Digitale Methoden | *Digital Methods*

Michael Achmann-Denkler, Christian Wolff

Beyond the Feed: A Computational Blueprint for Multimodal Analysis of Ephemeral Instagram Stories

Abstract: Since their introduction in 2016, Instagram Stories have become a core feature of the platform, offering users short-lived posts that disappear after 24 h. Although research on Instagram content is growing, Stories remain understudied, possibly due to challenges in data collection. This chapter addresses that gap by presenting a comprehensive workflow for the computational analysis of Instagram Stories. The workflow outlines steps for data collection, preprocessing, and analysis of textual and visual content, utilising tools such as optical character recognition and automated transcription to capture Stories' multi-layered nature. By deconstructing each Story into its text, audio, and visual components, researchers can analyse these elements separately or in combination to uncover patterns. The adaptable framework supports various disciplines and research questions, including historical sciences or political communication, the latter being illustrated through a fictional case study of Martian election campaigns.

Additionally, the chapter explores using large language models like GPT-4 to automate content classification, showing how these tools assist in analysing text and images. The workflow emphasises computational approaches while advocating for the inclusion of human annotations to ensure accuracy. This method makes computational multimodal analysis more accessible to researchers with limited technical expertise. Practical guidance, including example notebooks, is available online to help researchers easily apply this methodology.

1 Introduction

Instagram Stories represent a unique type of social media post, characterised by their ephemeral nature, disappearing from a user's profile 24 h after posting. This format, pioneered by Snapchat and adopted by Instagram in 2016, has been hailed as a significant factor in the platform's success.[1] Unlike traditional posts, stories are archived and remain accessible to the author only, with the option to transform these fleeting moments into permanent highlights on the profile. Instagram

---

1 Tama Leaver et al., Instagram: Visual Social Media Cultures, Cambridge 2020.

provides several editing tools, including adding drawings, filters, text, emojis, and various stickers,[2] showcasing the app's "affordances" – features that enhance and shape user interactions and content creation.[3] Each story, essentially an image or video accompanied by metadata,[4] contributes to a chronological narrative and is subject to expiration.

Despite a growing research interest in Instagram content,[5] the focus predominantly rests on permanent posts, leaving a notable gap in the study of Instagram Stories. This may be attributed to two primary challenges: the ephemeral character of stories necessitating real-time data collection and the lack of tools to capture this content effectively. Posts, in contrast, may be collected retrospectively, and tools like *CrowdTangle* provided researchers with extensive datasets readily available for analysis.[6] Consequently, researchers may opt for the path of least resistance by using only the Instagram post data available through *CrowdTangle*, potentially introducing bias by overlooking other formats – such as stories, reels, or images in album posts – that the tool was not able to access.[7]

Addressing this gap, our contribution outlines a comprehensive workflow for collecting, preprocessing, and computationally analysing Instagram Stories, addressing textual and visual content. This blueprint is aimed at social scientists and (digital) humanists alike, offering adaptability across various research questions and disciplines. We demonstrate our approach through a hypothetical case

---

2 Alexandra Georgakopoulou, Sharing the moment as small stories: The interplay between practices & affordances in the social media-curation of lives, in: Narrative Inquiry 27/2 (2017), p. 311–333, doi.org/10.1075/ni.27.2.06geo; Lucia Bainotti et al., From archive cultures to ephemeral content, and back: Studying Instagram Stories with digital methods, in: New Media & Society, 23/12 (2020), doi.org/10.1177/1461444820960071.

3 For a discussion of platform affordances see: Michael Bossetta, The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election, in: Journalism & Mass Communication Quarterly 95/2 (2018), p. 471–496, doi.org/10.1177/1077699018763307.

4 Information like the date and time of posting or the author account handle constitute metadata. Below, we discuss in detail the different types of information that belong to a single story.

5 For example, Abderahman Rejeb et al., The big picture on Instagram research: Insights from a bibliometric analysis, in: Telematics and Informatics 73 (2022), doi.org/10.1016/j.tele.2022.101876.

6 *CrowdTangle* was shut down in August 2024 – which was not foreseeable at the time of writing. The tool provided researchers with retrospective access to Instagram and Facebook posts. Meta moderated access to *CrowdTangle*. As of the summer of 2024, the Meta Content Library has replaced the tool. A third party, the Inter-university Consortium for Political and Social Research (ICPSR), now moderates access, for which researchers still have to apply. See: transparency.meta.com/de-de/researchtools/meta-content-library/

7 Domenico Trezza, To scrape or not to scrape, this is dilemma. The post-API scenario and implications on digital research, in: Frontiers in Sociology 8 (2023), doi.org/10.3389/fsoc.2023.1145038.

study of an election campaign using computational content analysis techniques. The blueprint can be adapted to questions from other disciplines like public history. Using the same techniques, we might collect a different corpus, e.g. to understand what and how users post about their museum visits or about historical topics in general. Readers primarily interested in manual content analysis are encouraged to bypass the computational sections and proceed directly to the evaluation section. Our recommendations for data annotation during evaluation are fully compatible with manual content analysis. As this text only provides a high-level overview of the process, we invite readers to visit our website at social-media-lab.net for comprehensive tutorials and ready-to-use Jupyter notebooks. These resources are designed to assist with the practical aspects of social media research.

This contribution only touches on ethical and legal considerations. Our perspective is based on German law, where non-profit researchers benefit from certain exemptions under copyright law (Urheberrecht) for text and data mining. Researchers should familiarise themselves with the applicable laws before collecting data, obtain informed consent when working with personal profiles, and carefully consider privacy and data security throughout their research. The Association of Internet Researchers provides guidelines for navigating these ethical concerns.[8] We recommend McCrow-Young's reflections on their Instagram research for further insights.[9]

2 Workflow

We suggest a workflow grounded in the ideas of *Cultural Analytics*[10] by Lev Manovich and *Digital Methods*[11] by Richard Rogers. The concept of *Cultural Analytics* inspires social media analysis beyond text using computational approaches. It guides the general research workflow and its focus on content analysis to "see one billion images".[12] In contrast, the *Digital Methods* concept motivates data access through scraping and a particular focus on metadata, as Rogers proposes to use born-digital

---

8 aline shakti franzke et. al., Internet research: ethical guidelines 3.0: association of internet researchers (2020), aoir.org/reports/ethics3.pdf

9 Ally McCrow-Young, Approaching Instagram data: reflections on accessing, archiving and anonymising visual social media, in: Communication Research and Practice 7/1 (2021), p. 21–34, doi.org/10.1080/22041451.2020.1847820.

10 Lev Manovich, Cultural Analytics, Cambridge 2020.

11 Richard Rogers, Digital Methods, Cambridge 2013.

12 Title Chapter 1, Lev Manovich, Cultural Analytics, Cambridge 2020.

data and "the methods that are native to the medium".[13] Thus, we suggest a workflow that involves scraping for data collection, (visual) content analysis as a well-established method for converting visual and textual content into structured datasets for quantitative interpretation, a final analysis based on theoretical concepts like platform vernaculars or grammars, and an evaluation of the computational classifications (see Fig. 1).

Between data collection and content analysis, we need an additional preprocessing step. A single Instagram Story may contain several content layers: text overlays, spoken language, and the pivotal visual layer (compare Fig. 2). Humans can process multiple content layers at once: By being given an annotation manual and watching a video story, we quickly grasp the voice, the visuals, and embedded text elements, enabling us to categorise the video according to the manual and all layers. For computational analysis, we suggest disassembling a multimodal story into individual content layers and classifying each layer separately.[14] Disassembling occurs during the preprocessing step, where we use optical character recognition (OCR) and automated transcription software and services to turn embedded text and spoken language into plain text. This deconstruction is a technical necessity,[15] but it risks overlooking nuanced meanings that emerge from the interplay of different modalities in multimodal social media posts. At the same time, isolating and analysing each layer individually offer the advantage of reducing distractions from the whole picture, allowing for more precise annotation and classification.

The goal of our blueprint's workflow is to uncover patterns across stories. These patterns have been analysed through *platform vernaculars*[16] or *grammars of action*,[17] combining variables from content analysis with platform affordances and metadata. Thus, we see all variables as modular building blocks to be combined under a conceptual framework like the *grammars of action*, which refer to

13 Richard Rogers, Digital methods for web research, in: Emerging Trends in the Social and Behavioral Sciences, 15.5.2015, p. 1–22, doi.org/10.1002/9781118900772.etrds0076.

14 This is a suggestion: as of March 2024, multimodal abilities for large language models (LLMs) are developing rapidly. Future models might make this step obsolete.

15 As mentioned above, future multimodal models could render this deconstruction unnecessary.

16 Martin Gibbs et al., #Funeral and Instagram: death, social media, and platform vernacular, in: Information, Communication and Society 18/3 (2015), p. 255–268, doi.org/10.1080/1369118X.2014.987152.

