

## **Automating Intersectional Analysis on Canadian VOD Services**

**Stéfany Boisvert**  
**University of Quebec at Montreal**  
**boisvert.stefany@uqam.ca**

Since 2010, several French-language subscription video-on-demand (SVOD) and broadcaster video-on-demand (BVOD) services have emerged in Canada. While these services provide access to foreign series and films dubbed in French, they have endeavored to distinguish themselves by offering an increasing number of ‘original’ productions. Despite the cultural significance of these French-language Canadian streaming services, their original content has received scant scholarly attention. Beginning in 2020, I have therefore undertaken a comprehensive study to systematically catalog all original French-language content—encompassing series, films, and single programs—produced for Canadian-owned SVOD services. This study includes both qualitative and quantitative analyses to discern dominant narrative and aesthetic trends and assess the diversity of character representations (Boisvert, Gagnon & Boisvert, 2024; Boisvert, 2024). While analyzing a substantial corpus is crucial—or at least very useful—to understanding the production and commissioning strategies of SVOD services, I could not help but notice the significant limitations of manual analysis when confronting a large corpus, even with a team’s support.

Indeed, employing an intersectional approach to analyze dozens of series inevitably introduces limitations to the results obtainable. The extensive lengths of TV series present substantial challenges for conducting thorough narrative and character analyses, making the process both challenging and time intensive. Furthermore, although a detailed analytical framework is used, it is impossible for manual textual analysis to comprehensively address all variables at once, particularly when considering their intersections and the multifaceted implications they entail. This methodological limitation inherently curtails the ability to cross-reference data effectively, thereby constraining the scope of conclusions that can be drawn about both progress and ongoing marginalization within audiovisual content.

To address these challenges, the next phase of my research, which has just received funding by the Fonds de recherche du Québec – Société et culture (FRQSC), will be to conduct an automated analysis of audiovisual productions, an approach not yet pursued in Canada. Inspired by data feminism (Edwards & Esposito 2019; Bassett, Kember & O’Riordan 2020; D’Ignazio & Klein 2020; Losh & Wernimont 2019), as well as audiovisual studies adopting an automated approach (Avezzù & Rocchi 2023; Doukhan et al. 2018, 2022, etc.), the goal is to leverage algorithms and digital tools, in other words to add a computer-assisted analysis to my research framework, to conduct more comprehensive intersectional studies. Drawing inspiration from the method used by the Geena Davis Institute on Gender in Media (GD-IQ), my goal is to mobilize AI tools to: 1) enhance the richness of findings by providing robust statistical data (percentage of characters according to certain

sociodemographic categories, calculation of speaking time and screen presence, etc.); 2) facilitate the integration of diverse character data (complete the data and enable the crossing of a greater number of them); 3) allow for a more exhaustive and multidimensional analysis by integrating “unusual” perspectives and disregarded data. More specifically, a computer-assisted analysis shall enable: 1) soundtrack analysis to aid semantic assessments for information that cannot be simply deduced from visual data; and 2) statistical visual analysis of predefined textual data to enhance the precision, quantification, and contextualization of findings (e.g., quantified identification of character traits and identities).

As this research project is in its initial stages, I am currently evaluating various AI tools and models to determine which are best suited for my analysis. My presentation at the workshop will therefore focus on preliminary findings from these tests, fostering discussion on the potential advantages of integrating automated analysis techniques. Additionally, it will highlight key methodological challenges that must be addressed to effectively implement these technologies in our research framework.

Currently, two methodologies show promise and are under investigation. The initial method involves leveraging automated analysis with various open-source Python tools, such as those from the National Institute of Audiovisual (INA). This process entails:

- 1) Utilizing PySceneDetect for video segmentation into shots or scenes, allowing analysis on more manageable units (this step is optional).
- 2) Employing facial recognition to identify all characters (inaFaceAnalyzer). To ensure precision, the analysis could only include characters who appear for more than five minutes and have spoken at least once.
- 3) Conducting dialogue transcription using a tool such as InaSpeechSegmenter, which effectively segments audio tracks into music, noise, and speech components for accurate transcription (Doukhan et al., 2018).
- 4) Classifying and organizing character data with tools such as inaSpeechSegmenter and inaFaceAnalyzer, based on both visual and auditory recognition.

