring the dataset. Following this, tools like Table2Net (Jacomy, 2013) help in extracting a network from a comma-separated value file, converting it into a graph exchange XML format (.GEXF). This is where we should inform how connections are made between images and the Vision API features, including node attributes (numerical or textual).

A graph file format (GEXF or GDF) is required to visualize the network (Figure 6), which gains a shape through the work of a force-directed algorithm and life through its interpretation and results yet to come. The visualization is mediated by Gephi (Bastian et al., 2009) —or similar software— and

DISEÑA 19
AUG 2021
ARTICLE 4

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

ForceAtlas2 (Jacomy et al., 2014) which has a specificity in its force-directed drawing, "placing each node depending on the other nodes. This process depends only on the connections between nodes" (Jacomy et al., 2014, p. 2). Eventual attributes of nodes (what changes their size or color) are meaningless in the spatialization of the network, although reflecting analytical decisions. What spatializes the nodes are the weighted connections (see Venturini et al., 2015), referring to the relational nature of platform data.

In computer vision networks, nodes are spatialized according to how the computer vision feature is associated (or not) to an image. Here, researchers choose the eventual node attributes according to the character and features of the image collection and according to the research questions. To depict images, researchers can either install a Gephi plug-in (e.g., ImagePreview)[7] or use a file script (e.g., image-network-plotter).[8] To make the network easier to read, in Gephi the nodes with ineffective image descriptions may be removed, e.g., 'product'. By acquiring another layer of meaning from Gephi and ForceAtlas2, new arrangements are imprinted in images.

The analytical process is guided by visual network analysis (see Jacomy, 2021; Venturini et al., 2019). This technique draws our attention to the position, size, and color of nodes as key interpretative aspects, highlighting the relational nature of digital data in tandem with the technicity of the medium. At this stage, the intervention of the researcher(s) is crucial. Interpretation means to navigate back and forth between the Gephi *overview/data laboratory* and the web environment, between the spreadsheet and the big screen, while making annotations in printed versions of the network or online whiteboards. A navigational procedure in the analysis is not required but mandatory.

When analyzing the network, researchers should recall where the images come from, what informs the position of nodes and their attributes, and the specificities of the vision API features. Here, the researcher knows that the final visual metaphor neither entirely represents the image dataset (extracted from a digital platform), nor entirely exposes the Vision API in its full potential. The network may reflect *a second order of grammatization* co-created with what the images represent (and associated metadata), the vision API, and research software outputs, yet mediated by the researcher intervention (Omena, in press).

In the last step of the protocol (Figure 8), the visualization needs to be adequately presented and organized by adding the title and the key, and converting any previously handwritten annotations into digital format, taking into consideration how it will be experienced (printed or digital) and the research questions previously asked.

7   A plug-in for Gephi by the Yale Computer Graphics Group, available from https://gephi.org/plugins/#/plugin/image-preview
8   A file script by André Mintz, available from https://github.com/amintz/image-network-plotter

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
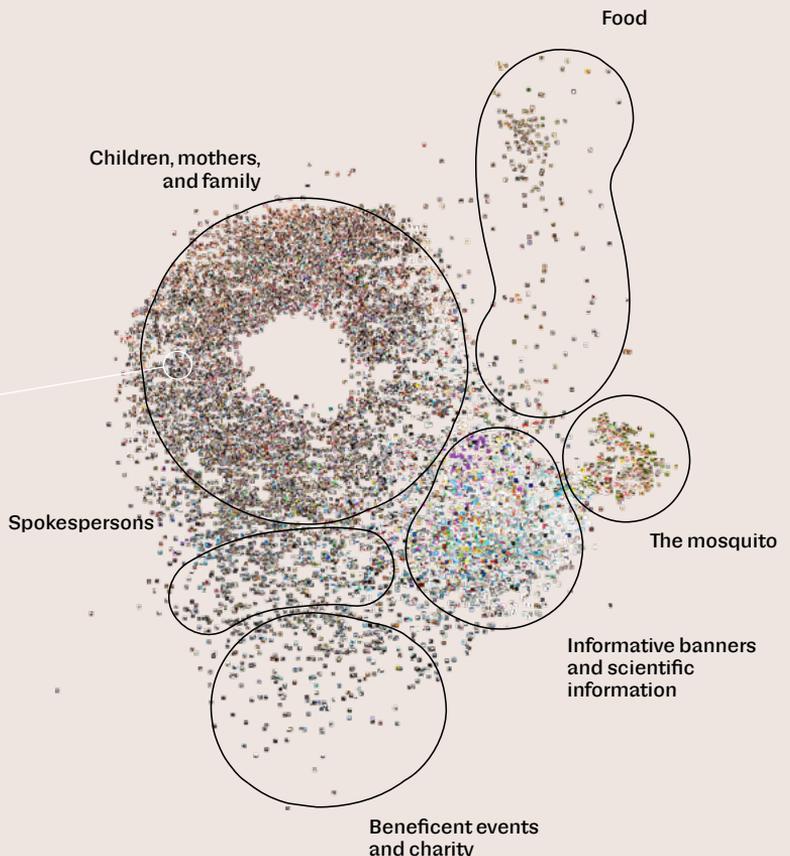TO STUDY NATIVELY DIGITAL IMAGES