17 For example, by Janna Joceli Omena et al., Digital methods for hashtag engagement research, in: Social Media + Society 6/3 (2020), doi.org/10.1177/2056305120940697.
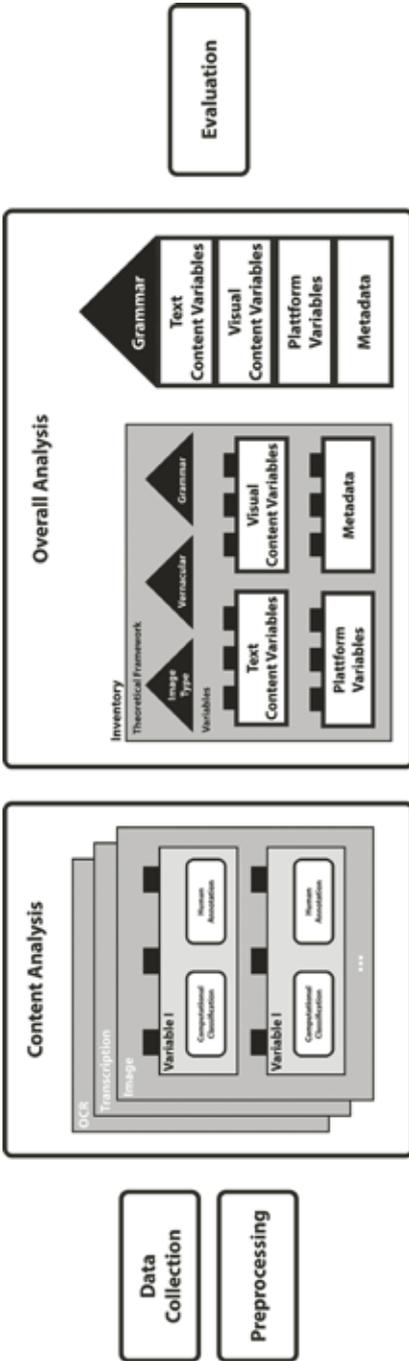
Fig. 1: Our proposed workflow for multimodal data like Instagram Stories.

Fig. 2: An illustration of content layers in Instagram Stories: To the left is the unaltered story as posted by *@gruenebayern* on 2.9.2023. The second image shows the visual layer, the third image is the embedded text layer, and the final image illustrates the sticker layer. We used Adobe Firefly to create the second image. Videos add another layer: Audio.

the underlying rules, norms, and conventions shaping how users interact with plat-form-specific features to create and consume content. This provides a structured approach to analysing behaviour and communication patterns on digital platforms.

3 A Fictional Case Study: The Martian Election

To illustrate these considerations, let us imagine the most recent elections on the planet Mars: Two parties, the red and the blue, are campaigning to win the Mar-tians' votes. The political groups heavily use Instagram Stories towards the hot phase of the election campaign. As researchers, we are curious to see how Martian election campaigns operate and, more importantly, what strategies the two parties and their front-runners adopt. Reviewing the literature on human election campa-igns, we identify two key analytical dimensions: (1) the presence and nature of poli-tical issues referenced in shared stories[18] and (2) the emphasis on political leaders within the campaign narrative.[19]

The first analytical dimension is operationalised through a dichotomous vari-able (presence vs. absence of political issues) and a categorical variable detailing the specific issues (e.g. Terraforming and Interplanetary Relations). For the second dimension, we measure concentrated visibility[20] using two variables: visibility of the front-runner (visible vs. not visible) and a categorical variable quantifying the individuals in the image (i.e. 0, 1, and 2+).

Viewing our analytical dimensions as building blocks, we can stack them together to uncover the overall pattern, following the *grammars of action* theory. Through this lens, we can identify each account's content strategies and analyse the differences between front-runner accounts and party accounts or between the two competing parties. Thus, we might ask the following research ques-tions: (RQ1) How does the presence and nature of political issues vary between front-runner and party accounts? (RQ2) How does visual individualisation vary

---

18 This research interest, including the operationalisation, is inspired by: Jörg Haßler et al., Politi-cians over issues? Visual personalization in three Instagram election campaigns, in: Information, Communication and Society 27/5 (2023), p. 815–835, doi.org/10.1080/1369118X.2023.2227684; Jörg Haßler et al., Instagram and political campaigning in the 2017 German federal election. A quantita-tive content analysis of German top politicians' and parliamentary parties' posts, in: Information, Communication and Society 26/3 (2021), p. 530–550, doi.org/10.1080/1369118X.2021.1954974.
19 Inspired by: Peter Van Aelst et al., The personalization of mediated political communication: A review of concepts, operationalizations and key findings, in: Journalism 13/2 (2012), p. 203–220, doi.org/10.1177/1464884911427802; Haßler et al., Politicians over issues?
20 Van Aelst et al. The personalization of mediated political communication.

across different parties? (RQ3) What patterns of Instagram Stories emerge in Martian elections?

To answer these questions, we start a prospective data collection over a set amount of time, e.g. 4 weeks before the election date. Once all data is collected, we start the preprocessing, recognising embedded text, transcribing video audio, and extracting image frames from videos. The results are three content layers: OCR text, transcriptions, and the visual layer. To answer RQ1, we annotate the presence of policy issues in all text layers and categorise them.[21] To answer RQ2, we look at the visual layer and annotate the presence of front-runners and the count of individuals in each image. To answer RQ3, we look at the interaction between the content variables, metadata, and platform affordances.

4 Data Collection

The data collection is the entry point of the proposed workflow. The selection of a sampling strategy precedes it. The sampling strategy depends heavily on the research interest and constitutes a different discussion. Bainotti et al., for example, searched for profiles using hashtags and observed them over a while.[22] Manovich suggests sampling along geographical locations for *Cultural Analytics*,[23] while Rose presents several sampling strategies for selecting images within a visual corpus.[24] In the context of our fictional case study, we choose front-runner and party accounts and observe them for 4 weeks.

Archiving Instagram Stories for research is more complex than archiving permanent posts, where researchers could – until August 2024 – use *CrowdTangle* for retrospective data access.[25] The situation with stories today is comparable to the

---

21 Policy issues could also be part of the visual layer, for example, in memes. Conversely, concentrated visibility might be part of the textual content layers through, i. e., mentions or quotes. Ultimately, we must decide on variables per content layer grounded in theory and preliminary corpus exploration. High-quality human annotations are labour-intensive. Hence, we suggest only focusing the annotation process on the most important variables per content layer.

22 For example, by Bainotti et al., From archive cultures to ephemeral content, and back.

23 Manovich, Cultural Analytics.

24 Gillian Rose, Visual methodologies: An introduction to researching with visual materials, 4th ed., Los Angeles 2016.

25 The tool has been replaced by Meta Content Library; researchers may apply for data access, and an independent institution processes the applications, the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan. At the time of writing, researchers receive free access. Several limitations regarding the verification status and follower count of target

time following the Cambridge Analytica scandal when Meta platforms closed down on application programming interfaces (APIs), which previously offered legal and accepted access.[26]

The lack of access leaves room for creativity: A master's thesis, for example, used screenshots for data collection.[27] This approach has its limitations: It does not scale and is feasible for only a few posts and limited data collection periods. Others used commercial tools like *4K Stogram*[28] or *StorySaver.*[29] These commercial tools are black boxes that offer image and video downloads for stories but lack metadata files. We have experimented with the selenium framework[30] and in situ annotations[31] and settled for expanding the digital method's Zeeschuimer[32] Firefox plugin to collect Instagram Stories while visiting the website using the Firefox browser. We called the project *Tidal Tales,* as we suggest collecting stories twice daily to eliminate the risk of overlaps and add a margin for error.

Tidal Tales scrapes the metadata and media files while browsing stories on the Instagram website. It intercepts the server and website communication, looking

accounts apply. We have yet to have the chance to test the new tool. See: transparency.meta.com/de-de/researchtools/meta-content-library/

26 Alessandro Caliandro, Repurposing digital methods in a post-API research environment: Methodological and ethical implications, in: Italian Sociological Review, 11/4S (2021), doi.org/10.13136/ISR.V11I4S.433; Trezza, To scrape or not to scrape; Ally McCrow-Young, Approaching Instagram data.

27 Marina Amancio, "Put it in your Story": Digital storytelling in Instagram and Snapchat Stories, 2017, www.diva-portal.org/smash/record.jsf?pid=diva2%3A1111663&dswid=-5700

28 That is, Terri L. Towner/Caroline Lego Muñoz, A Long Story Short: An Analysis of Instagram Stories during the 2020 Campaigns, in: Journof of Political Marketing 21/3–4 (2022), doi.org/10.1080/15377857.2022.2099579.

29 For example, by Bainotti et al., From archive cultures to ephemeral content, and back.

30 *Selenium* is a Python library that allows for the automation of web browsers. Used in: Michael Achmann/Christian Wolff, Policy issues vs. Documentation: Using BERTopic to gain insight in the political communication in Instagram Stories and posts during the 2021 German Federal election campaign, in: Digital Humanities in the Nordic and Baltic Countries Publications 5/1 (2023), p. 11–28, doi.org/10.5617/dhnbpub.10647.

