

2. Exploring TV Seriality and Television Studies through Data-Driven Approaches

Mirko Degli Esposti and Guglielmo Pescatore

◀ ABSTRACT

The article discusses the use of data-driven approaches in television studies, which has become possible due to the increasing availability of digital data. Computational techniques can be used to analyze cultural artifacts, gain insights into audience reactions to specific shows or episodes, and investigate patterns and trends in television programming over time. The article also highlights the challenges of analyzing television series, which are complex open systems that interact with external factors such as the production process, audience feedback, and cultural and social context. Content analysis, which involves qualitative and quantitative methods based on the coding process and data collection, can be used to analyze various elements of a TV series.

Generative AI is also discussed, which refers to the use of deep multimodal algorithms to generate new content such as images, speech, and text. Generative methods like Generative Adversarial Networks (GANs) and Stable Diffusion can create new content that is almost indistinguishable from real data. While generating videos is more challenging, Recurrent Neural Networks (RNNs) like LSTMs can capture the temporal dynamics of the scenes to create interesting and promising applications for complex, but short duration videos.

KEYWORDS

Television series; television studies; data-driven audiovisual analysis; generative AI; content analysis.

Datafication and Data-Driven Approaches in the Humanities

Data-driven approaches have become increasingly popular in the humanities for a variety of reasons. The growing availability of digital data means that researchers can analyze, model and interpret cultural artifacts like books, paintings, photographs, and music using computational tools and techniques. This has opened up new avenues for research and analysis that were not previously possible. One example is digital humanities, an interdisciplinary field that emerged from the use of computational methods to study history, literature, arts, and other forms of human culture.

Data-driven approaches allow for the processing of large amounts of data quickly and efficiently, which would be difficult or impossible to do manually. This makes it possible to identify patterns and trends in the data, leading to new insights and discoveries. Data-driven approaches are inherently interdisciplinary, drawing on techniques and methods from computer science, statistics, mathematical physics, and other fields. This has fostered new collaborations and partnerships between humanities scholars and researchers in other disciplines, leading to new perspectives and new areas of inquiry that were previously impossible or impractical to investigate. For example, it is now possible to use computational tools to analyze the spread of ideas and cultural trends across different geographic regions, using data drawn from social networks or other digital sources such as the Web and search engines. Data-driven approaches also allow research findings to be communicated in innovative ways. Interactive visualizations, digital archives, and multimedia presentations enable broader public engagement with humanities research and improve its dissemination.

Applications in Television Studies

The field of data-driven approaches has extended in recent years to the realm of television studies, particularly in the analysis of audience reception and engagement in television programming. By analyzing data from social media platforms like Facebook, Instagram, and especially Twitter researchers can identify trends in how audiences respond to specific shows or episodes. This approach can provide insights into how audiences interpret and react to various elements of a programme, such as characters, storylines, or themes (Antelmi et al. 2018, Burgess and Bruns 2015, Crespo-Pereira and Juanatey-Boga 2017, Ibrahim et al. 2018, Molteni and Ponce de Leon 2016). Social media has become a valuable source of data for analyzing audience reactions to television shows. There are several methods of using social media data to gain insights into audience responses. Sentiment analysis is one such approach, which involves analyzing the perception of social media posts related to a particular television show. By identifying patterns in audience reactions, such as which characters or storylines are most popular or controversial, sentiment analysis can provide valuable insights into how audiences interpret and respond to different aspects of a show (Antonakaki et al. 2021, Giachanou and Crestani 2016, Rocchi 2022a, 2022b, Scharl et al. 2016). Of particular interest is the use of social media as a second screen: second screen engagement allows viewers to share their opinions, reactions, and commentary with a broader audience and to engage in online live conversations around TV shows (Buschow et al. 2014, Pehlivan 2021, Williams and Gonlin 2017). Another approach is to analyze the use of hashtags related to a television show (Pilař et al. 2021). Both second screen and hashtag analysis can help to identify trends and themes in audience reactions by providing insights into how audiences are discussing the show on social media.

Network analysis involves mapping the relationships between different social media users and their interactions with each other. By identifying key influencers and opinion leaders within the audience, network analysis can provide insights into how opinions and reactions are shared and spread across social media (Hecking et al. 2017, Rocchi 2022b). Content analysis is a further method of analyzing social media data related to a television show. This involves analyzing the text and images of social media posts to gain insights into the specific themes and topics that audiences are discussing (Hagy et al. 2019). Text mining and Natural Language Processing (NLP)

can also be used to analyze large amounts of textual data, such as reviews, to identify patterns in audience responses. By identifying common themes and issues that emerge across different reviews, researchers can gain a deeper understanding of how audiences interpret and respond to cultural artifacts. A very useful tool for processing the data provided by content analysis is topic modeling, a linguistic analysis technique discussed below.