READING NETWORKS OF IMAGES BUILT UPON GOOGLE VISION API

Image-Label Network: Image descriptions from its material content

This case study addresses the spread of the Zika virus, which the World Health Organization declared a Public Health Emergency of International Concern (see Barreto et al., 2016). This epidemiological scenario prompted warnings to pregnant women in Brazil and other several Latin American and Caribbean countries, raising questions of how families affected by the emerging Zika virus infection were coping with the situation. In this respect, we asked: How has the general public *visually* reacted to the Zika epidemic issue? What can the visualities associated to #microcefalia (microcephaly) tell about Brazilian families impacted by the virus?

In figure 9, the image-label network gathers 10.797 images published by Instagram users between June 2012 and October 2017, who used the hashtag #microcefalia. The data was collected over seven months (from July 2017 to January 2018), using Visual Tagnet Explorer to call the deprecated Instagram Platform API with the query 'microcefalia'. The analyses were performed

Figure 9: Image-Label Network of Microcephaly on Instagram. 10.797 hashtagged images (#microcefalia) published between June 2012 and October 2017. Source: The authors.

Food

Children, mothers, and family

Spokespersons

The mosquito

Informative banners and scientific information

Beneficent events and charity

DISEÑA 19
AUG 2021
ARTICLE.4

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

in January 2018 when the platform registered a total number of 16,049 existing publications containing #microcefalia. The resulting network was then used for qualitative exploration (or relabeling) of the image clusters, allowing the interpretation of how health issues are inscribed in Instagram.
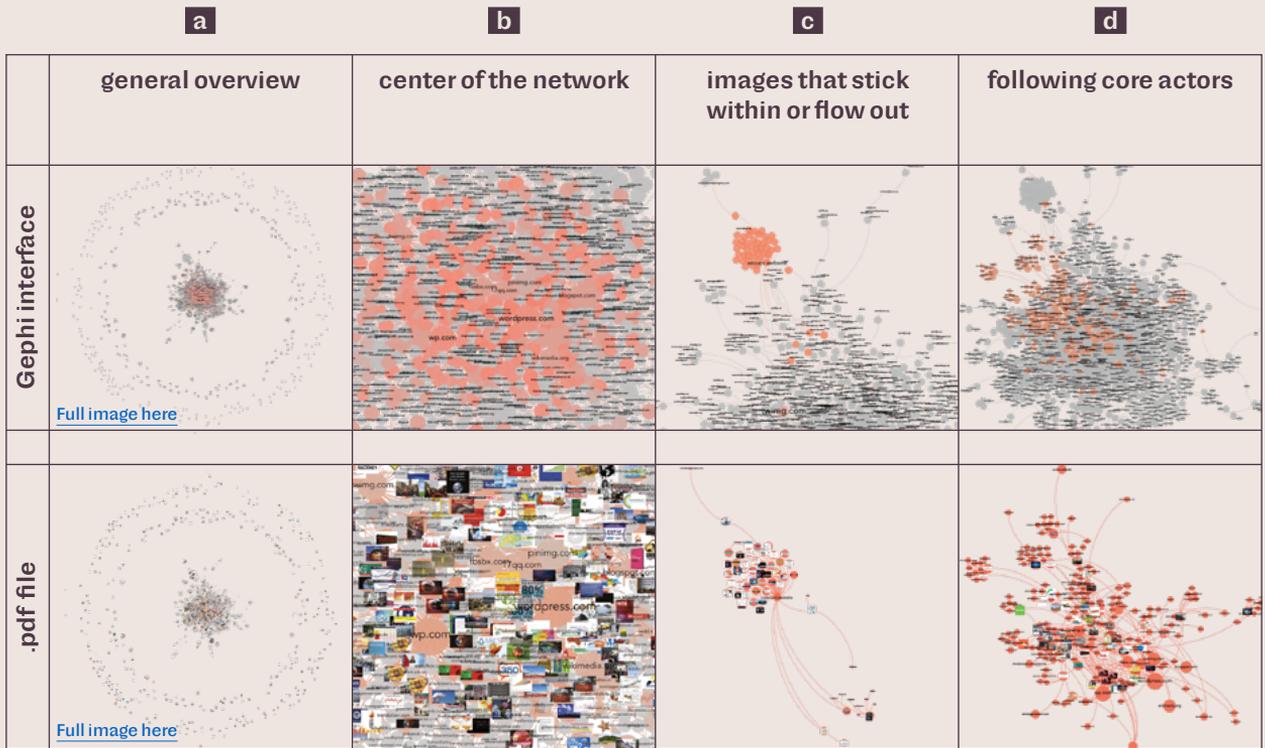
The strength of the image-label approach lies in the capacity of exploring variously situated forms of visual expression engaged in the public negotiation of the issue in question. As we can see in the case of the #microcefalia image-label network (Figure 9), a group of images is placed together when they share the same or related *labels*, forming clusters that address different concerns with the Zika virus epidemic in Brazil (Figure 9). The imagery of #microcefalia on Instagram associates with the everyday life of microcephalic kids. We see infants and their families engaging in social activities or physiotherapy sessions, illustrative explanations of microcephaly congenital disabilities, including the image of the mosquito that transmits the virus. The hashtag-specific image content also exposes events, spokespersons, and food related to microcephaly.