Utilizing Python tools for analysis offers considerable flexibility and customization. It enables the selection of the most appropriate tools from a comprehensive library of open-source packages tailored for data manipulation, machine learning, and image analysis. This adaptability allows for modifications to the algorithms according to specific research needs. However, as I will discuss, this method is inherently complex and time-consuming. It necessitates a good knowledge of Python programming—or, in my case, extended periods of learning and trial-and-error. Without such expertise, researchers may find themselves overly reliant on technical assistants.

With the rapid advancements in artificial intelligence and the development of large language models (LLMs), a new method of automated analysis using a multimodal model is now possible. This approach is considerably simpler to initiate, as it operates through textual prompts, allowing researchers to harness the power of deep learning without requiring extensive machine learning expertise. However, this method may incur higher

costs unless an effective open access AI model suitable for video analysis is available, which remains a significant challenge. Additionally, while the analysis is easier to initiate, its effectiveness closely depends on the researcher's skills in prompt engineering, and knowledge about the operation of multimodal models with moving images is still relatively scarce.

While acknowledging these challenges is essential, the initial testing phase using a multimodal model, specifically Gemini (Flash and Pro versions), has yielded interesting results. For this preliminary phase, I analyzed an episode from a TV series, to test the model's capabilities, refining prompts to fine-tune its performance, and assessing the utility and limitations of this approach. This initial exploration is crucial in determining the feasibility of deploying this technology for analyzing a larger corpus.

As I will demonstrate, analysis with a multimodal model allows for a quick and easy compilation of certain data. For instance, facial recognition is conducted quite easily and accurately, enabling the compilation of basic identification data (gender, approximate age, skin color, etc.). With accurately parameterized instructions, the AI model can identify all relevant characters for further analysis. LLMs are also very accurate in sentiment analysis and identifying characters' main personality traits. Furthermore, although it requires a bit more effort in prompt engineering, a premium/Pro model can be used to calculate a character's speaking time and screen presence. This capability enhances our ability to conduct a nuanced intersectional analysis, assessing the true extent of diversity and inclusion within television series.

However, during the initial phase of my testing, some challenges remained while employing Gemini. Particularly, it has difficulty accurately synchronizing text with images, which results in errors in identifying character names. When analyzing scripted content, the prompting must also enable the AI model to clearly distinguish between the character and the performer. Additionally, filters and "safety settings" can cause it to be overly cautious in identifying character traits or, conversely, to overinterpret certain behaviors as potentially offensive. The analysis thus requires vigilant monitoring by the researcher and often requires manual data correction—which, although time-consuming, is not necessarily problematic, as it ensures rigor in the analysis. It is also necessary to circumvent ethical issues related to gender attribution, since "data on gender identities that fall beyond the binary [e.g., trans, non-binary, etc.] is a complex question for big-data research" (Barotsi, Fanchi & Tarantino, 2023, 99).

The crafting of prompts (analytical instructions) must also be meticulously detailed, especially concerning the units of analysis, as the AI model often struggles to define what constitutes a 'scene'. More fundamentally, the challenge of prompt engineering is critical: even minor adjustments in the prompts can lead to variations in the results. To prevent this, more nuanced and elaborate instructions beyond the conventional input-output (IO) method is necessary, to ensure effectiveness in analysis. Another significant challenge, which I am currently exploring with a computer analyst specializing in LLMs, involves standardizing prompts to ensure consistency across a large corpus of audiovisual

productions. In other words, the challenges are substantial, but so are the possibilities for analysis.

Beyond basic data compilation tasks, the testing phase with the multimodal model indeed revealed unexpected applications, particularly by facilitating exploratory research approaches. One of the most significant outcomes was the AI model's ability to highlight elements within the datasets that were not initially apparent. While some discoveries were merely anecdotal, others, including specific dialogue moments, images, and narrative elements, provided unexpected insights. These were elements that had been overlooked during manual analysis but proved very useful. Given the extensive lengths of series, manual analysis inevitably misses certain potentially significant narrative, visual, and sonic details. Therefore, incorporating a LLM-assisted analysis into our methodology can unearth additional insights into character identities or power dynamics, "identify patterns and trends in the data, [thus] leading to new insights and discoveries" (Esposti & Pescatore, 2023, p. 24)

Another promising method involves requesting the AI model to compose a critical and intersectional summary of a series' narrative and character portrayal. For instance, during the testing phase, Gemini Pro generated a critical intersectional summary for the series under review (*L'Académie*), which highlighted class dynamics that had been previously overlooked. Employing such synthetic analyses across multiple series, and then combining them using the AI model, could thus yield substantial insights into the representational trends of a specific SVOD service. This approach could enhance our understanding of how diversity is incorporated, and the prevailing ideological frameworks within a specific platform's original content.

