# Visual Methods for Digital Research

*Digital Methods Summer School 2024*

**Tutorials**

**1 – 12 July 2024**

https://bit.ly/dmi24-ss-tutorials

***#dmi24***

**Digital Methods Initiative**
**University of Amsterdam**
Turfdraagsterpad 9, 1012 XT Amsterdam
https://wiki.digitalmethods.net/Dmi/SummerSchool2023

# Contents

# Template

*!! Instructions: Please copy-paste the template below for all tutorials.*

| | |
|---|---|
| **1. Tutorial title:** | *Please provide a title that succinctly captures the tutorial.* |
| **2. Facilitator(s):** | *Please name the facilitators.* |
| **3. Short description:** | *Summarise the contents of the tutorial, and the training aim or takeaway.* |
| **4. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# Multi- and cross-platform data analysis with 4CAT

| | |
|---|---|
| **1. Tutorial title:** | Multi- and cross-platform data analysis with 4CAT |
| **2. Facilitator(s):** | Sal Hagen<br>Stijn Peeters<br><br>**Prerecorded tutorials:**<br>● Creating a dataset with 4CAT<br>● Analysing a dataset using processors<br>**Exercises:**<br>● 4CAT exercises |
| **3. Short description:** | In this tutorial we discuss and demonstrate how to create and analyse datasets with 4CAT, the 4CAT Capture & Analysis Toolkit. We use the 4CAT tool to build datasets according to a specific query and run subsequent textual and statistical analysis on them. Data from the following platforms is available through 4CAT (some more complete than others):<br><br>● 4chan<br>● 8kun<br>● BitChute |

| | |
|---|---|
| | ● Parler<br>● Telegram<br>● Douban<br>● Tumblr<br>● Instagram, via CrowdTangle (no capture, analysis only)<br>● TikTok and Instagram (see the 'Capturing TikTok with Zeeschuimer' tutorial)<br><br>Examples of analyses that we discuss are the generation of image walls; frequency diagrams of post behaviour or characteristics for specific publics; identification of neologisms and domain-specific vocabulary; extracting YouTube metadata and thumbnails from videos linked in the dataset; scanning the dataset for hateful language; and more generally the generation of issue-specific data sets and visualisations sourced from these platforms.<br><br>No specific preparation is required, as 4CAT is a web-based tool, which we will provide access to. It helps to be somewhat familiar with the platforms at hand (particularly 4chan and ~~Reddit~~); if you would like to follow this tutorial, we recommend that you browse the platforms you would like to query beforehand. Note that especially 4chan and 8chan may contain shocking content. |
| **4. Related project(s):** | Useful for probably most Summer School projects |

# # Capturing TikTok (and Instagram) data with Zeeschuimer and 4CAT

| | |
|---|---|
| **1. Tutorial title:** | Capturing TikTok (and Instagram) data with Zeeschuimer and 4CAT (Presentation slides here). |
| **2. Facilitator(s):** | Stijn Peeters<br>Lucia Bainotti |
| **3. Short description:** | In this tutorial we discuss and demonstrate how to capture data from TikTok using the browser extension Zeeschuimer, and then analyse it with the 4CAT Capture and Analysis Toolkit.<br><br>Capturing TikTok data with Zeeschuimer allows one to capture the metadata and video/thumbnail URLs of TikTok posts seen while browsing the TikTok website. It is a way to study TikTok data as it is seen by a user, enabling analysis of e.g. its recommendation algorithm or affording a TikTok-based walkthrough analysis. This can be contrasted with specialised scraping tools, which capture TikTok data outside its own interface and are arguably less suitable for exploratory, critical analyses.<br><br>Zeeschuimer can also be used to capture data from Instagram. While this is not the primary focus of the tutorial, instructions should allow participants to additionally study that platform too.<br><br>Participants are recommended (but not required) to either have familiarity with 4CAT or to additionally follow the tutorial about that tool (see above). There is also another tutorial on further analysis of TikTok data - after it has been captured - with ImageJ and Memespector (see further down). As Zeeschuimer is a Firefox extension, participants need to install that browser to use it.<br><br>A worksheet outlining the protocol discussed in this tutorial is available here: https://tinyurl.com/nmrw-zeeschuimer-tiktok. |
| **4. Related project(s):** | Particularly useful for TikTok and Instagram analyses |