Further explorations of the metadata connecting the images are also possible, e.g., images can be sized according to audience reactions (likes, comments) or the year/date of publication, guiding the interpretative process in a different way. When reading this type of network, it is crucial to keep in mind the vision API's degree of specificity in labeling (questioning to what extent this can be useful) and the ethical issues that specific collections of images may raise.

Image-Web Entities Network: Image descriptions from web content
Figure 10 zooms in 500 #femalepresentingnipples Tumblr images classified by Google Vision *web entities*. In this network, each visible web entity is connected to at least five images. The size of the image nodes corresponds to the total number of reblogs, comments, and likes or 'notes', highlighting memes that went viral on Tumblr in December 2018 (shortly after the platform announced its new porn-banning content policy). The data was retrieved with the Tumblr Tool (Rieder, 2015) through a series of time- and issue-specific hashtag queries, as part of a larger set of 15,158 images circulating in the wake of the Tumblr porn ban, over ten months (Pilipets et al., 2020; Pilipets & Paasonen, 2020). The focus on #femalepresentingnipples was specified based on users' memetic tactics of mocking Tumblr's updated community guidelines, which state that "Adult content primarily includes photos, videos, or GIFs that show real-life human genitals or female-presenting nipples, and any content —including photos, videos, GIFs and illustrations— that depicts sex acts" (Tumblr, 2018). Producing sarcastic commentary far beyond the platform itself, the #femalepresentingnipples hashtag involved a wide range of artistic, activist, and popular culture connotations that Figure 10 addresses in response to the following question: What can we learn about the memetic scenar-
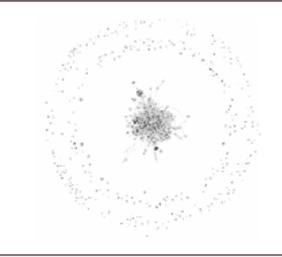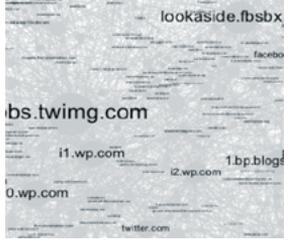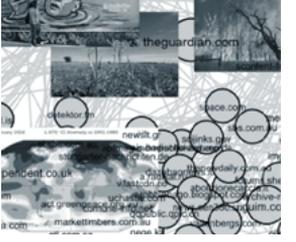
DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

Figure 10: Image-Web entities network: a zoom-in view focusing on the Internet meme cluster of 500 #femalepresentingnipples images in the context of the Tumblr porn ban in December 2018. Source: The authors.

ios of Tumblr porn purge by looking at hashtag-specific user-generated content and its corresponding web references?