31 We tested a *Chrome* extension to annotate stories while browsing. The extension added an overlay to the Instagram page with an annotation interface. The initial idea was not to save the actual media files but rather some metadata like time posting, username, and human annotations. This would have circumvented any issues regarding scraping and automation. We discontinued this approach due to methodological issues regarding annotation quality. See: Michael Achmann et al., Studying the ephemeral, cultures of digital oblivion. Identifying patterns in Instagram Stories, presented at the 8. Jahrestagung des Verbands "Digital Humanities im deutschsprachigen Raum, Potsdam 2022, doi.org/10.5281/zenodo.6327900.

32 Stijn Peeters, Zeeschuimer (v1.8.0), doi.org/10.5281/zenodo.8399900 For the use of the tool in other contexts also see the text by Marloes Geboers and Elena Pilipets in this volume.

for data concerning stories. The *Firefox* plugin saves all metadata of each story item to a local database and triggers the video and image download in the background. Once all stories have been viewed, the metadata can be saved as a file with comma-separated values (CSV) with one row per story and several columns that include information on stickers and the filename of the corresponding image and video files.

Practically speaking, researchers identify accounts of interest, create a new Instagram profile, follow these "target" accounts, and view the stories of these accounts twice a day in *Firefox* with the *Tidal Tales* plugin installed. These visits may be automated using *Selenium*. The software and manuals are available at tidaltal.es.

The tools referenced, including *Tidal Tales*, utilise data scraping. Marres and Weltevrede define scraping as an automated process for collecting and extracting data from online sources.[33] This typically involves downloading the source code of web pages and identifying specifically structured data fields in an automated process. It is important to note that data scraping is subject to varying legal regulations across different jurisdictions. In Germany and the European Union, certain exceptions permit non-profit researchers to engage in data mining for analytical purposes under specific conditions.[34] Researchers must consult local laws and institutional guidelines to ensure compliance with all relevant legal and ethical standards.[35]

4.1 Data Management

For a successful social media analysis, systematic data management is vital. Following a computational approach and decomposing the content layers fuels this requirement as we work with metadata, textual, and visual data. Each story consists of up to three elements after the data collection: One image, video, and metadata file. When using *Tidal Tales,* the media files are stored in a predefined directory structure: `{username}/{ID}.jpg`. The metadata is exported as a CSV file, repre-

---

33 Noortje Marres/Esther Weltevrede, Scraping the social? Issues in real-time social research, in: Journal of Cultural Economy 6/3 (2013), doi.org/10.1080/17530350.2013.772070.

34 See Article 3, Directive (EU) 2019/790 (eur-lex.europa.eu/eli/dir/2019/790/oj) and §60d *Urheberrechtsgesetz* (German copyright law).

35 Ethical and privacy guidelines often involve the pseudonymisation or anonymisation of scraped research data, see for example: Franzke et al.: Internet research: ethical guidelines 3.0: association of internet researchers.

Fig. 3: The Tidal Tales Plugin is a Firefox extension. Once installed, the plugin saves the metadata of each story item locally, while the associated images and videos are downloaded directly to the computer's disk. After data collection, the metadata can be exported as a CSV file. Each entry in the CSV includes the file names corresponding to the downloaded media files.

senting one story by one row. Each row contains an `ID` and `username` value that is pivotal for mapping the metadata with the correct media files (compare Fig. 3).

Tables, as represented by CSV files, are central to our workflow. In the tutorial notebooks on our website, we use *pandas*, a data analysis library for *Python*. CSV files can be imported as *pandas DataFrames*, and *DataFrames* can be exported as CSV files. For a manual analysis without technical knowledge, the concept of Data-Frame can be transferred to *Excel*: Each table is a *Sheet* within one *Excel* file. We follow the concept of *Tidy Data*, where "each variable forms a column, each observation forms a row, and each type of observational unit forms a table".[36] Translated to Instagram Stories, we suggest a total of three tables for (1) Stories (Metadata), (2) Text, and (3) Images.[37] Within each table, we add one column after another. Each column represents a variable (see Fig. 4).

Story ID	Username	Text	Text Type	Text ID	Sentiment	...
1	Galaxy.Party	Lorem Ipsum	OCR	123	Positive	
1	Galaxy.Party	dolor sit amet	Transcription	456	Negative	
2	Alien.Alliance	consetetur sadipscing	Transcription	523	Neutral	

Fig. 4: The first line of a text table. Following the tidy data principle, each row corresponds to one text and each column to one variable. We recommend adding a unique Text ID to identify each row and keeping the Story ID and Username across all content tables to connect text rows to story metadata.

We recommend using the story ID and the username across all (three) tables to identify the relation between a single story and multiple rows in the text/image table related to this story. Using the ID and username allows us to reassemble everything at the end of the project.[38] To keep the references intact, we recommend using a systematic folder and filename structure (as suggested by *Tidal Tales*), optionally adding a column with the file location to the metadata table.

---

36 Hadley Wickham, Tidy Data, in: Journal of Statistical Software 59/10 (2014), doi.org/10.18637/jss.v059.i10.

37 One might argue that we are missing videos. We will, however, decompose videos to text and images, see below.

38 Our proposal assumes the analysis of public figures' Instagram accounts. In most other cases, researchers should pseudonymise or anonymise the username column. In Germany, for example, researchers can share data for peer review, see *Urheberrechtsgesetz* §60d (4). Pseudonymisation or anonymisation becomes even more important when sharing data.

4.2 Preprocessing

Instagram Stories often contain embedded text, and videos additionally contain audio. We need to transform both into computer-readable text. From a technical point of view, embedded text is not directly accessible: The images and videos posted as stories are raster graphics. While digitally created, these images are similar to a photo of a newspaper article: The text is not accessible without OCR. Thus, OCR is our first preprocessing step. Our website provides tutorials using two approaches: Using *Python* and the easyocr package, or commercial providers like *Google Vision AI* using *Memespector*.[39]

Our second recommended preprocessing step is the automated transcription of videos. This step involves *OpenAI's Whisper* model, introduced in September 2022. This model shows high accuracy. Our experiments, however, revealed one major disadvantage of *Whisper*: hallucinations. Sometimes, the model added additional content when the voices were not audible. The transcriptions can be controlled efficiently using human annotators and *Label Studio* (see Section 6), if needed.

Additionally, videos contain multiple image frames. Story videos can be up to 60 s long (as of February 2024). Assuming a video frame rate of 30 frames per second, we could extract up to 1,800 images per story item. To save processing time, distant viewing[40] approaches rely on an extraction strategy, i.e. one frame per second,[41] or using more sophisticated approaches to extract specific keyframes.[42] Since the Instagram app limits embedded text and stickers to be added only once per video, we suggest focusing on one frame per video: The *Tidal Tales* export contains a cover image provided by Instagram for each story.[43]

Overall, the preprocessing step added up to two rows per story to our text table – one for OCR and one for transcriptions. In our experience, both steps are sometimes overambitious and recognise unimportant text elements like street

---

39 Jason Chao, Memespector-GUI: Graphical User Interface Client for Computer Vision APIs (2023), doi.org/10.5281/zenodo.7704877.

40 Taylor Arnold/Lauren Tilton, Distant viewing: analyzing large visual corpora, in: Digital Scholarship in the Humanities 34/Supplement_1 (2019), p. i3–i16, doi.org/10.1093/llc/fqz013.

41 Alina El-Keilany et al., Distant viewing of the Harry Potter movies via computer vision, in: Karl Berglund/Matti La Mela/Inge Zwart (Ed.), Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries Conference (DHNB 2022), p. 33–49, ceur-ws.org/Vol-3232/paper03.pdf

42 Compare the following review for several approaches: Milan Kumar Asha Paul et al., Key-Frame Extraction Techniques: A Review, in: Recent Patents on Computer Science 11/1 (2018), p. 3–16, doi.org/10.2174/2213275911666180719111118.

43 This approach, while imperfect, reduces the effort. Applying OCR across multiple frames could capture any text added through external video editing software or subtitles. At the same time, we would have to compare the OCR results to deduplicate our text corpus.

signs or hallucinations in inaudible situations. Although these limitations might not significantly alter the analysis outcomes, acknowledging them is crucial for accurately interpreting our results.

5 Content Analysis

We are ready for the actual content analysis with all content layers prepared. As the workflow section outlines, we suggest considering each layer by itself to facilitate computational analysis.

One common approach in social media research is quantitative content analysis.[44] This method offers a well-established approach towards converting unstructured textual and visual data into a structured format, which can be used for statistical analysis and description.[45] Rose provides a streamlined version of the necessary steps for visual content analysis:[46] 1. Finding your images, 2. Devising your categories for coding, 3. Coding the images, and 4. Analysing the results. This chapter discusses steps 3 and 4: We have already selected and collected the image corpus, and we are reusing categories from the literature (see operationalisation above).[47]

We suggest a computational approach, followed by a validation of automated classifications,[48] detailed in the upcoming section. We end each classification

---

[44] That is, most studies on political communication: Jennifer Bast, Politicians, Parties, and Government Representatives on Instagram: A Review of Research Approaches, Usage Patterns and Effects, in: Review of Communication Research 9 (2021). Additionally, Milanesi and Guercini reviewed several studies using visual content analysis: Matilde Milanesi/Simone Guercini, Image-based Social Media and Visual Content Analysis: Insights from a Literature Review, in: Micro & Macro Marketing 3 (2020), p. 537–558, ideas.repec.org/a/mul/jyf1hn/doi10.1431-97640y2020i3p537-558.html. For a theoretical grounding and more information about the method, see: Philipp Mayring, Qualitative content analysis: theoretical foundation, basic procedures and software solution, nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

[45] For detailed instructions see: Nicola Döring/Jürgen Bortz, Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften, Berlin 2016, doi.org/10.1007/978-3-642-41089-5.

