

Comparatismi IO 2025

ISSN 2531-7547

<http://dx.doi.org/10.14672/20253098>

Reclaiming the Empirical: Artificial Intelligence and Literary Genres

Daniel Raffini

Abstract • The paper explores the relationship between AI and literary genres, highlighting their shared features. Both occupy a borderline space between creativity and interpretation, the author's intentionality and the critic's gaze. Another convergence is the tension between imposed rules and emerging patterns, which reflects the distinction between a normative and a descriptive conception. The study reviews several applied studies that employ AI for text classification, to bring out tendencies such as emotional dimension, cultural diversity, character networks, and paratextual aspects. Building on this body of work, it develops a theoretical reflection on how AI systems can contribute to redefining the category of literary genre. What emerges is a view of genre as an empirical and pragmatic construct, culturally situated and historically differentiated, composed of heterogeneous elements.

Keywords • Artificial Intelligence; Genre; Empirical; Theory

Abstract • Il saggio analizza il rapporto tra IA e generi letterari, mettendo in luce le caratteristiche comuni, come il collocarsi tra creatività e interpretazione, tra l'intenzionalità dell'autore e lo sguardo del critico, e la tensione tra regole imposte e pattern emergenti, che rinvia alla distinzione tra normatività e descrittivismo. Vengono poi passati in rassegna studi applicativi che impiegano l'IA per la classificazione dei testi, evidenziando tendenze rilevanti come la dimensione emozionale, la diversità culturale, le relazioni tra personaggi e gli aspetti paratestuali. Dalla revisione di tali contributi scaturisce una riflessione teorica su come i sistemi di IA possano contribuire a ridefinire la categoria di genere letterario. Ne emerge una concezione del genere come costruito empirico e pragmatico, culturalmente situato, storicamente differenziato e composto da elementi eterogenei.

Parole chiave • Intelligenza Artificiale; Genere; Empirico; Teoria

Ledizioni 

Reclaiming the Empirical: Artificial Intelligence and Literary Genres

Daniel Raffini

I. Between creation and interpretation

The notions of AI and literary genre have some aspects in common¹. The first one is that both concepts lie at the intersection of creativity and interpretation. Throughout its history, AI as a discipline has dealt with both designing creative tools and analytical tools, and in many cases the two elements are intertwined. Distinguishing between the two aspects is the first step to study the relationship between literature and AI. Since its origins, AI has been applied to both generation and analysis, albeit usually through different techniques and tools. The fact that today with the most recent Generative AI the same tool is employed both for the generation and the analysis of texts should not obscure the essential distinction between the creative act and the critical one. Like AI, literary genre is also a category that lies between creativity and interpretation, between the intentions of the author and those of the reader or the critic – and is indeed born from the encounter between the two. At the same time, however, one cannot overlook the unique convergence of creativity and interpretation that AI highlights – a convergence that recalls a long tradition in literary theory which conceives the artistic object as intrinsically bound to its interpretation. Literary genre, in this sense, is defined both by its textual tradition and by its critical reception.

Using AI for text creation or for text analysis entails different theoretical issues regarding literary genres. Today, we have many examples of literary text written with AI, that shows its potential as a creative tool². If we consider genre as a market category, the definitions of AI Novel and AI Poetry show strong appeal for publishers, as demonstrated by their insistence on highlighting the involvement of an AI system in the making of a work, which is conveyed through various paratextual elements, such as promotional bands on the cover, advertising, and explicit references in prefaces or on the back cover. This promotional dimension, moreover, frequently turns out to be both deceptive and sensationalistic: consider, for instance, the periodic announcement of the first novel written

¹ In this paper, I treat AI as an evolving concept from a diachronic perspective, considering everything that has been classified as AI at different points in history, ranging from rule-based and combinatorial systems to the latest machine learning and deep learning models. By doing so, I aim to overcome the so-called AI effect, i.e., the gradual exclusion of technologies from the domain of AI as they became more widespread and understandable to users (McCorduck, 2004). Instead, the goal is to view AI as a cultural modifier with a continuous history stretching from at least the mid-20th century to the present, following definitions like the one proposed by Farrell et al. (2025). I believe this is the most effective approach for integrating AI with a historicized concept such as literary genre and for evaluating its impact at a theoretical level.