The strength of the web entity approach to visual public controversies, derives from the ability of Google Vision to utilize the open web for annotating images according to the most common references that visually similar images are captioned with (Leetaru, 2019; Robinson, 2017). The 'Internet meme' cluster, for example, specifies memetic scenarios related to 'free the nipple' movement, Pikachu, Lisa Simpson, Harry Potter, The Office, and Star Wars, the latter prominently parodying Senator Amidala's line "So this is how liberty dies" with over 3,000 notes (see Figures 10 and 12). By drawing recent information about the images from the web environment with the aid of metadata connected to the ranking of each page, this approach provides a web-driven contextual perspective, which is not necessarily informed in the image content itself, but outside of it. Web entities such as 'Know Your Meme' or 'Imgur' suggest a certain correlation with the significance of the sites as memetic archives where the contents of popular memes and their web captions adapt in real-time to emerging events. The emphasis on the web environment of the image thus helps us to understand new frames of reference, in which artworks such as 'Portrait of a Young Woman as Saint Agatha' or 'Liberty Leading the People' can be *re*embedded in a new context (e.g., 'Tumblr') and provided with new associations ('female presenting nipples').

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

	a general overview	**b** center of the network	**c** images that stick within or flow out	**d** following core actors
Gephi interface	Full image here			
.pdf file	Full image here			

Figure 11: Image-Domain network. From left to right, a zoom-in view focusing on: a general overview of the network, dominant link domains (center of the network), the analysis of images that stick and flow out of web pages, ego networks to follow core actors, and associated visualities. Source: The authors.

On the one hand, this means that web entities allow researchers to grasp temporally situated transformations in the web-distributed topics and practices of contextualization that vernacular images, reaction GIFs, and Internet memes derive from. On the other hand, despite its affordances, the Google Vision API web entity feature adds some uncertainty in the process of interpretation, especially when it comes to the logic of content adaptation. As the relations of images, their textual captions, and matching image domains are constantly evolving, so do the annotations and the confidence scores provided by Google Vision API. In the process of network exploration, this requires paying attention to the *less* confident descriptions, since this is where contextual image modifications first register, each pointing to the temporal issue- and site-specific relevance of the images and their varying range of associations.

Image-Domain Network: The sites of image circulation across the web
This case study served as a methodological tool for exploring networks of image circulation. Figure 11 illustrates the networks and how we can navigate them. The two rows depict the same image dataset, helping to situate the network in different environments; its exploration in Gephi interface (top), followed, and in

Figure 12 (next page): Reading computer vision networks according to type, how/what we see, and what we read. Source: The authors.

TYPE OF NETWORK *platform and case study*	IMAGE-LABEL *platform* **INSTAGRAM** *case study* **#microcefalia**	IMAGE-WEB ENTITIES *platform* **TUMBRL** *case study* **#femalepresentingnipples**	IMAGE-DOMAIN *platform* **GOOGLE SEARCH IMAGES** *case study* **CLIMATE EMERGENCY**
HOW WE SEE			
WHAT WE SEE	Mother and her child with microcephaly 	Internet memes and GIFs 	Domains hosting the greater recurrence of image appearance
	Scientific explanations 	Star Wars adaptations 	Images shared by specific domains (newsmedia)
	The Mosquito 	Stock photography and free the nipple 	Domains sharing the same image
	Newsmedia updates 	Art references Saint Agatha 	The imagery that stick within and flows out of the platform
WHAT WE READ	**Literal descriptions of images based on existing labels.** *The content within the image matters.* *Interpreting the outputs of machine learning models to image classification.*	**Contextual descriptions of images based on existing web entities.** *The content within the image and the context outside of it matter.* *An interpretation from the web content/ perspective.*	**A dynamic view of images and their sites of circulation.** *Images fully matched and found in different web pages or URLs matter.* *An interpretation from the web content/ perspective.*

DISEÑA 19
AUG 2021
ARTICLE.4

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

parallel, by its exploration in PDF files (bottom). We see 992 images ranked by Google Image Search when searching for 'climate emergency', reflecting the top 100 images suggested by the search engine over the years (from 2008 to 2019). Google Image Extractor withdrew the image URLs from the search engine results and DownThemAll! facilitated the download of available images. In the network, nodes are top-level domains (pointing to where the images were found) and images (suggested by Google Image Search), connections indicating whether a given image was found in URLs.