Another way that data-driven approaches are being applied in television studies is through the use of digital archives. Public archives like the Internet Archive's TV News (<http://archive.org/details/tv>) provide access to an enormous amount of historical television content that can be analyzed using computational techniques. This allows researchers to investigate patterns and trends in television programming over time, such as changes in genres, formats, or styles. This kind of approach has already enabled, in the field of film, the valorisation of archives and audiovisual materials from the perspective of enhancing cultural heritage and public history. Many projects have been developed, both internationally, such as i-media-cities (<https://www.imediacities.eu/>, Sala and Bruzzo 2019) and pilot projects dedicated to a single archive, such as PH-Remix (<http://www.labcd.unipi.it/ph-remix/>). Such projects can be developed to enhance the enormous amount of archival material available in the field of television.

Beyond analyzing the reception and history of television programming, data-driven approaches are also being employed to study the production and distribution of television shows. For instance, network analysis can be used to map the relationships between different producers, networks, and platforms, shedding light on the complex networks that underlie the production and distribution of television content (Ruffino and Brembilla 2016, Fanchi and Tarantino 2020).

Artificial Intelligence

Undoubtedly, the application of Artificial Intelligence (AI) has brought a new dimension to the analysis of audiovisual text, which involves scrutinizing textual information within audio and video content. AI systems have been developed with a range of techniques to extract and analyze text from audio and video content. Speech recognition involves using AI algorithms to transcribe spoken language into text, while image recognition utilizes AI algorithms to identify and classify objects and text within images. Additionally, NLP is employed to analyze and understand human language

in text that is embedded in audio and video content, such as transcripts, subtitles, and closed captions. The most common applications of AI in the field of audio-visuals involve automatic transcription and translation. AI systems can automatically generate closed captions for videos by analyzing the spoken content and transcribing it into text and translate subtitles in videos from one language to another. Speech-to-text transcription is possible through AI systems, enabling users to search and analyze spoken content from audio and video recordings. AI systems can analyze video content to identify key topics, themes, and sentiment expressed in the video. Artificial intelligence is often categorized into different types based on their functions and capabilities, like classification and prediction, but also generating new artificial content, as we will discuss.

Classification is a type of AI that involves assigning a label or category to a given input or data point. This task is accomplished by training the AI system to recognize patterns in the input data and using those patterns to make predictions about which category the input belongs to. For instance, an AI system can identify and classify images into different categories such as animals, vehicles, or objects. Another example of classification is sentiment analysis where an AI system can analyze text data and classify it as positive, negative, or neutral based on the sentiment expressed in the text.

Topic modeling is a powerful technique used in NLP and machine learning to uncover the underlying themes and topics within a corpus of text data. One of the most widely used algorithms for topic modeling is Latent Dirichlet Allocation (LDA) (Blei et al. 2003). LDA is based on the assumption that each document in a corpus contains a mixture of topics, and each word in a document is generated by one of those topics. This probabilistic model uses Bayesian inference to estimate the topic distributions, making it a powerful tool for discovering the topics that exist in a text corpus. Another popular algorithm used for topic modeling is Non-negative Matrix Factorization (NMF) (Lee and Seung 1999), which decomposes a term-document matrix into two lower-rank matrices representing the topics and their corresponding weights. Unlike LDA, NMF is a non-probabilistic algorithm that requires an unsupervised learning algorithm to reduce the dimensionality of data into lower-dimensional spaces. Other algorithms that are quite widely used and often derived from the previous two are the Hierarchical Dirichlet Process (HDP), the Correlated Topic Model (CTM), and the Structural Topic Model (STM), that allow for the modeling of topics as a function of covariates, such as time, authorship, or genre.

Recently, there has been a growing interest in the use of deep learning models for topic modeling, particularly with the emergence of powerful pre-trained language models such as BERT and GPT-3. These models have been shown to perform well in a variety of NLP tasks, including topic modeling. One example of a deep learning-based topic modeling algorithm is Bertopic (Grootendorst 2022), which uses BERT as a feature extractor and applies clustering techniques to group similar documents into topics. Bertopic has been shown to outperform traditional topic modeling algorithms such as LDA and NMF on several benchmark datasets (Egger and Yu 2022). While deep learning-based topic modeling algorithms have shown promising results, they also require large amounts of training data and computational resources. Furthermore, their black-box nature can make it difficult to interpret the resulting topics, which may limit their usefulness in certain applications. Nonetheless, the use of deep learning models for topic modeling is an active area of research with potential for further advancements in the field.

Prediction is an important AI technique in the audiovisual field that involves using historical data to forecast future events or outcomes. AI models are trained on past audiovisual data and can learn to recognize patterns and relationships in that data. The system then utilizes these patterns to make predictions about future audiovisual events. For instance, an AI model can analyze historical data on user behavior and predict which streaming platform a user is likely to use based on their viewing history. This information can help streaming platforms optimize their content recommendations and improve user engagement. Another example is content prediction, where AI models can analyze historical data on viewing patterns to predict which type of content is likely to be popular in the future. These predictive algorithms can help streaming platforms reduce churn rates, improve retention, and increase revenue.