# # The Social Moving Image: Navigating TikTok

| | |
|---|---|
| **1. Tutorial title:** | *The Social Moving Image: Navigating TikTok* |
| **2. Facilitator(s):** | *Lucia Bainotti, Elena Pilipets, Marloes Geboers, (PRESENTATION) (WORKSHOP FOLDER)* |
| **3. Short description:** | ⚠️ **Attention** ⚠️ This tutorial is a natural follow up to the *Capturing TikTok (and Instagram) data with Zeeschuimer and 4CAT* tutorial provided by Stijn Peeters and Sal Hagen. But everyone is WELCOME.<br><br>It presents targeted case studies that showcase how TikTok data, captured with Zeeschuimer, can be analysed with sensitivity for the medium-specificities presented by moving images (video).<br><br>We ask: How to study the social moving image on TikTok? To which extent can we repurpose TikTok metadata for the analysis of networked video cultures? Which new forms of memeification and templatability can we identify?<br><br>We invite participants to discuss a series of techniques for Tiktok navigation by “metadating” collections of TikTok videos and “metapicturing” their different networked characteristics. If metadata, according to Lev Manovich, is what allows computers to “connect data with other data” (2002), “metapicturing” (Mitchell 1995; Rogers 2021) can be seen as an analytical technique for inventive, ethical, and contextual data remix.<br><br>Programme 65 min:<br><br>Short **introduction** (5 min)<br><br>**Part I with Lucia Bainotti**. This part focuses on the **listed sounds** searchable in the TikTok sound library and presents a technique for arranging audio-visual content in |

| | |
|---|---|
| | a specific type of metapicture—**video stripes**—with the 4CAT Capture and Analysis Toolkit (Peeters & Ha.gen 2022). Participants will learn how to "follow the sound" and create scene-by- scene video timelines, resulting in a horizontal collage of sequential frames that can be used to easily identify styles, patterns and recurring gestures in the dataset.<br><br>**Part II with Elena Pilipets** will focus on the detection of speech templates in the videos networked through polyvocal **original sounds** and show how to create **video stacks** based on a selection of frames from the videos that share the same aural pattern. Participants will learn how to work with tools such as Video Frame and Audio Extractor, Speech-To-Text Converter (both Chao 2022), and ImageJ (Rasband et al. 2018).<br><br>**Part III with Marloes Geboers** will discuss a series of methodological designs for contextualising video content through the analysis of linkages between visual and textual elements **(hashtags, stickers, emojis)** and specific practices **(such as effects or duetting)**. Different ways of data curation in Google spreadsheets exemplify the mindset of 'mix and match metadata' that she aims to install in participants. Another aspect is detecting patterns in eg. comments using text processing tools such as Polsys Text Analysis tool (Rieder 2018).<br><br>**Questions** (10-15 min)<br><br>🆓**Bonus** 🆓 Participants are welcome to use our 'walkthrough' documents with the step-by-step instructions for scraping, structuring, analysing, and visualising TikTok video data with different techniques. Every step of the tutorial is documented in these walkthroughs (including some considerations on the ethics of visualization). Please make yourself familiar with some of the techniques introduced below:<br><br>1. Learn how to scrape and navigate TikTok metadata with **Capturing TikTok Data with Zeeschuimer and 4CAT** |

| | |
|---|---|
| | 2. Download TikTok videos and/or thumbnails to a local folder using TikTok URLS/thumbnails with **TikTok Scraping and Video Download**<br><br>3. Learn different techniques of video frame extraction and convert TikTok's 'original sounds' to text with **Video Frame Extractor and Speech-To-Text Converter**<br><br>4. Learn how to contextualize, visualize, and ethically interpret collections of TikTok videos with **Video Analysis with ImageJ (Plot, Montage, Stack)** |
| **4. Related project(s):** | *Any project interested in studying TikTok* |

# Content moderation research: theory and tools for the study of moderation practices

| | |
|---|---|
| **1. Tutorial title:** | Content moderation research: theory and tools for the study of moderation practices |
| **2. Facilitator(s):** | Emillie de Keulenaar |
| **3. Short description:** | This tutorial recaps current methods for the capture and analysis of content moderation practices, and then offers a few hands-on tools for participants to try those methods themselves.<br><br>First, we will go through a bit of the methodological literature behind content moderation studies, discussed in the literature below. Then, we will set up the right environments to begin MCAT, which is a collection of Python scripts that analyse content moderation policies from a variety of platforms, and get statuses from Facebook, Instagram, YouTube and Twitter posts.<br><br>With those scripts, we will learn to:<br>1. Compare differences between platform policies, as well as upload them to 4CAT for NLP processing, or Prompt Compass for systematic analysis for a variety of LLMs;<br>2. Upload a CSV file from the YouTube Data Tools, Crowdtangle or Zeeschuimer, and then check the status of each post;<br>3. Trace YouTube video (and possibly tweet) demotions or "shadowbanning".<br><br>This tutorial is particularly **useful for those using archives** (i.e., older datasets) containing social media posts that may have been moderated at some point or another (for example: COVID, electoral and other politically sensitive material such as the YouTube radicalization dataset (2006-2018)). |