In the network overview (Figure 11a), the two rings around the main component reveal not fully matched images on the web (first ring) or images only found in specific URLs (second ring). We may ask what images (and domains) are these, or why they have not circulated elsewhere. Looking at the center of the network (Figure 11b), we find the domains (outlined in red) where most of the images related to climate emergency were found. Not unexpectedly, we see popular platforms (twimg.com, fbsbx.com, pinimg.com), sources of reputation information (wikimedia.org), and expert sources of information (blog spheres e.g., blogspot.com, wp.com). Results reflect Google's ranking systems and search quality rater guidelines (Google User Content, 2020), also exposing the hierarchical structure of the web, which delivers content arranged in a particular order (see Berners-Lee, 1995; Jacomy, 2019): content known of all, known of amateurs and experts, forgotten (Jacomy, 2019). Here, social media and Wikipedia belong to the top layer. Therefore, when reading networks of image circulation, researchers must understand that what they are seeing includes the layered structure of online connectivity through Google's eyes.

This case study points towards three analytical strategies to explore sites of image circulation. First, to identify images that stick within or flow out of platforms by visually detecting a group of images that appear only in a domain. As we see in Figure 11c, diagrams and graphics appear only on *secure.website*, while a few images are found within this domain but also on Twitter and WordPress. The second technique consists of following core (or low profile) actors within the network, e.g., using Twitter's neighbors (Figure 11d) as an information resource for the analysis of both the images and associated actors. The third points to a qualitative analysis of one or more images associated with a group of link domains, revealing the type of visuality shared by specific actors, e.g., looking at images appearing in a cluster of local and international media.

The strength of the image-domain approach lies in the study of image circulation and actors to whom specific visualities matter. By considering the latter, the image-domain network sheds light on questions such as: What are the sites of image circulation (the sites hosting fully or partially matching images)? Who share the same or similar visual content? Which actors appear to be more

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

influential in terms of distributing images across the web? Which images tend to stick within and flow out of platforms?

DISCUSSION AND CONCLUSIONS

In this article, we presented the potentials of Google Vision API-based networks for studying online images, covering three important modalities, which are part of a critical visual methodology: the image itself, its specific 'audiencing' through references obtained from the web, and the sites of circulation (Rose, 2016). To this end, we defined (conceptually and technically) three types of computer vision networks (Figure 12) and presented a digital methods protocol (Figure 3) to build and interpret these networks. Drawing on the case studies, we identified the following points of departure for a critical understanding of Google Vision API-based networks: (1) the composition of contents assembled within the image; (2) the references to an image obtained from web content; and (3) the web pages or URLs hosting fully matching visual objects. By allowing to shift between different perspectives on the same image dataset, each point of departure helps to address the multiplicity of *natively* digital images. The networks offer visual methodologies with new approaches to image analysis.

Networks of image description (arranged through Google Vision API *label* and *web entities* features) enable researchers to, respectively, make sense of literal and contextual descriptions of a collection of images. The former informs what is *in* the image; the latter provides images with web-driven annotations. In both types of networks, clusters of images sharing the same or similar content guide the process of interpretation. In networks of image circulation, researchers can explore clusters of images that relate to only one particular platform that appear across multiple sites (including mainstream platforms and news media websites) or individual images (outliers) that appear only in specific websites (Figure 12). This helps to detect which visual content sticks within and flows out of a given platform, opening further possibilities for qualitative analysis of visual vernaculars (see Gibbs et al., 2015; Pearce et al., 2020).

We argued that engaging with technical practices inherent to the digital methods approach, through the research diagram protocol, is an essential element of building and interpreting computer vision networks. Network interpretation is informed by the art of querying platforms (resulting in a good collection of images), from what an image depicts to which interactions it involves, where the image is hosted (e.g., social media, search engines) and how it is rendered visible by the platform mechanisms (e.g., ranking systems). This points to three challenges when using computer vision networks for research purposes. First, since thick layers of technical mediation characterize the work with these networks, it requires certain proximity with software in its language, content, and performance

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

(see Omena, in press; Rieder, 2020). Second, researchers' need to understand technical practices and the role of these practices, in order to successfully implement the methodology (see Figure 3). Third, understanding how images are rearranged and modified in the process of analysis through different types of network design. Different features of computer vision APIs suggest different directions in the process of visual network exploration and analyses.