These models use a variety of statistical techniques to analyze large datasets and identify relationships and patterns that can be used to predict future outcomes. In machine learning, predictive models are typically trained using supervised learning techniques, which involve providing the algorithm with labelled training data. The algorithm then uses this data to identify patterns and relationships between the input features and the output variable, and then generates a predictive model that can be used to make predictions on new, unseen data. Deep learning models, on the other hand, are a type of artificial neural network that are designed to learn and make predictions, partially inspired by how the human brain work. Deep

learning models typically use multiple layers of artificial neurons to analyze and interpret data, and are particularly effective at processing large, complex datasets.

The most important predictive application in the audiovisual domain is that of recommendation engines (Avezzi 2017, Schrage 2020). The algorithms used by recommendation engines for streaming platforms are typically complex and rely on various machine learning and/or deep learning techniques. The most commonly used approaches include collaborative filtering, content-based filtering and hybrid filtering. Collaborative filtering is based on the idea that users who have similar consumption habits are likely to have similar preferences for songs, movies or TV shows. Collaborative filtering looks at a user's past behavior, such as the movies they have watched and rated, and then recommends new content based on the viewing history of other users who have similar tastes. Content-based filtering uses features of the movies or TV shows themselves, such as genre, director, or actors, to recommend new content to users. Hybrid filtering combines collaborative filtering and content-based filtering to provide more accurate recommendations. These approaches can be used by implementing various machine learning algorithms or to extract the features of a deep learning model that identifies patterns and relationships between movies or TV shows and user behavior.

Data-Driven TV Series Analysis

TV series can be complex in a number of ways, and these complexities can pose both creative and analytical challenges. In a more formal way we might say that TV series are (multimodal, multiscale) open complex systems. One aspect of TV series' complexity is their narrative structure. TV series often feature multiple plotlines (Pescatore and Rocchi 2019, Rocchi and Pescatore 2022) that intersect and evolve over the course of many episodes, and on different time scales. This can create a sense of continuity and depth that is difficult to achieve in other storytelling formats, like movies. Several approaches have recently been introduced to use quantitative and automatic data extraction and analysis methods for understanding and revealing these complexities in TV series and movies, sometimes using the temporal and statistical properties of shoots and scenes (Cutting et al. 2010, Guha et al. 2015, Huang et al. 2018, Miech et al. 2019, Rao et al. 2020). This also poses a challenge for the so-called *automated summarization*, an-

other interesting problem in TV series' quantitative analysis (Chen et al. 2004, Bost et al. 2019).

Another aspect of TV series' complexity is in their character development. TV series often feature large ensemble casts with complex relationships and motivations (Pescatore et al. 2014, Pescatore and Innocenti 2018). Characters may evolve over time, revealing new facets of their personality and backstory as the series progresses. A very instructive and interesting approach to quantitative *character identification*, without the script, is found in Nagrani and Zisserman (2018), but other approaches also appear in Zhang and colleagues (2008), Maldi and colleagues (2011), Tapaswi and colleagues (2012), and Naldi and Dalla Torre (2022). Restricted to audio, this becomes the well-known problem of *diarization*, where algorithms are developed to learn how to predict who is talking at a given time (Bredin and Gelly 2016, Skowron et al. 2016, Park et al. 2021).

TV series can also be complex in terms of their thematic content. An interesting and relatively new approach to the analysis of television series is that of content analysis, a mix of qualitative and quantitative methods based on the coding process and data collection. Content analysis can be used to quantify and analyze different elements of a TV series, such as formulaic aspects of a particular genre, e.g., legal drama (Rocchi and Sonogo 2022), the frequency of certain types of scenes or dialogue, the emotions portrayed by characters, or the cultural representations depicted in the show. It can also be used to identify patterns and trends over time, such as changes in the depiction of certain social issues or the portrayal of certain types of characters (Rocchi and Anselmeti 2021). Many series tackle complex social and political issues and explore nuanced moral and ethical questions. TV series are not only complex in terms of their internal narrative structures and character development, but they are also open systems that interact with a range of external factors (Pescatore 2018), such as the production process, the audience, and the broader cultural and social context. One way in which TV series interact with the production process is through the development and execution of the series itself. TV series are typically produced by a team of writers, directors, producers, and other professionals, who work together try to create a coherent and engaging story arc. The production process can have a significant impact on the final product, as decisions around casting, editing, and marketing can all influence the way the series is received by audiences. TV series are designed to be consumed over a period of time, and as such, they create a relationship between the audience and the story.

Viewers become invested in the characters and their fates and may engage in a range of activities to deepen their understanding of the series, such as online discussion forums, fan fiction, and social media engagement. The feedback and reactions of viewers can also influence the direction of the series, as writers and producers may adjust the storyline or characters in response to audience feedback.