[46] Gillian Rose, Visual Methodologies: An Introduction to Researching with Visual Materials, 4th Ed., London 2016.

[47] Our website provides several approaches for text and visual exploration helpful for developing custom codes and explorative analyses.

[48] Christian Baden et al., Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda, in: Communication Methods and Measures 16/1 (2022), doi.org/10.1080/19312458.2021.2015574; Yilang Peng et al., Automated Visual Analysis for the Study of Social Media Effects: Opportunities, Approaches and Challenges, in: Communication Methods and Measures 18/2 (2023), p. 163–185, doi.org/10.1080/19312458.2023.2277956.

example with a suggestion for the actual analysis and interpretation of the results. Additionally, each content variable is a building block for an overall analysis of patterns across stories. The computational steps below may be replaced with human annotation projects for manual content analysis.

5.1 Textual Content Analysis and Text as Data

As noted by Bengfort, Bilbro, and Ojeda, "language is *unstructured* data that has been produced by people to be understood by other people".[49] This characterisation of language highlights its contrast with *structured* or *semi-structured* data. As unstructured data, language is not readily interpretable by computers and requires more advanced techniques for analysis. In the context of Instagram Stories, *Tidal Tales* metadata exports contain *structured data* columns such as "User Name", "Time of Posting", or "Stickers". These pieces of data are quantifiable and can be quickly sorted, filtered, or counted, e.g. using tools like *Excel* or *pandas*. For instance, we can quickly determine the most active users by counting the rows associated with each username.

In contrast, *unstructured* data is not organised in a predefined manner and is typically more challenging to process and analyse. In the context of stories, the OCR and transcription results from the preprocessing are unstructured data. These texts, composed of paragraphs or sentences, require different analytical approaches to extract meaningful insights. Unlike structured data, we cannot simply count or sort these texts straightforwardly. In our context, we often refer to the texts we analyse as a "corpus". Each text is called a "document" and makes up one row in our text table (see Section 4.1 and Figure 6). Each document can be broken down into smaller units known as "features". Features can be words, phrases, or even patterns of words, which we then use to quantify and analyse the text.[50]

Let us return to our analysis of the Martian elections. A computational content analysis can help answer the first research question: How do the presence and nature of political issues vary between front-runner and party accounts? Our operationalisation stipulates two classifications: One binary classification for the presence or absence of policy issues and one multi-class classification to determine the actual policy issues a given text mentions. We propose to use GPT-4 via API, a large language model (LLM) by OpenAI, for both tasks. First, we would obtain the binary

---

49 Benjamin Bengfort et al., Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning, Beijing 2018.
50 Compare Mario Haim, Computational Communication Science: Eine Einführung, Wiesbaden 2023, p. 230.

classification for **each text document**. Next, we can assign the policy issues mentioned within each text document. We send one request per classification and text document to the GPT API. Each request consists of a text document and a so-called prompt, natural language instructions for the classification task (cf. Fig. 5 for an example).

You're an advanced classifying AI.

Objective: Determine the presence of direct references, discussions, implications, or suggestions concerning specific policy issues, governance measures, or legislative details in German language social media texts.

Instructions:
1. Return 'True' if the text directly, indirectly, or implicitly mentions, discusses, suggests, or indicates a specific policy, governance measure, or legislative detail. This includes concrete proposals, changes, or calls to action related to governance.
2. Return 'False' if the text merely references general political events, activities, schedules, or generic statements that do not elaborate on a specific policy, governance measure, or legislative detail.
3. In cases of uncertainty, or if a statement is ambiguous without clear policy details, return 'False'.

Formatting: Return only True or False.

Fig. 5: A prompt example for classifying the presence or absence of policy issues in a given text. The prompt above instructs the large language model to return the binary classification as True or False.

In contrast to machine learning approaches, LLMs have a zero-shot classification ability, meaning that in some scenarios, their classifications reach a satisfactory quality through nothing more than a proper prompt. In some scenarios, we can improve the classification quality by adding a small set of examples to the prompt. This approach is called few-shot prompts.[51] In contrast, a custom classification model using machine learning would require more training data. To improve the classification quality, we can use the emerging field of prompt engineering, which explores the creation of questions and inputs so the model can more easily find the right information.[52] Törnberg provides an overview of all steps towards automated

---

51 Tom B. Brown et al., Language Models are Few-Shot Learners, in: arXiv Computation and Language (28.5.2020), arxiv.org/abs/2005.14165
52 Jindong Gu et al., A Systematic Survey of Prompt Engineering on Vision-Language Foundation Models, in: arXiv Computer Vision and Pattern Recognition (24.7.2023), arxiv.org/abs/2307.12980

text annotation using LLMs including prompt engineering.[53] It is based on the understanding that different questions can produce more or less accurate results, so adjusting the format and examples of the prompt is key to getting the best results.[54] The field of prompt engineering involves different ways of making these prompts. One can decide to create prompts manually or use automated methods.[55]

Each text document's classification results (=response of the language model) are added to the corresponding row in our text table. We add one column for each content variable, for example, *Positioning* and *Policy Issues*. In preparation for the analysis and to answer our research questions, we can add the columns *OCR Positioning, Transcript Positioning*, and *Policy Issues* to the overall story table and fill them with the values from the text table (see Fig. 6). We can then determine policy issues' overall presence or absence in a new *Positioning* column based on these columns. Assuming the story table contains a column *Account Type*,[56] we can check for significant differences between the appearance of policy issues between front-runners and party accounts using, e.g. a χ^2 test.

This section highlighted computational content analysis for text based on operationalisations from the literature and used LLMs like GPT. The computational classifications may be replaced by human annotations for manual content analysis or by other models and services for automated analysis. Next, we will look at the visual layer using visual content analysis.

5.2 Visual Content Analysis and Images as Data

Within the computational social science community and their neighbouring disciplines, the concept of *Images as Data* is currently being established,[57] equivalent

53 Petter Törnberg, Best Practices for Text Annotation with Large Language Models, in: arXiv (5.2.2024), arxiv.org/abs/2402.05129

54 Zihao Zhao et al., Calibrate Before Use: Improving Few-shot Performance of Language Models, in: Marina Meila/Tong Zhang (ed.), Proceedings of the 38th International Conference on Machine Learning, s.l. 2021, proceedings.mlr.press/v139/zhao21c.html

55 Pengfei Liu et al., Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing, in: ACM Computing Surveys 55/9 (2023), S. 1–35, doi.org/10.1145/3560815.

56 Based on a manually curated list categorising accounts (usernames) as front-runners and party accounts.

57 Peng et al., Automated Visual Analysis for the Study of Social Media Effects; Jungseock Joo/Zachary C. Steinert-Threlkeld, Image as data: Automated content analysis for visual presentations of political actors and events, in: Computational Communication Research 4/1 (2022), doi.org/10.5117/ccr2022.1.001.joo.
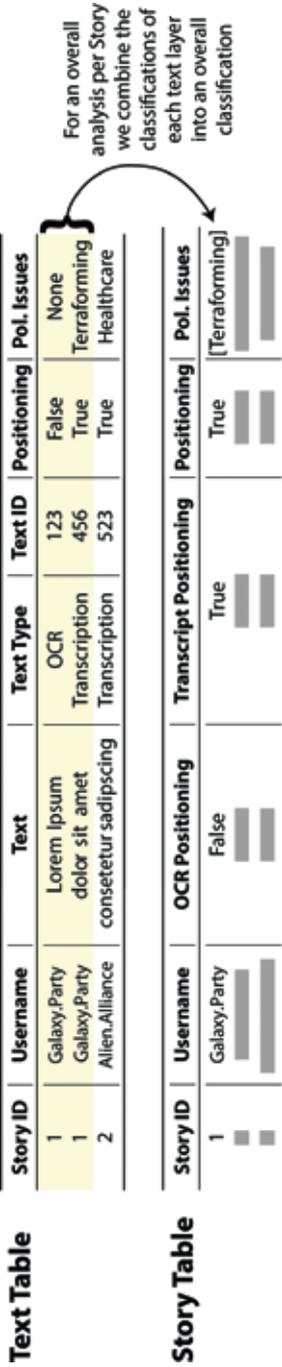
Text Table

Story ID	Username	Text	Text Type	Text ID	Positioning	Pol. Issues
1	Galaxy.Party	Lorem Ipsum dolor sit amet	OCR	123	False	None
1	Galaxy.Party	consetetur sadipscing	Transcription	456	True	Terraforming
2	Alien.Alliance		Transcription	523	True	Healthcare

For an overall analysis per Story we combine the classifications of each text layer into an overall classification

Story Table

Story ID	Username	OCR Positioning	Transcript Positioning	Positioning	Pol. Issues
1	Galaxy.Party	False	True	True	[Terraforming]

Fig. 6: Once all text rows have been classified, we can merge the data from the text table with the story table. Looking at the positioning variable across multiple text layers, we can determine an overall absence or presence of policy issues for each story item.

to the *Text as Data* paradigm. From a Digital Humanities perspective, Arnold and Tilton have developed a concept of *Distant Viewing*.[58] Both concepts deal with the theory and application of visual analyses using computational methods, with very similar ends: As with computational text analyses, computational visual analyses promise to open a new, quantitative perspective incorporating volumes of data unfeasible for human processing.