This article opens up new opportunities for future research with computer vision-mediated visual methodologies, requiring sensitivity to what the machine perceives as relevant and how this relevance is co-constituted through both, the practice of digital methods, and the practices of image use on social media platforms and across the web. ▣

* * *

REFERENCES
AZAR, M., COX, G., & IMPETT, L. (2021). Introduction: Ways of Machine Seeing. *AI & Society*. https://doi.org/10.1007/s00146-020-01124-6

BARRETO, M. L., BARRAL-NETTO, M., STABELI, R., ALMEIDA-FILHO, N., VASCONCELOS, P. F. C., TEIXEIRA, M., BUSS, P., & GADELHA, P. E. (2016). Zika Virus and Microcephaly in Brazil: A Scientific Agenda. *The Lancet, 387*(10022), 919–921. https://doi.org/10.1016/S0140-6736(16)00545-6

BASTIAN, M., HEYMANN, S., & JACOMY, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media, 3*(1), 361–362. https://doi.org/10.1136/qshc.2004.010033

BERNERS-LEE, T. (1995). *Hypertext and Our Collective Destiny* [Video]. 1995 Vannevar Bush Symposium - Videos and Links - Doug Engelbart Institute. Retrieved September 20, 2020, from https://www.dougengelbart.org/content/view/258/000/

CHAO, J. (2021). Memespector GUI: Graphical User Interface Client for Computer Vision APIs [Computer Software] (Version 0.1) [C#] (Original work published 2021). https://github.com/jason-chao/memespector-gui

COLOMBO, G. (2018). *The Design of Composite Images. Displaying Digital Visual Content for Social Research* [Doctoral Dissertation, Politecnico di Milano]. http://hdl.handle.net/10589/141266

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

D'ANDRÉA, C., & MINTZ, A. (2019). Studying the Live Cross-Platform Circulation of Images with Computer Vision API: An Experiment Based on a Sports Media Event. *International Journal of Communication, 13*, 1825–1845.

GARIMELLA, K., & ECKLES, D. (2020). Images and Misinformation in Political Groups: Evidence from WhatsApp in India. *Harvard Kennedy School Misinformation Review, 1*(5). https://doi.org/10.37016/mr-2020-030

GEBOERS, M. A., & VAN DE WIELE, C. T. (2020). Machine Vision and Social Media Images: Why Hashtags Matter. *Social Media + Society, 6*(2), 2056305120928485. https://doi.org/10.1177/2056305120928485

GEBRU, T., MORGENSTERN, J., VECCHIONE, B., VAUGHAN, J. W., WALLACH, H., DAUMÉ III, H., & CRAWFORD, K. (2020). Datasheets for Datasets. *ArXiv*, 1803.09010 [cs].

GERLITZ, C., & RIEDER, B. (2018). Tweets Are Not Created Equal: Investigating Twitter's Client Ecosystem. *International Journal of Communication*, 12. https://hdl.handle.net/11245.1/4da1d406-1213-4103-8237-eef5ae786948

GIBBS, M., MEESE, J., ARNOLD, M., NANSEN, B., & CARTER, M. (2015). #Funeral and Instagram: Death, Social Media, and Platform Vernacular. *Information, Communication & Society, 18*(3), 255–268. https://doi.org/10.1080/1369118X.2014.987152

GOOGLE CLOUD. (2017, April 13). *Pricing | Google Cloud Vision API Documentation*. Wayback Machine. https://web.archive.org/web/20170413081619/https://cloud.google.com/vision/docs/pricing

GOOGLE USER CONTENT. (2020). *General Guidelines*. https://static.googleusercontent.com/media/guidelines.raterhub.com/pt-BR//searchqualityevaluatorguidelines.pdf

HOCHMAN, N. (2014). The Social Media Image. *Big Data & Society, 1*(2), 2053951714546645. https://doi.org/10.1177/2053951714546645

HOELZL, I., & MARIE, R. (2015). *Softimage: Towards a New Theory of the Digital Image*. Intellect.