² It is important to note that the most convincing evidence of the so-called AI Literature is found in cases where innovative processes of generation and human-machine interaction are conceived and directed by human artists. By contrast, the complete automation of writing based solely on prompts is not particularly convincing from a literary standpoint, as it tends to produce stereotyped or even biased forms. Approaches to human-machine collaboration are presented in Farina (2024) and Bajohr (2024).

with AI, or the self-publishing phenomenon (Cabezas-Clavijo, 2024). It becomes evident, therefore, that the phenomenon must be placed within a systematic and historical framework, that considers both the theoretical implications and the broader landscape of works produced with AI. From the standpoint of literary genre, these works raise important questions. For example, we may wonder how these texts relate to traditional literary genres or if literary texts produced by AI constitute a new literary genre. Broadening the perspective, this leads us to ask whether the modes of text generation (in this case, the use of AI) can be included within the definition of literary genre. This is a first example of how engaging with AI can help us raise new questions about the notion of literary genre and redefine its boundaries.

The systematization of AI Literature, however, exceeds both the scope and the length of this paper and would require more detailed work. Here, I will focus instead on the use of AI as a tool for text analysis from the perspective of literary genre. As we will see, this approach proves equally fruitful for deepening our understanding and for redefining the concept of literary genre within the contemporary literary and transmedia landscape. Used as an analytical tool, AI demonstrates its potential in managing large corpora and extracting features that can be considered in the attribution of texts to a literary genre. The very process through which AI isolates and identifies features diverges markedly from human analytical reasoning: it may attend to minimal traces or seemingly meaningless details that, to a human observer, would appear irrelevant. At the same time, the decision-making trajectories of AI systems often resist reconstruction, thereby precluding any full verification of the elements underlying their outcomes³. In short, there are many differences between human and AI analysis that need to be considered. In this regard, the distinction between the human and the artificial gaze proposed by Arielli is particularly relevant. The first feature of the artificial gaze is the “innocent eye” (Arielli, 2024, p. 95), meaning that AI perceives forms in a way comparable to a child’s perception, without awareness of their significance. At the same time, however, AI can reveal the “data unconscious” (Arielli, 2024, p. 107), its ability to process massive datasets enables it to uncover patterns beyond human perception. Another characteristic of AI is its role as a “quintessence machine”: it extracts a kind of Platonic essence of style from the texts it is trained on or analyzes, transforming “artifacts defined by their uniqueness into instances of a general idea, that could be potentially materialized in an infinity of variations” (Arielli, 2024, p. 108). Finally, AI’s imagination is counterfactual: “The use of AI in culture suggests that each cultural product is just one possibility within a vast latent space of alternatives, some of which may be equally or even more interesting and good” (Arielli, 2024, p. 110). Instead of treating such differences as insurmountable – and thus irreconcilable with a humanistic object of study like the literary text – one might regard the alternative modalities introduced by AI as an epistemological challenge: an opening toward the reframing of fundamental categories of literary theory, such as the notion of literary genre.

2. Given rules or emerging patterns?

³ It is the problem of non-explainability, which means that the system could make highly accurate predictions, but it is often unclear how or why it reached a decision. This lack of transparency makes it difficult for users to trust results, verify fairness, or detect biases. In sensitive fields like education, healthcare, or law, non-explainability raises concerns because decisions may affect people without offering understandable reasons behind them (McDermid 2021).