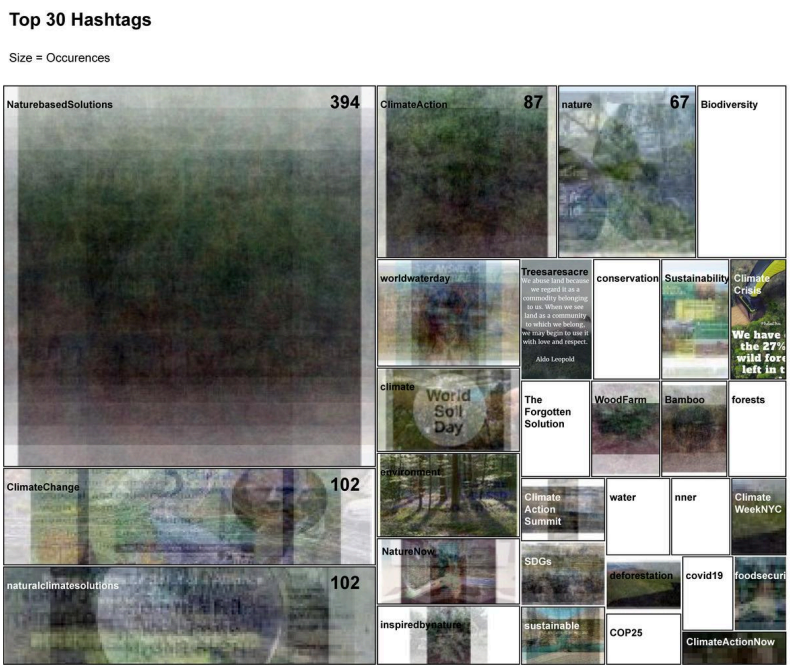
| | |
|---|---|
| | To prepare for this tutorial, we recommend going through this chapter first:<br><br>de Keulenaar E and Rogers R (2024) After deplatforming: the return of trace research for the study of content moderation. In: Venturini T, Acker A, Plantin J-C, et al. (eds) *The SAGE Handbook of Data and Society: An Interdisciplinary Reader in Critical Data Studies*. London: SAGE. Available at: https://drive.google.com/file/d/1P2V_l6cz4Q52V850Agb9FYPF_zziZrWc/view?usp=share_link. |
| **4. Key links** | 1. MCAT notebook: https://colab.research.google.com/drive/1Fut81WJYpsUOtau_9o6jZ8vFzoALpy1z?usp=sharing<br>2. Moderation capture and analysis scripts: https://github.com/edekeulenaar/MCAT-moderation-analysis-toolkit<br>3. Method: https://drive.google.com/file/d/1P2V_l6cz4Q52V850Agb9FYPF_zziZrWc/view?usp=share_link<br>4. Method presentation: https://drive.google.com/file/d/1mmDQoL73au7NE8eiaQficrkZnDOHMLCp/view?usp=sharing<br>5. Folder with datasets and other examples: https://drive.google.com/drive/folders/1earySxMFy2DGO1rZjuuNX-9HX9mjrLEI |
| **6. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# # Query Design for Social Media Research

| | |
|---|---|
| **Tutorial title:** | Query Design for Social Media Research |

| | |
|---|---|
| **Facilitators:** | *Richard Rogers*<br>Pre-recorded video |
| **Short description:** | The question of what constitutes a keyword is the starting point for query design, for that is what makes querying (and search as research) practically part of a research strategy. When formulating a query, one often begins with keywords so as to ascertain who is using them, in which contexts and with which spread or distribution over time. In the tutorial a particular keyword query strategy or design is put forward, whereby one queries competing keywords, asking whether a particular term is winning favour and amongst whom. But where to make the queries? The tutorial also provides heuristics for curating as well as deriving source sets. Such source sets may be social media collections from Twitter, Facebook, Instagram or elsewhere. They may be from the web. Ultimately, queries are designed so as to answer a research question, which is where the tutorial both begins and concludes.<br><br>See R. Rogers (2017) Foundations of Digital Methods: Query Design. In: Mirko Tobias Schäfer, Karin van Es (Eds.): The Datafied Society: Studying Culture through Data. Amsterdam: Amsterdam University Press, pp. 75–94. DOI: https://doi.org/10.25969/mediarep/12536. |