TV series are also influenced by broader cultural and social trends. They may reflect or challenge prevailing cultural norms and values and can be used as a tool for social commentary or political messaging. The broader cultural context in which a series is produced can also influence its reception by audiences, as different social groups may interpret the series in different ways depending on their own cultural and ideological perspectives. This is especially true for long-running series distinctive of broadcast television, which adopt a particular production model. Episodes are produced in parallel with the airing of the series, a few months in advance. The peculiar property of TV series being an open system interacting with the 'physical' world comes with another characteristic: a very low latency with respect to external stimuli and the capability of embedding in their narrative, almost immediately, events or facts from the real world. As an example, the Supreme Court ruling on *Roe vs. Wade*, which made abortion illegal in many U.S. states, can be cited. The ruling is dated June 24, 2022. The medical series *New Amsterdam* devoted an extremely critical episode to it, which aired on November 1, 2022. Considering the production and airing times, the reaction time to a major external event can be estimated to be a couple of months. Actually we believe that there is also a retro action in this interaction: changes in the narrative structure might impact on real cultural or social processes.

Generative AI

Another interesting direction that in a relative short future might have a huge impact on the television industry and media studies is the generation of new content using deep multimodal algorithms. A deep multimodal algorithm, as noted above, is an approach to machine learning that combines information from multiple sources, such as text, images, and audio. It is designed to take advantage of the complementary nature of different modalities to improve the accuracy and richness of the analysis and generation of artificial data.

Deep multimodal algorithms have become increasingly popular in recent years, particularly in the fields of NLP and computer vision. They are used for a wide range of tasks, from image and speech recognition to sentiment analysis and recommendation systems. At the core (learning phase) of most deep multimodal algorithms is the Transformer model, which was introduced by Vaswani and colleagues (2017). Transformers use the so-called self-attention mechanisms to encode and decode information in a very efficient way, through a neural network architecture that is capable of processing sequential data, such as text, as well as non-sequential data, such as images. It has been used successfully since then in a variety of natural language and computer vision tasks, often also in multimodal environment (Xu et al. 2022). Example of this include the Vision-and-Language pre-training models designed to learn joint representations of visual and textual information, enabling it to perform a range of multimodal tasks, such as image captioning and visual question answering (Li et al. 2020, Chen et al. 2023).

Generative methods, such as Generative Adversarial Networks (GANs) and Stable Diffusion, have gained immense popularity in recent years for their ability to create new content (Shahriar 2021). These multimodal generative algorithms can be used in a variety of applications such as image synthesis, speech synthesis, and NLP. For instance, in the field of image synthesis, multimodal generative algorithms can be used to generate photo-realistic images that do not exist in reality but look like they could be real. In speech synthesis, they can be used to generate human-like speech that is almost indistinguishable from the real thing. And in natural language processing, they can be used to generate text that is coherent and semantically meaningful.

GANs (Goodfellow et al. 2014) work by training two neural networks simultaneously: a generator and a discriminator. The generator creates new data by generating samples from a random noise input, while the discriminator evaluates the generated samples and decides whether they are real or fake. The two networks are then trained together in a process called adversarial training, where the generator learns to create more realistic samples by fooling the discriminator. Over time, the generator becomes better at generating realistic samples that can fool the discriminator, and the discriminator becomes better at identifying real from fake samples (Karras et al. 2018, Brock et al. 2019, Huang et al. 2021).

Stable Diffusion (Sohl-Dickstein et al. 2015, Ho et al. 2020), on the other hand, is a generative method that involves iteratively refining a noise

input to generate a sample that matches a given target distribution, empirically defined by real images. It uses the principles of diffusion processes in physics (in particular non-equilibrium statistical mechanics) to transform real images through steps of “small gaussian noise”, until all the information is lost, while a deep network learns how to invert the global (irreversible) process. This backward process is then used to generate artificial high-quality images from noise, creating artificial but realistic images of human faces, animals, and landscapes (Ho et al. 2020, 2021, Rombach et al. 2022). Furthermore, the attention mechanism embedded in the transformer can be used to ‘condition’ the generated image to be constrained to some given text, producing amazing artificial images (but also music) with a given style or specific content (see, for example: <https://stablediffusionweb.com/>).

GANs and diffusion models have shown great success in generating high-quality images and text in a multimodal setup. However, generating videos is more challenging due to the additional temporal dimension involved. In fact, to date, apart from a few exceptions, all applications for video exist in the so-called *deepfake*, i.e. the exchange of faces of a character in a movie scene, often porn. In order to generate videos, the models need to be able to capture the temporal dynamics of the scenes, for example how the objects and people in the video move and interact with each other over time. This requires incorporating the temporal dimension into the models, which can be done in several ways. The most mature approach is to use Recurrent Neural Networks (RNNs), such as LSTMs (Long Short-term Memory network), to model the temporal dependencies in the video frames. The RNNs can be trained to predict the next frame in the video sequence given the previous frames, and eventually combined with GAN architectures for interesting and promising applications to create complex but short duration videos (Clark et al. 2014, Vondrick et al. 2016, Munoz et al. 2020). Attempts have also applied diffusion methods to videos that forecast interesting future directions (Ho et al. 2022a, 2022b).