JACOMY, M. (2013). Table 2 Net [Computer Software]. https://medialab.github.io/table2net/

JACOMY, M. (2019). The Web as Layers [Video]. *HYPHE - Session 4*. https://panopto.aau.dk/Panopto/Pages/Viewer.aspx?id=48cfe5ff-5503-431b-887f-ab53007ef5c4

JACOMY, M. (2021). *Situating Visual Network Analysis* [Doctoral Dissertation, Aalborg University]. https://reticular.hypotheses.org/1879

JACOMY, M., VENTURINI, T., HEYMANN, S., & BASTIAN, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE, 9*(6), e98679. https://doi.org/10.1371/journal.pone.0098679

LEETARU, K. (2019, June 1). Using Google Vision AI's Reverse Image Search To Richly Catalog Television News. *Forbes*. https://www.forbes.com/sites/kalevleetaru/2019/06/01/using-google-vision-ais-reverse-image-search-to-richly-catalog-television-news/

MACKENZIE, A., & MUNSTER, A. (2019). Platform Seeing: Image Ensembles and their Invisualities. *Theory, Culture & Society, 36*(5), 3–22. https://doi.org/10.1177/0263276419847508

MANOVICH, L. (2020). *Cultural Analytics*. MIT Press.

MAURI, M., BRIONES, M. A., GOBBO, B., & COLOMBO, G. (2020). Research Protocol Diagrams as Didactic Tools to Act Critically in Dataset Design Processes. *INTED2020 Proceedings*, 9034–9043. https://library.iated.org/view/MAURI2020RES

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

MINTZ, A., SILVA, T., GOBBO, B., PILIPETS, E., AZHAR, H., TAKAMITSU, H., OMENA, J. J., & OLIVEIRA, T. (2019). *Interrogating Vision APIs*. #SMARTDataSprint. https://smart.inovamedialab.org/past-editions/smart-2019/project-reports/interrogating-vision-apis/

NIEDERER, S., & COLOMBO, G. (2019). Visual Methodologies for Networked Images: Designing Visualizations for Collaborative Research, Cross-platform Analysis, and Public Participation. *Diseña*, (14), 40–67. https://doi.org/10.7764/disena.14.40-67

OMENA, J. J. (in press). *Digital Methods and Technicity-of-the-Mediums. From Regimes of Functioning to Digital Research* [Doctoral Dissertation]. Universidade NOVA de Lisboa.

OMENA, J. J., & AMARAL, I. (2019). Sistema de leitura de redes digitais multiplataforma. In J. J. Omena (Ed.), *Métodos Digitais: Teoria-prática-crítica* (pp. 121–140). ICNOVA.

OMENA, J. J., & GRANADO, A. (2020). Call into the Platform! *Revista ICONO14 Revista Científica de Comunicación y Tecnologías emergentes, 18*(1), 89–122. https://doi.org/10.7195/ri14.v18i1.1436

OMENA, J. J., RABELLO, E. T., & MINTZ, A. G. (2020). Digital Methods for Hashtag Engagement Research. *Social Media + Society, 6*(3), 2056305120940697. https://doi.org/10.1177/2056305120940697

PAGLEN, T. (2014, March 13). *Seeing Machines*. Is Photography Over? Fotomuseum Winterthur Series. https://www.fotomuseum.ch/en/2014/03/13/seeing-machines/

PARIKKA, J., & DVOŘÁK, T. (2021). Introduction: On the Scale, Quantity and Measure of Images. In T. Dvořák & J. Parikka (Eds.), *Photography Off the Scale: Technologies and Theories of the Mass Image* (pp. 1–24). Edinburgh University Press.

PEARCE, W., ÖZKULA, S. M., GREENE, A. K., TEELING, L., BANSARD, J. S., OMENA, J. J., & RABELLO, E. T. (2020). Visual Cross-platform Analysis: Digital Methods to Research Social Media Images. *Information, Communication & Society, 23*(2), 161–180. https://doi.org/10.1080/1369118X.2018.1486871

PILIPETS, E., FLORES, A. M. M., FLAIM, G., SKAZEDONIG, M., SEPÚLVEDA, R., & DEL NERO, S. (2020). *From "Tumblr Purge" to "Female Nipples": Telling a Story of Platform Censorship Critique through Memes and Digital Methods*. #SMARTDataSprint. https://smart.inovamedialab.org/2020-digital-methods/project-reports/tumblr-purge-female-nipples/

PILIPETS, E., & PAASONEN, S. (2020). Nipples, Memes, and Algorithmic Failure: NSFW Critique of Tumblr Censorship. *New Media & Society*, 1461444820979280. https://doi.org/10.1177/1461444820979280

RETTBERG, J. W. (2020). Situated Data Analysis: A New Method for Analysing Encoded Power Relationships in Social Media Platforms and Apps. *Humanities and Social Sciences Communications, 7*, Art. 5. https://doi.org/10.1057/s41599-020-0495-3

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

RIEDER, B. (2015). YouTube Data Tools [Computer Software] (Version 1.11). https://tools.digitalmethods.net/netvizz/youtube/

RIEDER, B. (2020). *Engines of Order: A Mechanology of Algorithmic Techniques*. Amsterdam University Press.