# # RAWGraphs 2.0: from spreadsheet to visualization

| | |
|---|---|
| **1. Tutorial title:** | RAWGraphs 2.0: from spreadsheet to visualization<br>https://bitly.cx/GrnoW |
| **2. Facilitator(s):** | *Anna Cattaneo, Noemi Capparelli (Week 1)*<br>*Ángeles Briones (Week 2)* |
| **3. Short description:** | *This tutorial will introduce RAWGraphs 2.0 as a free, open-source, and easy-to-use tool to produce data visualizations from various data formats. This tutorial will teach you how to use RAW2.0's user interface and which visual models suit which data structures. Additionally, we will introduce a prototype to produce a treemap with images like this one:*<br>  <br><br>*Recording of the tutorial:*<br>*https://youtu.be/suQY9jk7QFI* |

| | |
|---|---|
| **4. Related project(s):** | *Previous DMI School projects are:*<br>- *Understanding Normiefication: A Cross-Platform Analysis of the QAnon Conspiracy (Winter School 2019)*<br>*Image Circulation of Russian State-Controlled and Independent Media on Google.com and Yandex.ru (Summer School 2022)*<br>- *Slicing Bread-tube: Leftist Discourse Networks (Summer School 2019)* |

# # YouTube Data Tools

| | |
|---|---|
| **Tutorial title:** | YouTube Data Tools |
| **Facilitators:** | *Daniel Jurg*<br>Series of pre-recorded videos from creator Bernhard Rieder (2023) |
| **Short description:** | This tutorial will show how to use the YouTube Data Tools (and hint at other useful tools and browser extension) to extract and analyze various datasets from YouTube channels, videos, comments, and recommendations.<br><br>There is an extensive 'YouTube cookbook' with detailed descriptions here (bit.ly/3WSAEWf). |
| **Related Projects:** | |

# # Collection Space Navigator: An Interactive Visualization Interface for Multidimensional Datasets

| | |
|---|---|
| **1. Tutorial title:** | Collection Space Navigator: An Interactive Visualization Interface for Multidimensional Datasets |
| **2. Facilitator(s):** | *Mar Canet Sola* |
| **3. Short description:** | This tutorial will show how to use *The Collection Space Navigator (CSN) which is an explorative visualization tool for researching image collections and their multidimensional representations. We designed this tool to better understand multidimensional data, its methods, and semantic qualities through spatial navigation and filtering. CSN can be used with any image collection and can be customized for specific research needs (see Jupyter Notebook or Google Colab).*   <br><br>*https://github.com/Collection-Space-Navigator/CSN (MIT licence)*<br><br>*Slides:*<br>*https://docs.google.com/presentation/d/1rBVTsCLhPJwnJuC67TRCSYZrHCIu9qphHhE8cj9vZZE/edit?usp=sharing*<br><br>*Article:*<br>*Ohm, T., Canet Sola, M., Karjus, A., & Schich, M. (2023, September). Collection space navigator: an interactive visualization interface for multidimensional datasets. In Proceedings of the 16th International Symposium on Visual Information Communication and Interaction (pp. 1-5).*<br><br>*List of projects using this tool:* |

| | |
|---|---|
| | • *Kinokroonika project CSN - Visualizing Estonian newsreels in the 20th century: https://csn.kinokroonika.ee/*<br>• *Datalab CSN, an interactive visualization interface that allows visitors to explore the collections of Badisches Landesmuseum Karlsruhe and Allard Pierson Amsterdam: https://datalab.landesmuseum.de/CSN/* |
| **4. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# # TeleCatch: filtering, visualizing and extracting Telegram messages

| | |
|---|---|
| **1. Tutorial title:** | TeleCatch: filtering, visualizing and extracting Telegram messages |
| **2. Facilitator(s):** | *Giosuè Ruscica*<br>*Giulia Tucci*<br>*Bia Carneiro* |
| **3. Short description:** | The tutorial will teach you how to use the TeleCatch tool to create Telegram datasets (with content and metadata of messages and image files). By feeding a list of Telegram public groups and/or channels into the tool, it is possible to view the most recent messages, as well as apply queries and extract only messages related to a specific topic. |
| **4. Related project(s):** | Uncovering visual narratives about human mobility on Telegram |