The TV series industry and media studies can in future gain from the development and consolidation of these generative methods, but in the short term the reverse is true: TV series could represent an ideal testing ground for the development of new deep multimodal algorithms for complex and long video analysis and generation. Developing multimodal deep algorithms suited to TV series data (for analysis or generative tasks) would also imply the creation of tools for a better exploration and understanding of certain facets of TV series’ complexity. As the technology advances and the data

becomes more diverse and abundant, we can expect to see more exciting and innovative ways of using these algorithms to enhance the TV viewing experience, as well as in production. One future application of generative multimodal deep learning algorithms in the realm of TV series could be, for example, in the creation of personalized interactive TV shows.

Overall, generative multimodal deep learning algorithms have the potential to revolutionize how we design, produce, create and consume TV shows, opening up new possibilities for personalization and interactivity. We cannot precisely predict the future of these generative AI algorithms as they apply to TV production, but change will come and, considering the accelerations of the last five years, mostly probably it will come sooner than expected. It is hard to predict the details of the impact on the TV industry and media studies, but there will be social, economic, and cultural impacts. Considering what is happening at time of writing (in April 2023) with “just one ChatBot” (<https://www.bbc.com/news/technology-65139406>, <https://www.wired.it/article/openai-blocca-chatgpt-in-italia/>), we could just about imagine what will happen when we can generate artificial media content of high complexity, long duration and realistic appearance.

BIBLIOGRAPHY

- Antelmi, Alessia, John Breslin and Karen Young (2018). “Understanding User Engagement with Entertainment Media: A Case Study of The Twitter Behaviour of Game of Thrones (GoT) Fans.” In *2018 IEEE Games, Entertainment, Media Conference (GEM)*, 1-9. <https://doi.org/10.1109/GEM.2018.8516505>.
- Antonakaki, Despoina, Paraskevi Fragapoulou and Sotiris Ioannidis (2021). “A Survey of Twitter Research: Data Model, Graph Structure, Sentiment Analysis and Attacks.” *Expert Systems with Applications* 164, 114006. <https://doi.org/10.1016/j.eswa.2020.114006>.
- Avezzi, Giorgio (2017). “The Data Don’t Speak for Themselves: The Humanity of VOD Recommender Systems.” *Cinéma & Cie* 17(29): 51-66.
- Blei, David, Andrew Ng and Michael Jordan (2003). “Latent Dirichlet Allocation.” *Journal of Machine Learning Research* 3: 993-1022.
- Bost, Xavier, Serigne Gueye, Vincent Labatut, Martha Larson, Georges Linares, Damien Malinas and Raphaël Roth (2019). “Remembering Winter Was Coming: Character-Oriented Video Summaries of TV Series.” *Multimedia Tools and Applications* 78(24): 35373-99. <https://doi.org/10.1007/s11042-019-07969-4>.
- Bredin, Hervé and Grégory Gelly (2016). “Improving Speaker Diarization of TV Series Using Talking-Face Detection and Clustering.” In *Proceedings of the 24th ACM International Conference on Multimedia*, 157-61. <https://doi.org/10.1145/2964284.2967202>.
- Brock, Andrew, Jeff Donahue and Karen Simonyan (2019). “Large Scale GAN Training for High Fidelity Natural Image Synthesis.” In *International Conference on Learning Representation*, 1-29. <https://doi.org/10.48550/arXiv.1809.11096>.
- Burgess, Jean and Axel Bruns (2015). “Easy Data, Hard Data: The Politics and Pragmatics of Twitter Research after the Computational Turn.” In *Compromised Data: From Social Media to Big Data*, edited by Greg Elmer, Ganaele Langlois and Joanna Redden, 93-111. New York: Bloomsbury Academic.
- Buschow, Christopher, Beate Schneider and Simon Ueberheide (2014). “Tweeting Television: Exploring Communication Activities on Twitter while Watching TV.” *Communications* 39(2): 129-149. <https://doi.org/10.1515/commun-2014-0009>.
- Chen, Fei-Long, Du-Zhen Zhang, Ming-Lun Han, Xiu-Yi Chen, Jing Shi, Shuang Xu and Bo Xu (2023). “VLP: A Survey on Vision-Language Pre-Training.” *Machine Intelligence Research* 20(1): 38-56. <https://doi.org/10.1007/s11633-022-1369-5>.