Another convergence between AI and literary genre lies in the dichotomy between rules and emerging patterns – that is, between an inductive and a deductive paradigm, a normative and an empirical approach. In the history of AI, this opposition is reflected in the alternation between rule-based and data-driven AI. The first relies on explicit rules and symbolic representations to reason in a logic-based way, while the latter learns patterns from large datasets through statistical and machine learning techniques. The first one implies rules given by humans; the second one identifies patterns and correlations without prior explicit instruction. AI systems that we use today are prevalently data-driven, but in the history of AI applied to literary creation we have many examples of rule-based systems. Some of the first story generators adopted an approach based on story grammar derived from literary genres (Pérez y Pérez, 2004). When it comes to analysis and genre identification, most of the methodologies rely on data-driven systems, but still in many cases it is necessary to annotate texts and even to give some rules to classify texts. We can see, then, that the interaction between given rules and emerging patterns in AI is quite relevant and not so easy to untangle, since the two elements may be intertwined.

Historically, literary genres have shown an affinity with the kind of formalization typical of the computational approach. In genre studies, scholars have long discussed the tension between fixed conventions and emerging patterns – a dichotomy closely aligned with the distinction between symbolic and data-driven AI. Traditional or classical conceptions emphasized the formal rules, structures, and conventions that defined a genre. From the Hellenistic interpretation of Aristotle's *Poetics* onward, emerges a normative notion of genre, where rules are applied to the creation and classification of cultural objects, just like in rule-based AI. With Romanticism, the situation shifts quite rapidly. In this regard, one may recall the well-known controversy over the Aristotelian unities, taken up for instance by August W. Schlegel and Alessandro Manzoni, who pointed out that Aristotle had in fact derived the unities of time, space, and action from his observation of the corpus of tragedies available in his time, as a simple statement of fact rather than as prescriptive rules. This raises the theoretical question underlying objections to a dogmatic view of literary genre: to what extent is it legitimate to transform observation into rules? Can practice be turned into norm? The Romantic – and later twentieth-century – reaction against a normative conception of literary genre rests precisely on a negative answer to these questions.

Starting from this, more contemporary theories stress that genres are not static entities, but dynamic, emergent practices shaped through social use and shifting patterns. Rhetorical genre studies propose that genres are best understood as recurring social actions, not merely textual structures (Miller 2015). The social value of the concept of genre is also emphasized by Todorov (1976), who links genres to the dominant ideology of a society and highlights the mutable nature of genres, that evolve through continuous negotiation, with boundaries growing fluid and permeable over time. Twentieth-century theoretical redefinitions also come to propose moves toward other categories with more flexible boundaries. Frye (1957) suggests shifting the concept of genre into the idea of modes and categorizes literature according to mode-driven frameworks, such as mythical, romantic, or ironic. Contemporary theory, thus, entails a descriptive notion of genre, where the analysis focuses on extracting patterns from a corpus of texts. The descriptive approach proves more suited to the analysis of the contemporary cultural landscape, and more generally one can argue for the greater relevance of the deductive approach in contemporary literary studies. Yet one cannot overlook the influence exerted by normative patterns on the very making of literature, also in contemporary writing. In the act of writing, an author carries with them a

repertoire of notions – or rules – concerning literary genres, which they may choose either to follow or to contravene, but which they cannot cease to know.

The tension between normative rule-based definitions and emergent adaptive patterns reflects the dichotomy between rules and data found in AI on multiple levels. A first level is a historical one: theory of genre and AI shows a similar evolution from a deductive to an inductive approach. In this evolution we can see the normative approach of genre theory coinciding with rule-based AI and the descriptive approach with data-driven AI. The analogy, of course, is not perfect, and it should be considered more broadly, taking into account also the combination of the two approaches. In the next paragraph we will see real cases of application of AI for genre classification. Most of them use machine learning, i.e. a data-driven approach. Nevertheless, we can see that both inductive and deductive methodologies are used. In fact, studies that employ AI for genre classification can be divided into those that provide the system with a definition of genre (based on rules or ontologies), and those that rely on machine learning to detect patterns and cluster texts into groups. In the following paragraph, we will examine some of these studies in greater detail, as they provide the basis for a subsequent theoretical reflection on the concept of literary genre in the context of digital contemporaneity.