# # Exploring Blockchain Fundamentals: Web3 Wallets and NFT Image Collections

| | |
|---|---|
| **1. Tutorial title:** | *Exploring Blockchain Fundamentals:*<br>*Web3 Wallets and NFT Image Collections* |
| **2. Facilitator(s):** | *Janna Joceli Omena, Alberto Cossu, Augusto Falcão*<br>*Pre-recorded tutorial* I *Slides* |
| **3. Short description:** | This tutorial will explore blockchain fundamentals from conceptual, practical, and technical perspectives. Participants will engage in simple hands-on activities, including installing Web3 browsers and browser extensions, creating software crypto wallets, and visualizing NFT image collections. The aim is to lay the groundwork for a fundamental understanding of blockchain while discussing potential methods to repurpose Web3 software wallets for research purposes. The tutorial is |

| | |
|---|---|
| | divided into three parts, including a collective discussion, as follows:<br><br>**Part 1:** An introduction to key concepts, history and terminology in blockchain technology and its role as a research area (Alberto Cossu, Digital Sociologist of Crypto-Finance)<br><br>**Part 2:** Understanding Web3 wallets and different types of actions to receive an airdrop (Augusto Falcão, Blockchain Practitioner and Web3 Expert; Janna Joceli Omena, Digital Methodologist)<br><br>**Part 3:** Creating Web3 software wallets and visualising NFT image collections (Janna Joceli Omena, Digital Methodologist)<br><br>**Collective Discussion:** How can Web3 wallets and NFT image collections be repurposed? What are the potential applications?<br><br>By the end of the session, attendees will have practical experience with Web3 software wallets and a grasp of blockchain fundamentals. This tutorial is ideal for anyone seeking practical skills and insights into Web3 wallets and advancing critical digital methods with them.<br><br>**Preparation:**<br><br>  Install Brave to explore web3 wallets<br>  Install Instant Data Scraper or  DownThemAll to download NFT images<br><br>Download and install (at least one) software to visualise image collections by Visual Computing Group:<br><br>ImageSorter for Windows XP / Vista / 7 / 8 / 10,<br>Mac OS X 10.6 Snow Leopard, Windows 10 x64<br><br>PicArrange for macOS 10.15 Catalina and above<br><br>ImageX for Windows 10 x64 (for old hardware), Mac OS X 10.13 High Sierra |

| | |
|---|---|
| **4. Related project(s):** | *Digital methods for Blockchain Research: Exploring Web3 software wallets and mapping Airdrop Farming Vernaculars on Social Media* |

# # Enriching image data with AI

| | |
|---|---|
| **1. Tutorial title:** | Enriching image data with AI (Cross-vision API studies) |
| **2. Facilitator(s):** | Jason Chao & Janna Joceli Omena<br>Pre-recorded video |
| **3. Short description:** | Participants will learn how to exploit AI technologies to enrich image datasets.<br><br>Participants will be introduced to the affordances of computer vision APIs supported by Memespector-GUI:<br>● Google Vision (proprietary)<br>● Microsoft Azure Cognitive Services (proprietary)<br>● Clarifai (proprietary)<br>● Image classifier based on Keras (open source)<br><br>Memespector-GUI is a tool with graphical user interface which helps researchers invoke proprietary and open source computer vision APIs to analyse images with ease.<br><br>Preparation: In this tutorial, we recommend that the participants try invoking at least one proprietary API to process an image dataset. Participants are advised to register with one of the APIs beforehand by following the instructions for Google Cloud, Microsoft Azure or Clarifai. (Google Cloud and Microsoft Azure may ask for bank card details. They will check whether the card is active but will not charge on the card in the process of registration.)<br><br>Note on proprietary APIs: Proprietary APIs are commercial services and using them is not necessarily free of charge. Google Cloud, Microsoft Azure and Clarifai provide each account with monthly quotas of free API requests. Moreover, new accounts usually receive free credits from Google Vision and Microsoft Azure which are adequate for processing tens of thousands of images.<br><br>Slides https://bit.ly/DMI-tutorial_MemespectorGUI |