- Chen, Hsuan-Wei, Jin-Hau Kuo, Wei-Ta Chu and Ja-Ling Wu (2004). "Action Movies Segmentation and Summarization Based on Tempo Analysis." In *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 251-58. <https://doi.org/10.1145/1026711.1026752>.
- Clark, Aidan, Jeff Donahue and Karen Simonyan (2019). "Adversarial Video Generation on Complex Datasets." *arXiv:1907.06571*. <https://doi.org/10.48550/arXiv.1907.06571>.
- Crespo-Pereira, Verónica and Óscar Juanatey-Boga (2017). "Spanish TV Series on Twitter: What Social Media Audiences Say." In *Media and Metamedia Management*, edited by Francisco Campos Freire et al., 435-440. Cham: Springer.
- Cutting, James E., Jordan E. DeLong and Christine E. Nothelfer (2010). "Attention and the Evolution of Hollywood Film." *Psychological Science* 21(3):432-39. <https://doi.org/10.1177/0956797610361679>.
- Egger, Roman and Joanne Yu (2022). "Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts." *Frontiers in Sociology* 7. <https://doi.org/10.3389/fsoc.2022.886498>.
- Fanchi, Mariagrazia and Tarantino Matteo (2020). "Has Digital Transformation Impacted Gender Imbalance in Italian Cinema? A Data analysis of Creative Clusters 2004-2016." In *Digital Transformation in the Cultural and Creative Industries*, edited by Marta Massi, Marilena Vecco and Yi Lin. London: Routledge.
- Giachanou, Anastasia and Fabio Crestani (2016). "Like it or not: A Survey of Twitter Sentiment Analysis Methods." *ACM Computing Surveys (CSUR)* 2: 1-41. <https://dx.doi.org/10.1145/2938640>.
- Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio (2014). "Generative Adversarial." In *Proceedings of the International Conference on Neural Information Processing Systems*, 2672-2680. <https://doi.org/10.1145/3422622>.
- Grootendorst, Maarten (2022). "BERTopic: Neural topic modeling modelling with a class-based TF-IDF procedure." *arXiv arXiv:2203.05794v1*. <https://doi.org/10.48550/arXiv.2203.05794>.
- Guha, Tanaya, Naveen Kumar, Shrikanth S. Narayanan and Stacy L. Smith (2015). "Computationally deconstructing movie narratives: An informatics approach." In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2264-68. <https://doi.org/10.1109/ICASSP.2015.7178374>.
- Hagy, Madeleine, Ansley Hamilton and Caroline Miller (2019). "The Golden Age of Data, #ThisIsWe. A Content Analysis of Live Tweets of This Is Us." In *The Golden Age of Data. Media Analytics in Study & Practice*, edited by Don Grady. New York: Routledge.
- Hecking, Tobias, Vania Dimitrova, Antonija Mitrovic and H. Ulrich Hoppe (2017). "Using Network-Text Analysis to Characterise Learner Engagement in Active Video Watching." In *Proceedings of the 25th International Conference on Computers in Education*, edited by Chen W. et al., 326-335.
- Ho, Jonathan, Ajay Jain and Pieter Abbeel (2020). "Denosing Diffusion Probabilistic Models." *Advances in Neural Information Processing Systems* 33: 6840-6851. <https://arxiv.org/abs/2006.11239>.

- Ho, Jonathan, Chitwan Saharia, William Chan, David J. Fleet, Mohammad Norouzi and Tim Salimans (2022). "Cascaded Diffusion Models for High Fidelity Image Generation." *Journal of Machine Learning Research* 23(47): 1-33. *arXiv* <http://arxiv.org/abs/2106.15282>.
- Ho, Jonathan, Tim Salimans, Alexey Gritsenko, William Chan, Mohammad Norouzi and David J. Fleet (2022). "Video Diffusion Models." *arXiv* <http://arxiv.org/abs/2204.03458>.
- Ho, Jonathan, William Chan, Chitwan Saharia, Jay Whang, Ruiqi Gao, Alexey Gritsenko, Diederik P Kingma, Ben Poole, Mohammad Norouzi, David J. Fleet and Tim Salimans (2022). "Imagen Video: High Definition Video Generation with Diffusion Models." *arXiv* <https://arxiv.org/abs/2210.02303>.
- Huang, Qingqiu, Yuanjun Xiong, Yu Xiong, Yuqi Zhang and Dahua Lin (2018). "From Trailers to Storylines: An Efficient Way to Learn from Movies." *arXiv* <https://arxiv.org/abs/1806.05341>.
- Huang, Zhizhong, Shouzhen Chen, Junping Zhang and Hongming Shan (2021). "PFA-GAN: Progressive Face Aging with Generative Adversarial Network." In *IEEE Transactions on Information Forensics and Security* 16: 2031-45. <https://doi.org/10.1109/TIFS.2020.3047753>.
- Ibrahim, Ania, Ahmed Elbagoury, Mohamed S. Kamel and Fakhri Karray (2018). "Tools and Approaches for Topic Detection from Twitter Streams: Survey." *Knowledge and Information Systems* 3: 511-539. <https://doi.org/10.1007/s10115-017-1081-x>.
- Innocenti, Veronica and Guglielmo Pescatore (2018). "The Evolution of Characters in TV Series: Morphology, Selection, and Remarkable Cases in Narrative Ecosystems." In *Reading Contemporary Television Universes. A Narrative Ecosystem Framework*, edited by Ilaria A. De Pascalis and Paola Brembilla. New York and London: Routledge.
- Karras, Tero, Timo Aila, Samuli Laine and Jaakko Lehtinen (2018). "Progressive Growing of GANs for Improved Quality, Stability, and Variation." *arXiv* <http://arxiv.org/abs/1710.10196>.
- Lee, Daniel and H. Sebastian Seung (1999). "Learning the Parts of Objects by Non-negative Matrix Factorization." *Nature* 401: 788-791. <https://doi.org/10.1038/44565>
- Li, Linjie, Yen-Chun Chen, Yu Cheng, Zhe Gan, Licheng Yu and Jingjing Liu (2020). "HERO: Hierarchical Encoder for Video+Language Omni-Representation Pre-Training." *arXiv* <http://arxiv.org/abs/2005.00200>
- Maidi, Madjid, Veronica Scurtu and Marius Preda (2011). "Characters Identification in TV Series." In *2011 IEEE International Symposium on Multimedia*, 143-48. <https://doi.org/10.1109/ISM.2011.31>.
- McCallum, Shiona (2023). "ChatGPT banned in Italy over privacy concerns." *BBC NEWS*
- Miech, Antoine, Dimitri Zhukov, Jean-Baptiste Alayrac, Makarand Tapaswi, Ivan Laptev and Josef Sivic (2019). "HowTo100M: Learning a Text-Video Embedding by Watching Hundred Million Narrated Video Clips." In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2630-40. <https://doi.org/10.1109/ICCV.2019.00272>.
- Molteni, Luca and Joaquin Ponce de Leon (2016). "Forecasting with Twitter Data: An Application to USA Tv Series Audience." *International Journal of Design & Nature and Ecodynamics* 3: 220-29. <https://doi.org/10.2495/DNE-V11-N3-220-229>.
- Mottaghinia, Zahib, Mohammad-Reza Feizi- Derakhshi, Leili Farzinvas and Pedram Salehpour (2021). "A Review of Approaches for Topic Detection in