Peng et al. regard images as data within the context of social media effect studies.[59] They focus on turning visual content into quantifiable insights using computational methods, allowing for data-driven analysis. They also stress the importance of validating computational classifications and identifying biases in automated visual analysis. Arnold and Tilton underline the significance of structured annotations in this process, as they play a crucial role in decoding and interpreting visual data. They also encourage users to be mindful of the biases and power dynamics inherent in the way we perceive images and the technological tools used for this purpose.[60]

In our fictional case study, visual analysis is part of the second research question on visual individualisation, which we operationalise through the dichotomous variable front-runner visibility and a categorical variable quantifying the individuals in the image. We suggest face recognition using the deepface[61] package for Python to identify front-runners in images and multimodal GPT-4 to count the individuals.

Identifying the politicians consists of two steps: face detection and face matching. The deepface library provides a unified interface to several models for both tasks. At first, we localise all faces in each image using the detection model. In the second step, we compare the recognised faces with one or more target images of our politicians. We can decide whether an image shows the target politician using the resulting similarity measure and a threshold. This result can be saved (as *True* or *False*) to a new column, *Front-Runner,* in our image table.

Counting the individuals using the multimodal GPT-4 model is similar to the text classification above. We craft a prompt that describes our goal, send each image to the API, and receive a classification (cf. Fig. 7 for an example prompt). We can save the result to the *Person Count* column of the image table.

---

58 Taylor Arnold/Lauren Tilton, Distant Viewing: Computational Exploration of Digital Images, Cambridge 2023.

59 Peng et al., Automated Visual Analysis for the Study of Social Media Effects.

60 Arnold/Tilton, Distant Viewing.

61 Sefik Ilkin Serengil/Alper Ozpinar, LightFace: A Hybrid Deep Face Recognition Framework, in: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), Istanbul 2020, p. 1–5, doi.org/10.1109/ASYU50717.2020.9259802.

Face Recognition Result

Prompt for Person Count


```
Imagine you are a social science researcher, conducting an
analysis of the 2021 German Federal election campaign. The
image shows the front-runner: {NAME}. We're interested in
measuring the individualization of candidates in the
election campaign.

Assess the image and provide the following information:

1. How many people are in the focus of the image?
   Select the right choice: 'GPT Count': ["0", "1", "2", "3+"]

2. Is there a crowd of spectators visible?
   Select the right choice: 'GPT Crowd': ["True", "False"]

Respond in valid JSON only.
```

Model Response

```
{
    "GPT Count": "1",
    "GPT Crowd": "True"
}
```

Fig. 7: To the left, an example of face matching results: The face detection model detected the face illustrated by the yellow frame. Using face matching, we can identify the name of the politician. To the right is an example of a face count prompt using a multimodal LLM and its response for the left-hand picture. Image source: Generated using Midjourney v6.1.

Since we only deal with a single visual layer,[62] we can answer RQ2 directly through the image table, assuming it contains a column *Party*. As above, we use a χ^2 test to check for significant differences in visual personalisation in Instagram Stories between the two Martian parties.

In this section, we introduced two automated approaches for visual analysis: face recognition and the use of large multimodal language models. Once more, the same analysis can be conducted manually, using human annotations. Peng et al. list many more models and technologies available for computational image analysis.[63]

---

62 When working with posts we often deal with multiple layers, similar to the text layers, as posts can contain multiple images in a gallery. Similarly, we could split videos into frames and view each frame as one layer of a story. In this case we could proceed similarly to the text layers and create an overall content variable combining all layers.

63 Peng et al., Automated Visual Analysis for the Study of Social Media Effects.

5.3 Overall Analysis

With the content variables prepared, we can answer the final research question: What Instagram story patterns emerge in Martian elections? We see all variables as building blocks to answer this question: The textual and visual content variables, the metadata and the platform affordances. Previous studies, for example, have observed these patterns through the lens of *grammars of action*.[64] Originally linked to research on privacy,[65] *grammars of action* have been extended into the domain of social media analysis, embodying the rules and norms governing content creation and consumption. They provide a framework for understanding how users, through their engagement with platform-specific affordances, produce, and navigate digital content. For instance, in analysing Instagram Stories, especially during the Martian elections, this approach enables us to identify recurring patterns and themes in political actors' communication, influence, and mobilisation efforts through stories. Furthermore, by employing *grammars of action*, we can decode the subtleties of digital communication, uncovering not only the content but also the manner and potential impact of these patterns on public discourse and the shaping of opinions. Utilising these grammars allows us to dissect the complex interplay of textual, visual, and metadata elements, offering a comprehensive understanding of social media's role in the Martian electoral processes.

Building on top of the variables from RQ1 and RQ2, we can discover grammars using the *k*-means algorithm. This clustering technique groups similar data points together based on a set of variables. Merging all content variables from the textual and visual analysis with the metadata table, we create a final table that contains all variables of interest (=building blocks). Once the clusters, or patterns, emerge, we can interpret them considering theory and related work. Alternatively, we can follow a manual approach using descriptive statistics and visualisations to discover patterns within our data.

Similar to grammars, Towner and Muñoz used an affordance perspective to analyse Instagram Stories, mapping content types to affordances.[66] Liebhart and Bernhardt introduced Image Types as an analytical framework to explore visual

---

64 Bainotti et al., From archive cultures to ephemeral content, and back; Carolin Gerlitz/Bernhard Rieder, Tweets are not created equal: Investigating Twitter's client ecosystem, in: International Journal of Communication Systems 12 (2018), p. 528–547, pure.uva.nl/ws/files/23266519/5974_30096_2_PB.pdf; Omena, Digital Methods for Hashtag Engagement Research.
65 Philip E. Agre, Surveillance and capture: Two models of privacy, in: The Information Society 10/2 (1994), p. 101–127, doi.org/10.1080/01972243.1994.9960162.
66 Towner/Muñoz, A Long Story Short.

themes across Instagram posts,[67] a concept that has been extended by Haßler et al.[68] These image types are similar to grammars as they combine textual and visual variables into one overarching categorisation for overall posts. Using platform vernaculars, Gibbs et al. examined the unique styles and logic influenced by platform features and user habits. When focusing on vernaculars or image types, our workflow allows us to selectively concentrate on specific building blocks that align with our framework's requirements. For instance, we might exclude affordances and metadata, opting instead to focus on a tailored combination of content variables specific to image types.

We can now answer all research questions for our fictional case study. However, we still need an important part of computational content analyses: external validation. Both textual and visual content analyses require an evaluation against human annotations to prove the automated analysis validity.

6 Evaluation

In computational social media analysis, validation is crucial for ensuring the accuracy and reliability of text analysis methods. As highlighted by Birkenmaier et al., validation entails both internal and external processes.[69] Internal validation assesses a model's plausibility and quality within the data context. It tends to rely on researchers' judgement, whereas external validation compares model outputs with external benchmarks such as human-annotated labels. Baden et al. further emphasise the significance of continuous evaluation and transparent reporting of validation steps and results.[70] We focus on external validation through human annotations using Label Studio for the present blueprint.

---

67 Karin Liebhart/Petra Bernhardt, Political storytelling on Instagram: Key aspects of Alexander Van der Bellen's successful 2016 presidential election campaign, in: Media and Communication 5/4 (2017), p. 15–25, doi.org/10.17645/mac.v5i4.1062.

68 Jörg Haßler et al., Influence of the pandemic lockdown on Fridays for Future's hashtag activism, in: New Media & Society 25/8 (2021), doi.org/10.1177/14614448211026575.

69 Lukas Birkenmaier et al., The Search for Solid Ground in Text as Data: A Systematic Review of Validation Practices and Practical Recommendations for Validation, in: Communication Methods and Measures 18/3 (2024), p. 249–277, doi.org/10.1080/19312458.2023.2285765.

70 Baden et al., Three Gaps in Computational Text Analysis Methods for Social Sciences.

6.1 Creating Annotations with Label Studio

Collecting human annotations is essential for automated and manual content analysis. For computational content analysis, we aim to gather annotations on a subsample to validate our classification externally. For manual analysis, the entire corpus requires labelling. Regardless, multiple annotators are essential for quality assurance. For computational classifications, the standard practice involves collecting three annotations and adopting the majority rule for ground truth. Alternatively, expert review or annotator discussions may resolve discrepancies.