| | |
|---|---|
| | Preparation:<br><br>Download Memespector-GUI at<br>https://github.com/jason-chao/memespector-gui/releases<br><br>Download a test image dataset at<br>https://drive.google.com/drive/folders/1fve2kS_kG3lXebr2B8qu-AnavXlxGLPl?usp=sharing<br><br>Download temporarily access keys for Google Vision API, Microsoft Azure Cognitive Services and Clarifai at (...) |
| **4. Related project(s):** | 'What should climate change look like? Mapping climate imaginaries across platforms and eco-fiction genres'<br>—> 2022 What is a meme, technically speaking?<br>—> 2021 Mapping deepfakes with digital methods and visual analytics<br><br>● This tutorial can assist projects aiming at:<br>○ Making sense of large collections of images with literal and contextual descriptions<br>○ Tracing the circulation of visual content across the web<br>○ Content moderation and NSFW analysis |
| **5. References / use cases:** | ● Computer Vision Networks. Developing digital visual methods for social and media research.<br>https://www.cais.nrw/en/fellow_omena_2021_en/<br>https://www.researchgate.net/publication/350153591_Computer_Vision_Networks_a_research_proposal<br><br>● Publications using computer vision for research purposes:<br><br>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.<br><br>Geboers, M. A., & Van De Wiele, C. T. (2020). Machine Vision and Social Media Images: Why Hashtags Matter. Social Media and Society, 6(2). https://doi.org/10.1177/2056305120928485<br><br>Mintz, A., Silva, T., Gobbo, B., Pilipets, E., Azhar, H., Takamitsu, H., … Oliveira, T. (2019). Interrogating Vision APIs. Lisbon. Retrieved from |

https://smart.inovamedialab.org/smart-2019/project-reports/interrogating-vision-apis/

Omena, J. J., Rabello, E. T., & Mintz, A. G. (2020). Digital Methods for Hashtag Engagement Research. Social Media + Society, (July-September), 1–18. https://doi.org/10.1177/2056305120940697

Omena, J. J., Elena, P., Gobbo, B., & Jason , C. (2021). The Potentials of Google Vision API-based Networks to Study Natively Digital Images. *Diseña*, (19), Article.1. https://doi.org/10.7764/disena.19.Article.1

# # Network analysis with Gephi

| | |
|---|---|
| **Tutorial title:** | Network analysis with Gephi |
| **Facilitators:** | *Carlo De Gaetano*<br>Pre-recorded tutorial |
| **Short description:** | Gephi is an open-source software for visualizing and analysing large networks graphs. This tutorial covers how to generate and analyze network visualisations from generic datasets.<br><br>We will discuss when and why it is interesting to visualise networks, and we will learn how to do so by exploring the basics of Gephi, an open-source network analysis and visualization software. We will learn how to format a dataset for Gephi, and how to use different visual features and layout algorithms to visualise the dataset as a network. We will discuss principles on how to read a network visualisation and make sense out of it.<br><br>This is a link to a Drive folder with the presentation, a sample dataset and some readings. |
| **Related projects** | For Winter School projects undertaking network analysis and visualisation |

# # Studying ‘Fake News’

| | |
|---|---|
| **Tutorial title:** | Studying ‘Fake News’ |
| **Facilitators:** | *Richard Rogers*<br>Pre-recorded video |
| **Short description:** | Studying ‘fake news’ walks the researcher through a three-step process whereby one first demarcates political spaces on a variety of platforms through query design, receives lists of sources ranked by interactions (or engagement) and finally uses media bias classification schemes to label the sources. Ultimately one is able to measure the extent to which problematic sources are presented in political spaces on social media platforms. In the video, I also discuss changing the classifications, showing for example that if one uses a narrow definition of ‘fake news’ the problem shrinks.<br><br>See R. Rogers (2020). The scale of Facebook’s problem depends on how ‘fake news’ is classified, *Harvard Misinformation Review*. |

# # Praat: doing phonetics by computer

*!! Instructions: Please copy-paste the template below for all tutorials.*

| | |
|---|---|
| **1. Tutorial title:** | *Praat: doing phonetics by computer* |
| **2. Facilitator(s):** | *Dirk Vet* |
| **3. Short description:** | *On Day 1, participants will be instructed how to run a pitch analysis of the live commentary in video highlights of football* |

| | |
|---|---|
| | *games via a live tutorial of 1 hour facilitated by Dirk Vet, phonetic science lab manager at the University of Amsterdam. Download the Praat app at https://www.fon.hum.uva.nl/praat/ and follow the introductory tutorial found under the Help menu.* |
| **4. Related project(s):** | *N/A* |