- Twitter.” *Journal of Experimental & Theoretical Artificial Intelligence* 5: 747-773. [10.1080/0952813X.2020.1785019](https://doi.org/10.1080/0952813X.2020.1785019)
- Munoz, Andres, Mohammadreza Zolfaghari, Max Argus, and Thomas Brox (2021). “Temporal Shift GAN for Large Scale Video Generation.” In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 3179-3188. *arXiv* <http://arxiv.org/abs/2004.01823>.
- Nagrani, Arsha and Andrew Zisserman (2018). “From Benedict Cumberbatch to Sherlock Holmes: Character Identification in TV Series without a Script.” *arXiv* <http://arxiv.org/abs/1801.10442>.
- Naldi, Maurizio and Paola Dalla Torre (2022). “Uncovering the Narrative Structure of “Breaking Bad” Through a Multi-Dimensional Quantitative Analysis.” *SERIES – International Journal of TV Serial Narratives* 8(1): 17-32. <https://doi.org/10.6092/issn.2421-454X/14503>
- Park, Tae Jin, Naoyuki Kanda, Dimitrios Dimitriadis, Kyu J. Han, Shinji Watanabe and Shrikanth Narayanan (2022). “A Review of Speaker Diarization: Recent Advances with Deep Learning.” *Computer Speech & Language* 72, 101317. *arXiv* <http://arxiv.org/abs/2101.09624>
- Pehlivan, Zeynep (2021). “Linking Twitter Archives with Television Archives.” In *The Past Web*, edited by Daniel Gomes, Elena Demidova, Jane Winters and Thomas Risse. Cham: Springer. https://doi.org/10.1007/978-3-030-63291-5_10.
- Pescatore, Guglielmo (ed.) (2018). *Ecosistemi Narrativi. Dal fumetto alla serie tv*. Roma: Carocci.
- Pescatore, Guglielmo and Marta Rocchi (2019). “Narration in Medica Dramas. Interpretative Hypotheses and Research Perspectives.” *La Valle dell’Eden* 34: 107-115.
- Pescatore, Guglielmo, Veronica Innocenti and Paola Brembilla (2014). “Selection and Evolution in Narrative Ecosystems. A Theoretical Framework for Narrative Prediction.” In *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW) IEEE Computer Society*, 1-6. <https://doi.org/10.1109/ICMEW.2014.6890658>.
- Pilař, Ladislav, Lucie Kvasničková Stanislavská, Roman Kvasnička, Petr Bouda and Jana Pitrová (2021). “Framework for Social Media Analysis Based on Hashtag Research.” *Applied Sciences* 11(8): 3697. <https://doi.org/10.3390/app11083697>.
- Rao, Anyi, Linning Xu, Yu Xiong, Guodong Xu, Qingqiu Huang, Bolei Zhou and Dahua Lin (2020). “A Local-to-global Approach to Multi-modal Movie Scene Segmentation.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10146-10155. *arXiv* <http://arxiv.org/abs/2004.02678>.
- Rocchi, Marta (2022a). “Applying Automatic Text Analysis Methodologies to Audiovisual Serial Product.” *Cinergie—Il Cinema e le altre Arti* 22: 173-187. <https://doi.org/10.3390/app11083697>.
- Rocchi, Marta (2022b). “Environmental Misinformation and Audiovisual Serial Narratives. An Automatic Analysis of the Twitter Social Discursiveness on ‘Seaspiracy’.” *Elephant & Castle* 28: 124-134.
- Rocchi, Marta and Allegra Sonogo (2022). “Tribunali e giustizia nelle narrazioni seriali del prime-time statunitense.” *Griseldaonline* 21(2): 79-93. <https://doi.org/10.6092/issn.1721-4777/15434>.