Automated analysis can also serve to incorporate divergent thinking into our research, thereby challenging and reassessing certain data points for their relevance. For instance, automated analysis can be enhanced using Chain of Thought Self-Consistency Prompting (CoT-SC): this technique involves prompting the model to generate various outputs. In doing so, the analysis is conducted by mobilizing a "mixture of experts," thus uncovering diverse interpretations of the same TV series or film. In this way, instead of merely recognizing the polysemy of media texts as an epistemological foundation, we could actively test the heuristic value of a "polysemy in action."

Ultimately, my research aims to provide a more comprehensive view of the current audiovisual production in Canada, and thus to contribute to a broader understanding of national services that remains in the shadow of global SVODs (Lobato, Scarlata & Cunningham, 2023). This work intends to inform public policy debates by positing hypotheses on how politico-economic contexts or specific systemic factors, such as new funding policies and hiring practices within the audiovisual sector, influence casting decisions, character representation, and thematic choices. While manual analysis has already highlighted disparities between public and private SVOD services (Boisvert 2024) and identified various factors influencing diversity in representation, integrating automated analysis could offer deeper insights and contribute significantly to a better understanding of diversity and inclusion in the streaming era.

## References

- Avezzù, Giogio & Marta Rocchi (eds.) (2023). *Audiovisual Data: Data-Driven Perspectives for Media Studies*. Bologna: Media Mutations Publishing.
- Barotsi, Rosa, Mariagrazia Fanchi & Matteo Tarantino (2023). Constructing an Open, Participatory Database on Gender (In)Equality in the Italian Film Industry. In Giogio Avezzù & Marta Rocchi (eds.), *Audiovisual Data: Data-Driven Perspectives for Media Studies*, pp. 87-104. Bologna: Media Mutations Publishing.
- Bassett Caroline, Sarah Kember & Kate O’Riordan (2020). *Furious: Technological Feminism and Digital Futures*. London: Pluto Press.
- Boisvert, Stéfany, Dominique Gagnon & Theo Boisvert (2024). Club Illico à la conquête des ados. *L’Académie, La Dérape* et la production de *teen series* québécoises à l’ère de la télévision en ligne. *Kinephanos*, 10(1).
- Boisvert, Stéfany (2024). Streaming *Diversité*: Exploring representations within French-language scripted series on Canadian SVOD services. *Convergence*, 30(4).
- D’Ignazio, Catherine & Lauren F. Klein (2020). *Data Feminism*. Cambridge: MIT Press.
- Doukhan, David, Géraldine Poels, Zojra Rezgui & Jean Carrive (2018). Describing gender equality in French audiovisual streams with a deep learning approach. *VIEW*, 7(14), pp. 103-122.
- Edwards, Erica B. & Jennifer Esposito (eds.) (2019). *Intersectional Analysis as a Method to Analyze Popular Culture: Clarity in the Matrix*. New York: Routledge.
- Esposti, Mirko Degli & Guglielmo Pescatore (2023). Exploring TV Seriality and Television Studies through Data-Driven Approaches. In Giogio Avezzù & Marta Rocchi (eds.), *Audiovisual Data: Data-Driven Perspectives for Media Studies*, pp. 23-40. Bologna: Media Mutations Publishing.
- Geena Davis Inclusion Quotient/Geena Davis Institute on Gender in Media (2022). *The Reel Truth: Women Aren’t Seen or Heard*: <https://seejane.org/wp-content/uploads/gdiq-reel-truth-women-arent-seen-or-heard-automated-analysis.pdf>
- Lobato, Ramon, Alexa Scarlata & Stuart Cunningham (2023). Conceptualizing the National and the Global in SVOD Original Production. In Amanda D. Lotz & Ramon Lobato (eds.), *Streaming Video: Storytelling Across Borders*, pp. 37-53. New York: New York University Press.
- Losh, Elizabeth & Jacqueline Wernimont (eds.) (2019). *Bodies of Information: Intersectional Feminism and the Digital Humanities*. Minneapolis: University of Minnesota Press.