# # Being Critical With AI Investigations

*!! Instructions: Please copy-paste the template below for all tutorials.*

| | |
|---|---|
| **1. Tutorial title:** | **Being critical with AI investigations** |
| **2. Facilitator(s):** | Matti Nelimarkka |
| **3. Short description:** | *This tutorial examines how using AI based models for social investigations brings in the unknown values and world views. For example, large language models are known to have more traditional and even conservative gender representations. If we use these tools to "measure the society" we automatically incorporate these worldviews into our analysis.*<br><br>*In the first part of the tutorial, we focus on image recognition tools and examine differences they have in labelling following methods proposed by Berg & Nelimarkka (2023). We discuss differences and implications these have and learn how to use the COSLAB tool presented.*<br><br>*In the second part of the tutorial, we embed ourselves into values in large language models. We walk through the code used for fine-tuning an LLM with Marx's thoughts and devise our own LLM to focus on a theoretical stance we have.* |

| | |
|---|---|
| | *If there is time after all of this, we look through how to use image fingerprinting to find image re-use.* |
| **4. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# # Dealing with image collections: a toolbox

| | |
|---|---|
| **1. Tutorial title:** | Dealing with image collections: a toolbox |
| **2. Facilitator(s):** | *Alessandra Facchin (Week 1)*<br>*Tbd (Week 2)* |
| **3. Short description:** | This tutorial will introduce tools for doing research with digital images. We will provide a quick overview of workflows to visualise a corpus of images with different techniques and layouts, i.e. ImageSorter, PicArrange, ImageJ, Pixplot, and Gephi.<br>During the tutorial, a hands-on activity will follow a workflow for segmenting a corpus of images using two libraries GroundingDino and Segment Anything.<br><br>Requirements:<br>- Bring your own laptop<br><br>Slides:<br>[DMIsummer24] Dealing with image collections: a toolbox |
| **4. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# # Prompting for inquiring: text-to-image generative models

| | |
|---|---|
| **1. Tutorial title:** | Prompting for inquiring: LLMs and Text-to-image generative models |
| **2. Facilitator(s):** | *Andrea Benedetti (Week 1)* |

| | |
|---|---|
| | *Tbd (Week 2)* |
| **3. Short description:** | In this tutorial you will be provided with an overview of approaches on how to prompt to inquire the inner workings of Text-to-Image generative models, such as StableDiffusion or Midjourney.<br>The tutorial presents *ambiguity*, *recursiveness*, and *isolation* of the prompt as the primary methods of inquiring the inner workings ofsuch models.<br><br>We highly recommend to follow also **"# Dealing with image collections: a toolbox"** to learn how to treat image collections generated with the methods that we will introduce.<br><br>Requirements:<br>- Bring your own laptop<br>- An active Midjourney + Discord account<br>- An active account on 4CAT to access StableDiffusion<br>- An active OpenAI account to access GPT-4<br><br>Slides: Prompting for inquiring |
| **4. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# # TIB AV-Analytics: A Computational Tool for Film and Video Analysis

| | |
|---|---|
| **1. Tutorial title:** | TIB AV-Analytics: A Computational Tool for Film and Video Analysis |
| **2. Facilitator(s):** | *Eric Müller-Budack (TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany)*<br>*Ralph Ewerth (TIB and Leibniz University Hannover)* |
| **3. Short description:** | We are excited to explore the numerous possibilities of film and video analysis with you in the upcoming tutorial on our TIB-AV-A video analysis platform available at: https://service.tib.eu/tibava<br><br>The goal of the tutorial is to introduce TIB-AV-A to participants from the digital humanities and other interested individuals. TIB-AV-A is a third-party funded project (Duration: January 2021 – June 2024) by the German Research Foundation (DFG; Deutsche Forschungsgemeinschaft). The platform is developed by the TIB – |

| | |
|---|---|
| | Leibniz Information Centre for Science and Technology in cooperation with film scholars from Johannes Gutenberg-Universität Mainz (JGU).<br><br>Unlike other video analysis tools, TIB-AV-A uses modern web technologies to provide users with access to state-of-the-art deep learning approaches for systematic film and video analysis, as well as manual annotations in an interactive user interface.<br><br>To ensure we can focus on in-depth content examination and analysis during the tutorial, we kindly ask you to prepare the following by Friday, July 5:<br><br>• Select up to three videos in an mp4 format (e.g., Video - H.265 + MP3 (MP4)) with a maximum file size of 500 MB. Alternatively, you can download sample videos (recommendation is “Silent_Love.mp4”) using the following link: https://tib.eu/cloud/s/FpRbgQRwASwGzkr<br>• Create an account and upload your video(s) to TIB-AV-A via the following link: https://service.tib.eu/tibava<br>• Optional: Watch the tutorial video available at https://service.tib.eu/tibava that is visible if you are not logged in. Alternatively, you can download it from here: https://tib.eu/cloud/s/sMmqWqWYict3Zpb/download/TIB-AV-A_Einfuehrung_2.mp4<br><br>This preparation will allow you to gain an initial insight into the platform in advance. During the tutorial, you can try out TIB-AV-A yourself. Please note that TIB-AV-A is still a prototype and only supports 5-10 parallel users. Thus, we would appreciate it if you can work in small groups during the tutorial.<br><br>Our team looks forward to a lively exchange and a productive tutorial day with you.<br><br>If you have any questions or encounter issues, please contact Eric Müller-Budack (eric.mueller@tib.eu). |
| **4. Related project(s):** | *List the School projects related to the tutorial (where applicable).* |