Annotations can be organised using *Excel* or specialised software like *MAXQDA*, a tool that supports qualitative data analysis. However, we recommend *Label Studio*, an open-source tool compatible with diverse data types, including Instagram Stories. Its features include programmable data input, a customisable interface, and support for multiple annotators in its enterprise version (see Fig. 6 for an example). We provide a detailed tutorial on our website for the automated creation of *Label Studio* projects and labelling interfaces. The tutorial includes several notebooks for calculating evaluation measures and their reporting.
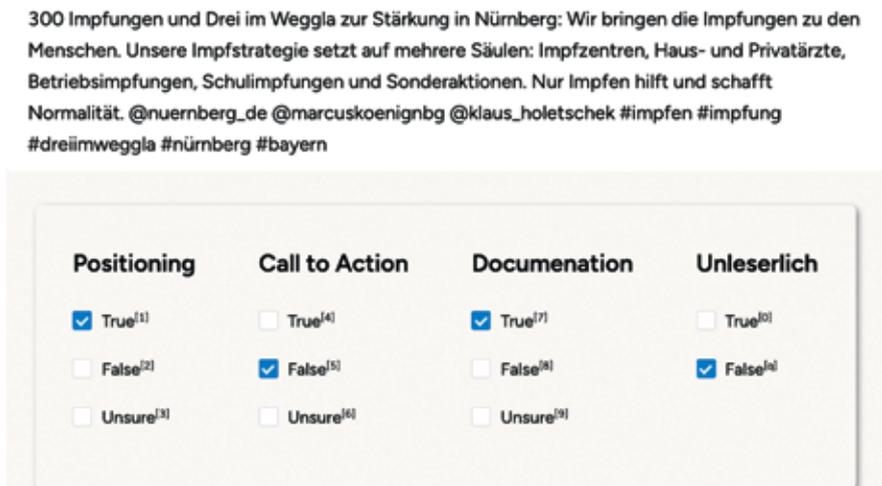
300 Impfungen und Drei im Weggla zur Stärkung in Nürnberg: Wir bringen die Impfungen zu den Menschen. Unsere Impfstrategie setzt auf mehrere Säulen: Impfzentren, Haus- und Privatärzte, Betriebsimpfungen, Schulimpfungen und Sonderaktionen. Nur Impfen hilft und schafft Normalität. @nuernberg_de @marcuskoenignbg @klaus_holetschek #impfen #impfung #dreiimweggla #nürnberg #bayern

Positioning	Call to Action	Documenation	Unleserlich
☑ True[1]	☐ True[4]	☑ True[7]	☐ True[0]
☐ False[2]	☑ False[5]	☐ False[8]	☑ False[a]
☐ Unsure[3]	☐ Unsure[6]	☐ Unsure[9]	

Fig. 8: A screenshot of a Label Studio annotation interface from a text annotation study. Participants annotated several dichotomous text content variables at once.

Label Studio's flexible interfaces allow adaptation to specific project needs. For manual analysis, a detailed interface can capture numerous variables simultaneously. In computational analysis, simpler non-expert annotations suffice for external validation. We advise setting up separate projects for different analy-

sis layers (OCR, transcription, visual) and limiting annotations to a few variables simultaneously. This approach enhances focus, facilitates quality control, and suits crowd-sourced annotations on platforms like *Mechanical Turk*.[71]

A well-laid-out and tested annotation manual is pivotal for collecting consistent, high-quality annotations. Niels Reiter provides an iterative process towards creating an annotation manual: Through several loops and test annotations, we keep improving the manual to reach an optimal interrater agreement for each variable.[72]

6.2 Measures and Reporting

Once the annotation process ends, we look at the interrater agreement between the human annotations to check the quality of our annotation dataset. We download the dataset and look at each variable's annotations. Returning to the concept of tables, we want to create a table where each annotated item (i.e. text or image) occupies precisely one row, and each annotator occupies one column. Using this format, we can easily calculate the agreement using standard measures. For multi-annotator scenarios, Krippendorff's alpha is a good choice. To calculate the score, we look at all annotator columns and make use of an open-source Python library.[73] We need to interpret the alpha score to determine the quality of our annotation dataset through interrater agreement. A Krippendorff's alpha of 0.8 or above is often considered an indication of solid agreement among annotators. However, in contexts where even slight discrepancies can have significant implications, aiming for an alpha closer to 1.0 is advisable.[74] A lower threshold may be acceptable for exploratory studies or preliminary research.[75]

---

71 Mechanical Turk and similar tools are online platforms where people are paid small amounts to complete tasks like annotations.
72 Niels Reiter, How to Develop Annotation Guidelines, in: nilsreiter.de (blog), 1.10.2017, nilsreiter. de/blog/2017/howto-annotation
73 grrr/bryant1410, krippendorf-alpha, github.com (18.9.2017), github.com/grrrr/krippendorff-alpha
74 See Silviu Paun et al., Using Agreement Measures for CL Annotation Tasks, in: Silviu Paun et al. (ed.), Statistical Methods for Annotation Analysis, Cham 2022, p.47–78, doi.org/10.1007/978-3-031-03763-4_3. They provide more information on the use of alpha on their website: sites.google.com/view/arrau/the-alpha-resources-page
75 Some computational analyses in social science contexts have reported far lower agreement levels, e.g. the average Krippendorff's alpha for Song et. al.'s literature review is 0,549 (our own calculation based on their published data, N=4). For pit-falls and problems of low gold standard data see their article: Hyunjin Song et al., In Validations We Trust? The Impact of Imperfect Human

Corresponds to the
ID in the text table.

Each table contains one
row per annotated text,
and one column per
annotator.

Text ID	Ann. 1	Ann. 2	Ann. 3
3733	True	True	True
4185	False	True	True
530	True	False	True
2721	False	False	False
3384	False	False	False

Fig. 9: This screenshot illustrates what an annotation table for one variable exported from Label Studio looks like. The unique Text ID is important to map the annotations to the computational classifications. One row corresponds to one text, one column to one annotator, and one table corresponds to one annotated variable, for example, Positioning.

Having established the quality of the human annotations, we can move forward to evaluate the annotation quality of our computational approach. This evaluation aims to compare the computational classification for each item and variable with the ground truth data and to calculate performance metrics for our model (or prompt). These metrics include accuracy, precision, recall, and the $F1$ score. The $F1$-score incorporates the precision (how many positively classified items are actually positive), and the recall (how many of the actually positive items were classified as positive), and higher $F1$-scores suggest a good balance between high precision and high recall. We use the same table structure as above, look at all human annotations per row, and add a Gold Standard (ground truth) column. We fill the column through a majority vote across the annotators' responses. In the case of categorical data, we might have to add an expert review or discussions to resolve ties. Once the gold standard has been established, we add the computational annotations to the table and calculate the performance metrics, comparing the gold standard column with the computational classification column.

The „Majority Decision" column is filled through a majority vote across the three annotators.

This column is filled with the results of the computational classification

Text ID	Ann. 1	Ann. 2	Ann. 3	Majority Decision	Model Classification
3733	True	True	True	True	True
4185	False	True	True	True	True
530	True	False	True	True	False
2721	False	False	False	False	False
3384	False	False	False	False	False

Fig. 10: A new Majority Decision column is added based on the majority vote. Additionally, we join the computational classification results with the annotation table. We can use this structure to calculate the model's/prompt's performance metrics, comparing the Majority Decision to the Model Classification across all rows. The table above represents the Positioning variable.

Let us return to the Martian election to illustrate the process: We classified the presence and absence of policy issues in OCR and transcription texts for RQ1. To evaluate the classification quality, we draw a random sample of 30% across all text documents (=rows in the text table). We create a Label Studio project with a simple annotation interface: Coders must decide between "True", "False", and "Unsure" using checkboxes to answer the following question: "Does the text displayed above contain any references to policy issues?". We recruit three annotators. They receive an extensive manual outlining the process and precisely describing their task. The annotation manual has been tested beforehand.

Once all three coders have annotated each document, we can export the results from Label Studio (see Fig. 9). Afterwards, we import the JSON file with the results into a *Jupyter* Notebook and calculate the interrater agreement. We can deduct the gold standard from the annotations if they are of sufficient quality (see Fig. 10). In case of insufficient quality, we must fix our annotation manual and start again. Using the gold standard data, we can calculate the performance metrics for our prompt and determine whether the quality is sufficient for our actual analysis to rely on the computational classifications. We report quality measures, human agreement, and the model's performance.

The annotation process needs to be repeated for each variable.[76] Thus, we also need to collect human codings for the categorical policy issues variable, meaning that we need humans to identify politicians in images and to count the people pictured. We would follow the same annotation process when working with machine learning models that need training data.