RIEDER, B., & RÖHLE, T. (2018). Digital Methods: From Challenges to Bildung. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society: Studying Culture through Data* (pp. 109–124). Amsterdam University Press. https://doi.org/10.1515/9789048531011-010

DISEÑA 19
AUG 2021
ARTICLE.1

JANNA JOCELI OMENA
ELENA PILIPETS
BEATRICE GOBBO
JASON CHAO

THE POTENTIALS OF GOOGLE VISION API-BASED NETWORKS
TO STUDY NATIVELY DIGITAL IMAGES

ROBINSON, S. (2017, June 1). *Exploring the Cloud Vision API*. Medium. https://medium.com/@srobtweets/exploring-the-cloud-vision-api-1af9bcf080b8

ROGERS, R. (2013). *Digital Methods*. MIT Press.

ROGERS, R. (2019). *Doing Digital Methods*. SAGE.

ROSE, G. (2016). *Visual Methodologies* (4th ed.). Open University.

RUBINSTEIN, D., & SLUIS, K. (2013). Concerning the Undecidability of the Digital Image. *Photographies, 6*(1), 151–158. https://doi.org/10.1080/17540763.2013.788848

SCHWEMMER, C., KNIGHT, C., BELLO-PARDO, E. D., OKLOBDZIJA, S., SCHOONVELDE, M., & LOCKHART, J. W. (2020). Diagnosing Gender Bias in Image Recognition Systems. *Socius, 6*, 2378023120967171. https://doi.org/10.1177/2378023120967171

SILVA, T., BARCIELA, P., & MEIRELLES, P. (2018). Mapeando Imagens de Desinformação e Fake News Político-Eleitorais com Inteligência Artificial. 3o CONEC: Congresso Nacional de Estudos Comunicacionais Da PUC Minas Poços de Caldas - Convergência e Monitoramento, 413–427. https://www.researchgate.net/publication/329525177_Mapeando_Imagens_de_Desinformacao_e_Fake_News_Politico-Eleitorais_com_Inteligencia_Artificial

SILVA, T., MINTZ, A., OMENA, J. J., GOBBO, B., OLIVEIRA, T., TAKAMITSU, H. T., PILIPETS, E., & AZHAR, H. (2020). APIs de Visão Computacional: Investigando mediações algorítmicas a partir de estudo de bancos de imagens. *Logos, 27*(1), 25–54. https://doi.org/10.12957/logos.2020.51523

STEYERL, H. (2009). In Defense of the Poor Image. *E-Flux Journal, 10*. https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/

SULLIVAN, D. (2020, May 20). *A Reintroduction to our Knowledge Graph and Knowledge Panels*. Blog.Google. https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/

SZELISKI, R. (2021). *Computer Vision: Algorithms and Applications* (2nd ed.). Springer.

TAIBI, D., ROGERS, R., MARENZI, I., NEJDL, W., AHMAD, Q. A. I., & FULANTELLI, G. (2016). Search as Research Practices on the Web: The Sar-web Platform for Cross-language Engine Results Analysis. *Proceedings of the 8th ACM Conference on Web Science*, 367–369. https://doi.org/10.1145/2908131.2908201

TUMBLR. (2018, December). Updates to Tumblr's Community Guidelines [Tumblr]. *Tumblr Support*. https://support.tumblr.com/post/180758979032/updates-to-tumblrs-community-guidelines

VENTURINI, T., JACOMY, M., & JENSEN, P. (2019). *What Do We See When We Look at Networks. An Introduction to Visual Network Analysis and Force-Directed Layouts* (SSRN Scholarly Paper ID 3378438). Social Science Research Network. https://doi.org/10.2139/ssrn.3378438

VENTURINI, T., JACOMY, M., & PEREIRA, D. (2015). Visual Network Analysis: The Example of the Rio+20 Online Debate. https://hal.archives-ouvertes.fr/hal-02305124

XI, N., MA, D., LIOU, M., STEINERT-THRELKELD, Z. C., ANASTASOPOULOS, J., & JOO, J. (2020). Understanding the Political Ideology of Legislators from Social Media Images. *Proceedings of the 14th International AAAI Conference on Web and Social Media, ICWSM 2020*, 726–737. http://arxiv.org/abs/1907.09594