# # Designing With Interactive Framework

| | |
|---|---|
| **1. Tutorial title:** | **Designing With Interactive Framework** |

| | |
|---|---|
| **2. Facilitator(s):** | *Antonella Autuori - Design Researcher SUPSI*<br>*Matteo Subet - Design Researcher SUPSI*<br>Prerecorded tutorial I Slides |
| **3. Short description:** | The tutorial will introduce participants to 'Designing With Interactive Framework', one of the outcomes developed by the research project "*Designing With. A New Educational Module to Integrate Artificial Intelligence, Machine Learning and Data Visualization in Design Curricula*", founded by Movetia in 2022 and in collaboration between SUPSI, EPFL and NOVA University.<br><br>The Designing With Interactive Framework (https://designingwithai.ch/interactive-framework) provides a comprehensive and organized overview of artificial intelligence system capabilities, highlighting a data-driven approach and incorporating AI tools at specific design process stages. This interactive i-cycle diagram, featuring an intuitive interface accessible on both desktop and mobile devices, is designed to assist design teachers and students in exploring and utilizing AI tools consciously.<br><br>The diagram includes five key components that simplify the process for designers to research and select AI tools suited to their practice needs: (a) design stages; (b) AI capabilities; (c) data input; (d) data output; (e) AI tools. |
| **4. Related project(s):** | 1. *Designing with AI, ML and DV Workshop*<br>2. *Co-Design With AI Workshop* |
| **5. References** | Botta, M., Autuori, A., Subet, M., Terenghi, G., Omena, J.J., Leite, E., Kim, F.C., (2024). Designing With: A New Educational Module to Integrate Artificial Intelligence, Machine Learning and Data Visualization in Design Curricula. designingwithai.ch<br><br>Botta, M., Autuori, A., Terenghi, G., and Subet, M. (2024) A Design-Stage-Oriented Framework to Introduce Artificial Intelligence and Machine Learning in Design Education, in Gray, C., Ciliotta Chehade, E., Hekkert, P., Forlano, L., Ciuccarelli, P., Lloyd, P. (eds.), DRS2024: Boston, 23–28 June, Boston, USA. https://doi.org/10.21606/drs.2024.535<br><br>Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17), Association for Computing Machinery, USA, 278–288. https://doi.org/10.1145/3025453.3025739 |

# # How to use Looker for “quick ‘n’ dirty” data explorations

| | |
|---|---|
| **1. Tutorial title:** | How to use Looker (formerly *Google Data Studio*) for “quick ‘n’ dirty” data explorations |
| **2. Facilitator(s):** | Fabio Daniele |
| **3. Short description:** | Traditionally meant for data analysis in enterprise contexts, Business Intelligence tools such as *Power BI* or *Tableau* can be of great help when we are in need of exploring huge datasets. Unfortunately, they can also be quite cumbersome to set up and the learning curve is somewhat steep for new adopters.<br>In this tutorial, we will focus on Google’s BI solution named Looker. Compared to other platforms, Looker strikes a good balance between ease of use and features, enabling researchers to quickly create interactive dashboards, produce both live and offline reports, or gather a quick overview over their latest acquired dataset.<br><br>The tutorial will touch upon:<br>1. How to upload a dataset in Looker<br>2. How to create a report<br>3. Best practices for data exploration<br>4. Advanced tips and tricks for complex datasets<br><br>The tutorial will be interactive, participants are welcome to join the exercise if they wish to do so.<br><br>Requirements:<br>- A Google account is required to access Looker |
| **4. Related project(s):** | Can be applied to any project where data exploration and/or data reporting is required, in absence of coding experience.<br><br>Recommended for project “*Exploring political narratives around the 2024 European Elections*”. |