When working with LLMs, we should reserve a small portion of the annotated data for prompt engineering. Törnberg suggests a process similar to creating an annotation manual: going back and forth between automated coding and improving the prompt until reaching a desired performance. The final evaluation of the prompt uses the annotated ground truth data minus the random sample used during prompt engineering.[77]

7 Conclusion

This chapter presented a modular blueprint for the computational analysis of ephemeral Instagram Stories. Our blueprint provides a flexible framework for social media analysis that, while practical, opens the door to more extensive theoretical and methodological exploration. Those interested in implementing these methods are encouraged to visit our website,[78] where practical tutorials and code are available for collecting and analysing Instagram data. Additionally, we introduce computational techniques for exploring multimodal data, emphasising their applicability beyond Instagram Stories to other content formats and platforms, such as Facebook posts.

Our approach emphasises using LLMs, which we believe represent cutting-edge technology highly accessible to scholars with limited technical backgrounds. The key advantage of LLMs in the humanities lies in their ease of use – unlike traditional machine learning models, which often require substantial expertise to train and fine-tune, LLMs can be used with minimal technical skills. This lowers the barrier for historians and scholars, allowing them to experiment with computational methods without extensive technical training. Consequently, LLMs offer a promising entry point into quantitative research. However, while these models provide a strong foundation, historians would benefit from a basic understanding

---

76 It is also possible to collect annotations for several variables simultaneously. We found that the quality of annotations increased when concentrating on a single task at a time – with student annotators.

77 Törnberg, Best Practices for Text Annotation with Large Language Models.

78 Michael Achmann-Denkler, Notes on Computational Social Media Research, social-media-lab.net

of data science principles such as data management and interpretation. Incorporating these topics into curricula might help scholars navigate the preparation of data and the interpretation of results without the need for programming knowledge, which might otherwise feel overwhelming.

We also proposed a workflow for prospective data collection. *Tidal Tales* is compatible with story highlights, facilitating access to expired stories. These highlights constitute a gap in the current research as an intermediate between ephemeral and permanent post types. Additionally, researchers might collect stories using data donations, enabling retrospective data collection combined with informed user consent.

Though we focused on Instagram and political communication, the blueprint is adaptable to other visual platforms, such as TikTok, and has potential applications in various fields. For instance, in a museum context, researchers could collect visitor-shared social media posts to better understand preferences and experiences. Here, LLMs could classify content, identify popular objects, or analyse sentiment in user feedback, providing insights into visitor engagement.[79]

Despite its advantages, our approach has limitations. General-purpose LLMs like GPT-4 are versatile, but specialised models may outperform them if fine-tuned for narrow tasks.[80] Furthermore, reliance on external servers for processing, as in the case of OpenAI's GPT, raises concerns about data privacy. Using open-source models like Llama 3[81] and hosting them locally would be preferable.[82] It offers greater control but requires substantial hardware resources for larger, more capable models.

Our blueprint is a conceptual guide to inspire future research that bridges social media studies, computational analysis, and humanities. By demonstrating the potential of LLMs, we hope to encourage scholars to explore data-driven approaches and adapt them to their research needs.

---

79 Bo-A, Rhee et. al., Analyzing the museum experience through the lens of Instagram posts, Curator: A Quarterly Publication of the American Museum of Natural History 64/3 (2021), p. 529–547, doi.org/10.1111/cura.12414.

80 For example, our call to action classification model based on GBERT: Michael Achmann-Denkler et al., Detecting Calls to Action in Multimodal Content: Analysis of the 2021 German Federal Election Campaign on Instagram, in: arXiv (4.9.2024), arxiv.org/abs/2409.02690

81 ai.meta.com/blog/meta-llama-3/

82 Törnberg, Best Practices for Text Annotation with Large Language Models.

Bibliography

Achmann, Michael/Hampel, Lisa/Asabidi, Ruslan/Wolff, Christian: Studying the Ephemeral, Cultures of Digital Oblivion, DHd 2022 Kulturen des digitalen Gedächtnisses. 8. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum" (DHd 2022), Potsdam 2022, doi.org/10.5281/zenodo.6327900 (6.9.2024).

Achmann, Michael/Wolff, Christian: Policy Issues vs. Documentation: Using BERTopic to Gain Insight in the Political Communication in Instagram Stories and Posts During the 2021 German Federal Election Campaign, in: Digital Humanities in the Nordic and Baltic Countries Publications 5/1 (2023), p. 11–28, doi.org/10.5617/dhnbpub.10647 (5.6.2024).

Achmann-Denkler, Michael/Fehle, Jakob/Haim, Mario/Wolff, Christian: Detecting Calls to Action in Multimodal Content: Analysis of the 2021 German Federal Election Campaign on Instagram, in: arXiv [cs.SI] (4.9.2024), arxiv.org/abs/2409.02690 (6.9.2024).

Aelst, Peter van/Sheafer, Tamir/Stanyer, James: The Personalization of Mediated Political Communication: A Review of Concepts, Operationalizations and Key Findings, in: Journalism 13/2 (2012), p. 203–220, doi.org/10.1177/1464884911427802 (5.6.2024).

Agre, Philip E.: Surveillance and capture: Two Models of Privacy, in: The Information Society 10/2 (1994), p. 101–127, doi.org/10.1080/01972243.1994.9960162 (5.6.2024).

Amancio, Marina: "Put it in Your Story": Digital Storytelling in Instagram and Snapchat Stories, 2017, www.diva-portal.org/smash/record.jsf?pid=diva2%3A1111663&dswid=7890 (5.6.2024).

Arnold, Taylor/Tilton, Lauren: Distant Viewing: Analyzing Large Visual Corpora, in: Digital Scholarship in the Humanities 34/Supplement_1 (2019), p. i3–i16, doi.org/10.1093/llc/fqz013 (5.6.2024).

Arnold, Taylor/Tilton, Lauren: Distant Viewing: Computational Exploration of Digital Images, Cambridge 2023.

Asha Paul, Milan Kumar/Kavitha, Jeyaraman/Jansi Rani, Arockia P.: Key-Frame Extraction Techniques: A Review, in: Recent Patents on Computer Science 11/1 (2018), doi.org/10.2174/2213275911666180719111118 (5.6.2024).

Baden, Christian/Pipal, Christian/Schoonvelde, Martijn/der Velden, Mariken van A. C. G.: Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda, in: Communication Methods and Measures 16/1 (2022), doi.org/10.1080/19312458.2021.2015574 (5.6.2024).

Bainotti, Lucia/Caliandro, Alessandro/Gandini, Alessandro: From Archive Cultures to Ephemeral Content, and Back: Studying Instagram Stories with Digital Methods, in: New Media & Society 24/12 (2020), doi.org/10.1177/1461444820960071 (5.6.2024).

Bast, Jennifer: Politicians, Parties, and Government Representatives on Instagram: A Review of Research Approaches, Usage Patterns, and Effects, in: Review of Communication Research 9 (2021).

Bengfort, Benjamin/Bilbro, Rebecca/Ojeda,Tony: Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning, Beijing 2018.

Birkenmaier, Lukas/Lechner, Clemens M./Wagner, Claudia: The Search for Solid Ground in Text as Data: A Systematic Review of Validation Practices and Practical Recommendations for Validation, in: Communication Methods and Measures 18/3 (2024), p. 249–277, doi.org/10.1080/19312458.2023.2285765 (5.6.2024).

Bo-A, Rhee/ Federico, Pianzola/ Gang-Ta, Choi: Analyzing the museum experience through the lens of Instagram posts, Curator: A Quarterly Publication of the American Museum of Natural History 64/3 (2021), p. 529–547, doi.org/10.1111/cura.12414 (6.9.2024).

Bossetta, Michael: The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election, in: Journalism & Mass Communication Quarterly 95/2 (2018), p. 471–496, doi.org/10.1177/1077699018763307 (5.6.2024).

Brown, Tom B./Mann, Benjamin/Ryder, Nick/Subbiah, Melanie/Kaplan, Jared/Dhariwal, Prafulla/ Neelakantan, Arvind/Shyam, Pranav/Sastry, Girish/Askell, Amanda/Agarwal, Sandhini/ Herbert-Voss, Ariel/Krueger, Gretchen/Henighan, Tom/Child, Rewon/Ramesh, Aditya/Ziegler, Daniel M./Wu, Jeffrey/Winter, Clemens/Hesse, Christopher/Chen, Mark/Sigler, Eric/Litwin, Mateusz/Gray, Scott/Chess, Benjamin/Clark, Jack/Berner, Christopher/McCandlish, Sam/ Radford, Alec/Sutskever, Ilya/Amodei, Dario: Language Models are Few-Shot Learners, in: arXiv (22.7.2020), arxiv.org/abs/2005.14165 (6.9.2024).

Caliandro, Alessandro: Repurposing Digital Methods in a Post-API Research Environment: Methodo-logical and Ethical Implications, in: Italian Sociological Review 11/4S (2021), doi.org/10.13136/ISR. V11I4S.433 (5.6.2024).

Chao, Jason: Memespector-GUI: Graphical User Interface Client for Computer Vision APIs (2023), doi. org/10.5281/zenodo.7704877 (6.9.2024).

Döring, Nicola/Bortz,Jürgen: Forschungsmethoden und Evaluation in den Sozial- und Humanwissen-schaften, Berlin 2016, doi.org/10.1007/978-3-642-41089-5 (5.6.2024).