3. Insights from applicative studies

In the fields of computer science and digital humanities, we find several applied studies on literary genres that investigate methods for analyzing and classifying texts in relation to common and recurring features, beginning with the identification of patterns emerging from large corpora⁴. Such studies rely on computational methodologies that make it possible to analyze vast amounts of textual data, identifying trends, similarities, and differences that are not always evident through traditional analysis. Many of these studies are grounded more in computer science than in literary theory. For this reason, the concept of genre does not always coincide with its meaning in literary theory and often covers a broader range of meanings and artistic expressions – such as music or visual art – as well as other disciplines, including media theory, linguistics, and rhetoric. In many of these studies, genre identification is directed more toward computational tasks than literary ones – for example, information retrieval, natural language processing, or data management. As a result, such research often devotes limited theoretical attention to the concept of literary genre. This does not prevent us, however, from analyzing their findings and reintegrating them more precisely into the field of literary studies.

These studies reveal a number of interesting, partly innovative features that may contribute to the discussion on literary genre. Such features arise less from the codified tradition of genre theory than from the “pure” perspective of the AI gaze – as described by Arielli – and from the texts themselves. This is especially evident in studies that adopt an inductive approach, relying less on rules and annotation and more on the extraction of patterns.

3.1 Emotions

One of the most significant elements that emerges is the attention to the emotional component. Recent research has shown that sentiment analysis – the automatic detection

⁴ For an overview on the different methodologies for automatic genre recognition - whose detailed treatment lies beyond the scope of this study – see Kuzman, 2023.

of emotions and attitudes in texts – can be surprisingly effective for identifying literary genres. Samothrakis and Fasli (2015) investigate whether the emotional content of sentences in works of fiction can be used to predict their genre. Using a corpus from Project Gutenberg, in which texts are tagged with genre metadata, each sentence was automatically annotated according to Ekman's six basic emotions, and the resulting emotional patterns were analyzed through machine learning. The analysis showed correlation between emotion and genres: for example, humor is negatively correlated with fear and positively correlated with surprise, but also with sadness; horror is strongly correlated with fear and somewhat with disgust; science fiction is negatively correlated with joy; western is positively correlated with anger and joy, negatively with sadness. Results show that emotional patterns nearly doubled classification accuracy compared to baselines.

Kim, Pado, and Klinger (2017), using again novels from the Gutenberg Project database, compared simple word-based methods with emotion-based ones. When they used only an emotion lexicon, i.e. a dictionary mapping words to basic emotions, the system correctly classified the genre of a text about 80% of the time, nearly as well as a traditional word-frequency method. More interestingly, when they combined emotional profiles with word-based features, accuracy rose to about 84%, suggesting that genres are partly defined by their characteristic emotional patterns. The combined approach highlights the complexity of the notion of genre, which cannot be identified through a single feature but rather arises from the interplay of multiple elements. Yako (2021) reached similar conclusions: in her study of eight genres, a model that included emotional features achieved a 75% accuracy rate, compared to 71% when relying only on vocabulary. The study shows that including emotion-related features improves classification performance compared to traditional methods, such as Bag of Words⁵. The results confirm that models based on emotional features are particularly effective for genres with distinctive emotional content, such as poetry and short stories, while they are less effective for genres with less specific or overlapping emotions, such as historical fiction and romance. The study introduces, thus, a difference among genres, suggesting that not all genres can be defined by the same features and encouraging a holistic approach to the notion of literary genre.

These findings point to an important shift: whereas genre theory has often relied on themes, forms, or stylistic conventions, computational analysis highlights the role of emotional structure. In this perspective, genres emerge not only as categories of style or content but also as affective configurations, defined by recurring distributions of emotions or by emotional arcs that unfold through a narrative. We should not think of a closed classification where each genre corresponds to a single emotion; rather, the cited studies reveal a system of interactions among emotions, whose distribution and correlation determine the genre⁶. In this sense, AI proves particularly useful in identifying the emotional conjunction in which a text is situated and whether that conjunction links it to other texts. AI thus pushes us to reconsider genre as a more fluid and probabilistic notion, where hybrid works can be situated in overlapping affective spaces rather than within rigid taxonomies.