- Rocchi, Marta and Guglielmo Pescatore (2022). "Modeling Narrative Features in TV Series: Coding and Clustering Analysis." *Humanities and Social Sciences Communications* 9(333). <https://doi.org/10.1057/s41599-022-01352-9>
- Rocchi, Marta and Martina Anselmeti (2021). "L'evoluzione del legal drama, stereotipi di genere e il caso di For the People (ABC 2018-2019)." *Ocula* 22(25): 229-245. <https://doi.org/10.12977/ocula2021-16>.
- Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser and Björn Ommer (2022). "High-resolution Image Synthesis with Latent Diffusion Models." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10684-10695. *arXiv* <http://arxiv.org/abs/2112.10752>.
- Ruffino, Marco and Paola Brembilla (2016). "Narrative Ecosystems Through the Network Analysis Lens. Step One: The Production of U.S. TV Series, Between Capital and Labor Strategies." *SERIES – International Journal of TV Serial Narratives* 2(1): 55-68. <https://doi.org/10.6092/issn.2421-454X/6164>.
- Sala, Teresa M. and Mariona Bruzzo (eds) (2019). *I-MEDIA-CITIES Innovative e-environment for Research on Cities and the Media*. Barcellona: Edicions de la Universitat de Barcelona.
- Scharl, Arno, Alexander Hubmann-Haidvogel, Alistair Jones, Daniel Fischl, Ruslan Kamolov, Albert Weichselbraun and Walter Rafelsberger (2016). "Analyzing the Public Discourse on Works of Fiction-Detection and Visualization of Emotion in online Coverage about HBO's Game of Thrones." *Information Processing & Management* 52(1): 129-138. <https://doi.org/10.1016/j.ipm.2015.02.003>.
- Schrage, Michael (2020). *Recommendation Engines*. Boston: Mit Press.
- Shahriar, Sakib (2021). "GAN Computers Generate Arts? A Survey on Visual Arts, Music, and Literary Text Generation using Generative Adversarial Network." *arXiv* <http://arxiv.org/abs/2108.03857>.
- Skowron, Marcin, Martin Trapp, Sabine Payr and Robert Trapp (2016). "Automatic Identification of Character Types from Film Dialogs." *Applied Artificial Intelligence* 30(10): 942-73. <https://doi.org/10.1080/08839514.2017.1289311>.
- Sohl-Dickstein, Jascha, Eric A. Weiss, Niru Maheswaranathan and Surya Ganguli (2015). "Deep Unsupervised Learning using Nonequilibrium Thermodynamics." In *International Conference on Machine Learning*, 2256-2265. PMLR. *arXiv* <http://arxiv.org/abs/1503.03585>.
- Tapaswi, Makarand, Martin Bauml and Rainer Stiefelbogen (2012). "Knock! Knock! Who Is It; Probabilistic Person Identification in TV-Series." In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2658-65. <https://doi.org/10.1109/CVPR.2012.6247986>.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin (2017). "Attention is All you Need." In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000-10.
- Vondrick, Carl, Hamed Pirsiavash and Antonio Torralba (2016). "Generating Videos with Scene Dynamics." *Advances in Neural Information Processing Systems* 29. *arXiv* <http://arxiv.org/abs/1609.02612>.
- Williams, Apryl and Vanessa Gonlin (2017). "I Got All My Sisters with Me (on Black

- Twitter): Second Screening of How to Get Away with Murder as a Discourse on Black Womanhood.” *Information, Communication & Society* 20(7): 984-1004. <https://doi.org/10.1080/1369118X.2017.1303077>.
- Xu, Peng, Xiatian Zhu and David A. Clifton (2022). “Multimodal Learning with Transformers: A Survey.” *arXiv* <https://doi.org/10.48550/arXiv.2206.06488>.
- Zhang, Yi-Fan, Changsheng Xu and Hanqing Lu (2008). “Automatic Character Identification in Feature-length Films.” In *2008 IEEE International Conference on Multimedia and Expo*, 1469–72. <https://doi.org/10.1109/ICME.2008.4607723>.
- Zimbra, David, Ahmed Abbasi, Daniel Zeng and Hsinchun Chen (2018). “The State-of-the-art in Twitter Sentiment Analysis: A Review and Benchmark Evaluation.” *ACM Transactions on Management Information Systems (TMIS)* 2: 1-29. <https://doi.org/10.1145/3185045>
- Zorloni, Luca (2023). “OpenAi blocca ChatGPT in Italia” *Wired*. <https://www.wired.it/article/openai-blocca-chatgpt-in-italia/> (last accessed 20-04-23).

EXPLORING TV SERIALITY AND TELEVISION STUDIES THROUGH DATA-DRIVEN APPROACHES



© The Author(s) 2023 <https://doi.org/10.21428/93b7ef64.ec022085>.

In Giorgio Avezzù and Marta Rocchi (eds). *Audiovisual Data: Data-Driven Perspectives for Media Studies. 13th Media Mutations International Conference*. Bologna: Media Mutations Publishing. <https://doi.org/10.21428/93b7ef64.ddf0a578>.