El-Keilany, Alina/Schmidt, Thomas/Wolff, Christian: Distant Viewing of the Harry Potter Movies via Computer Vision, in: Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries Conference (DHNB 2022), Uppsala Sweden, p. 33–49, ceur-ws.org/Vol-3232/paper03.pdf (5.6.2024).

franzke, aline shakti/Bechmann, Anja/Zimmer, Michael/Ess, Charles M.: Internet research: ethical guidelines 3.0: association of internet researchers (2020), aoir.org/reports/ethics3.pdf (6.9.2024).

Georgakopoulou, Alexandra: Sharing the Moment as Small Stories: The Interplay Between Practices & Affordances in the Social Media-Curation of Lives, in: Narrative Inquiry 27/2 (2017), p. 311–333, doi.org/10.1075/ni.27.2.06geo (5.6.2024).

Gerlitz, Carolin/Rieder, Bernhard: Tweets are Not Created Equal: Investigating Twitter's Client Ecosystem, in: International Journal of Communication Systems 12 (2018), p. 528–547, pure.uva. nl/ws/files/23266519/5974_30096_2_PB.pdf (5.6.2024).

Gibbs, Martin/Meese, James/Arnold, Michael/Nansen, Björn/Carter, Marcus: #Funeral and Instagram: Death, Social Media, and Platform Vernacular, in: Information, Communication and Society 18/3 (2015), p. 255–268, doi.org/10.1080/1369118X.2014.987152 (5.6.2024).

Gu, Jindong/Han, Zhen/Chen, Shuo/Beirami, Ahmad/He, Bailan/Zhang, Gengyuan/Liao, Ruotong/Qin, Yao/Tresp, Volker/Torr, Philip: A Systematic Survey of Prompt Engineering on Vision-Language Foundation Models, in: arXiv (24.7.2023), arxiv.org/abs/2307.12980 (5.6.2024).

Haim, Mario: Computational Communication Science: Eine Einführung, Wiesbaden 2023.

Haßler, Jörg/Kümpel, Anna Sophie/Keller, Jessica: Instagram and Political Campaigning in the 2017 German Federal Election. A Quantitative Content Analysis of German Top Politicians' and Parlia-mentary Parties' Posts, in: Information, Communication and Society 26/3 (2021), p. 530–550, doi. org/10.1080/1369118X.2021.1954974 (5.6.2024).

Haßler, Jörg/Wurst, Anna-Katharina/Jungblut, Marc/Schlosser, Katharina: Influence of the Pandemic Lockdown on Fridays for Future's Hashtag Activism, in: New Media & Society, 25/8 (2021), doi. org/10.1177/14614448211026575 (5.6.2024).

Haßler, Jörg/Wurst, Anna-Katharina/Pohl, Katharina: Politicians Over Issues? Visual Personalization in Three Instagram Election Campaigns, in: Information, Communication and Society 27/5 (2023), p. 815–835, doi.org/10.1080/1369118X.2023.2227684 (5.6.2024).

Joo, Jungseock/Steinert-Threlkeld, Zachary C.: Image as Data: Automated Content Analysis for Visual Presentations of Political Actors and Events, in: Computational Communication Research 4/1 (2022), doi.org/10.5117/ccr2022.1.001.joo (5.6.2024).

Leaver, Tama/Highfield, Tim/Abidin, Crystal: Instagram: Visual Social Media Cultures, Cambridge 2020.

Liebhart, Karin/Bernhardt, Petra: Political Storytelling on Instagram: Key Aspects of Alexander Van der Bellen's Successful 2016 Presidential Election Campaign, in: Media and Communication 5/4 (2017), p. 15–25, doi.org/10.17645/mac.v5i4.1062 (5.6.2024).

Liu, Pengfei/Yuan, Weizhe/Fu, Jinlan/Jiang, Zhengbao/Hayashi, Hiroaki/Neubig, Graham: Pre-Train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing, in: ACM Computing Surveys 55/9 (2023), doi.org/10.1145/3560815 (5.6.2024).

Manovich, Lev: Cultural Analytics, Cambridge 2020.

Mayring, Philipp: Qualitative content analysis: theoretical foundation, basic procedures and software solution, nbn-resolving.org/urn:nbn:de:0168-ssoar-395173 (06.09.2024).

McCrow-Young, Ally: Approaching Instagram Data: Reflections on Accessing, Archiving and Anonymising Visual Social Media, in: Communication Research and Practice 7/1 (2021), p. 21–34, doi.org/10.1080/22041451.2020.1847820 (5.6.2024).

Milanesi, Matilde/Guercini, Simone: Image-Based Social Media and Visual Content Analysis: Insights from a Literature Review, in: Micro & Macro Marketing, 3 (2020), p. 537–558, ideas.repec.org/a/mul/jyf1hn/doi10.1431-97640y2020i3p537-558.html (5.6.2024).

Marres, Noortje/Weltevrede, Esther: Scraping the social? Issues in real-time social research, in: Journal of Cultural Economy 6 (2013), doi.org/10.1080/17530350.2013.772070 (14.10.2024).

Omena, Janna Joceli/Teixeira Rabello, Elaine/Goes Mintz, André: Digital Methods for Hashtag Engagement Research, in: Social Media + Society 6/3 (2020), doi.org/10.1177/2056305120940697 (5.6.2024).

Paun, Silviu/Artstein, Ron/Poesio, Massimo: Using Agreement Measures for CL Annotation Tasks, in: Paun, SilviuArtstein, Ron/Poesio, Massiomo (ed.), Statistical Methods for Annotation Analysis, Cham 2022, p. 47–78, doi.org/10.1007/978-3-031-03763-4_3 (5.6.2024).

Peeters, Stijn: Zeeschuimer (v1.8.0), doi.org/10.5281/zenodo.8399900 (5.6.2024).

Peng, Yilang/Lock, Irina/Salah, Albert Ali: Automated Visual Analysis for the Study of Social Media Effects: Opportunities, Approaches, and Challenges, in: Communication Methods and Measures 18/2 (2023), p.163–185, doi.org/10.1080/19312458.2023.2277956 (6.5.2024).

Reiter, Niels: How to Develop Annotation Guidelines, in: nilsreiter.de (blog), 1.10.2017, nilsreiter.de/blog/2017/howto-annotation (14.10.2024).

Rejeb, Abderahman/Rejeb, Karim/Abdollahi, Alireza/Treiblmaier, Horst: The Big Picture on Instagram Research: Insights from a Bibliometric Analysis, in: Telematics and Informatics 73 (2022), doi.org/10.1016/j.tele.2022.101876 (5.6.2024).

Rogers, Richard: Digital Methods for Web Research, in: Emerging Trends in the Social and Behavioral Sciences, 15.5.2015, p. 1–22, doi.org/10.1002/9781118900772.etrds0076 (5.6.2024).

Rogers, Richard: Digital Methods, Cambridge 2013.

Rose, Gillian: Visual Methodologies: An Introduction to Researching with Visual Materials, 4th Ed., Los Angeles 2016.

Serengil, Sefikm Ilkin/Ozpinar, Alper: LightFace: A Hybrid Deep Face Recognition Framework, in: Innovations in Intelligent Systems and Applications Conference (ASYU), Istanbul 2020, p. 1–5, doi.org/10.1109/ASYU50717.2020.9259802 (5.6.2024).

Song, Hyunjin/Tolochko, Petro/Eberl, Jakob-Moritz/Eisele, Olga/Greussing, Esther/Heidenreich, Tobias/Lind, Fabienne/Galyga, Sebastian/Boomgaarden, Hajo G.: In Validations We Trust? The Impact of Imperfect Human Annotations as a Gold Standard on the Quality of Validation of

Automated Content Analysis, in: Political Communication 37/4 (2020), p. 550–572, doi.org/10.108 0/10584609.2020.1723752 (14.10.2024).

Towner, Terri L./Lego Muñoz, Caroline: A Long Story Short: An Analysis of Instagram Stories during the 2020 Campaigns, in: Journal of Political Marketing 21/3–4 (2022), doi.org/10.1080/15377857.2022 .2099579 (5.6.2024).

Törnberg, Petter: Best Practices for Text Annotation with Large Language Models, in: arXiv (5.2.2024), arxiv.org/abs/2402.05129 (5.6.2024).

Trezza, Domenico: To Scrape or Not to Scrape, This Is Dilemma. The Post-API Scenario and Implications on Digital Research, in: Frontiers in Sociology 8 (2023), doi.org/10.3389/ fsoc.2023.1145038 (5.6.2024).

Wickham, Hadley: Tidy Data, in: Journal of Statistical Software 59/10 (2014), doi.org/10.18637/jss.v059. i10 (5.6.2024).

Zhao, Zihao/Wallace, Eric/Feng, Shi/Klein, Dan/Singh, Sameer: Calibrate Before Use: Improving Few-Shot Performance of Language Models, in: Meila, Marina/Zhang, Tong (ed), Proceedings of the 38th International Conference on Machine Learning, vol. cxxxix, Proceedings of Machine Learning Research, 2021, p. 12697–12706, proceedings.mlr.press/v139/zhao21c.html (5.6.2024).