⁵ The Bag of Words (BoW) methodology is a technique in natural language processing (NLP) and information retrieval for representing text data numerically. The Bag of Words model represents a text (such as a sentence or document) as an unordered collection of its words, disregarding grammar and word order but keeping track of the frequency of each word in the text.

⁶ Analyzing the emotional component of a text raises several issues that go beyond the scope of this study. However, it is useful to at least mention the distinction between considering the emotion in the texts or in the reader, which helps us understand the emotional component in a non-univocal dimension (Winko, 2023).

3.2 Cultural diversity

Although the main centers of power in AI are located in the United States, from a research standpoint the field is widely spread in non-Western contexts as well. Accordingly, the use of AI in relation to literary genres spans different geographical settings, breaking with the West-centric paradigm. This points toward a conception of genre that is more attentive to cultural diversity. Because it adapts to various datasets and learns from them, AI supports systematic machine-learning classification of non-Western corpora, but also texts from different historical periods, promoting an investigation of cultural contexts where the function of literary genre may differ. AI holds the potential to reshape the study of genre by making visible the cultural diversity that lies within literary traditions. In a study devoted to Uzbek poetry, Mengliev et al. (2024) demonstrate how computational models can distinguish between poetic forms by drawing on features specific to local traditions, showing that genre is encoded not only in universal structures but also in culturally embedded markers of expression. Galib et al (2023) propose a parallel investigation into classical Bengali literature, which underlines this point: by curating dedicated corpora and training language-specific models, researchers have revealed how South Asian literary heritage carries its own distinctive patterns of genre formation. In the Afrikaans context, Kotzé and Burgert (2024) have explored how multilingual language models can be fine-tuned to recognize genre in a low-resource literary tradition, further underscoring the value of AI in supporting minority and regional languages. Zhu et al. (2023) proposes an AI-based method to automatically classify modern Chinese poetry styles to support learners and researchers. The article shows that the classification of modern Chinese poetry relies on four culturally specific schools, which differ from Western genre systems. By using AI to classify these styles more accurately and objectively, the study highlights how literary genres are culturally diverse constructs.

Similar works confirm that AI-driven approaches to literary genre are not confined to the Western canon, but instead open pathways to compare and connect diverse literatures on their own terms. The broader implication for genre theory is significant: rather than viewing genres as stable, universal categories, these AI-based studies highlight their contingent, culturally situated nature. From this perspective, genres appear as evolving configurations shaped by local histories, linguistic traditions, and social practices. When connected with transcultural literary theory, this approach allows us to trace affinities and overlaps between traditions without reducing them to fixed taxonomies, making space for hybrid and diasporic forms that resist clear classification.

3.3 Character networks

Character networks – understood as the mapping of interactions between characters in a story – offer a way of translating narrative into a system of relations. The idea was introduced in literary studies by Moretti (2011), who applied network theory to plot analysis. In character networks, each character becomes a node and the links between them represent encounters, dialogues, or shared episodes. Plot is redrawn as a pattern of connections, where protagonists cluster at the center and minor figures orbit around them. Such an approach has been taken up by computational analysis to investigate how the structure of these relations might reveal something about genre. Rahul et al. (2021) directly attempted to classify books into genres by comparing the shapes of their character networks. They achieved encouraging results, reporting accuracy rates of around two-thirds on a large dataset of fanfiction, showing that the way characters interrelate can indeed

encode signals of genre, even if these signals are not perfectly consistent. Elson, Dames, and McKeown (2010) emphasize that much depends on how one extracts the networks in the first place: a network drawn from dialogue, for example, reflects not only the story's social structure but also the conventions of reported speech in that genre. Likewise, Hettinger et al. (2015), working with German novels, found that while social networks contribute useful information, genre classification improves significantly when such relational data are combined with linguistic or thematic features. By analyzing the bibliography, Labatut and Bost (2019) underscore the ambivalence: character networks have been employed to classify texts, detect roles, or trace social structures, but the process is often noisy and deeply shaped by corpus selection.

For literary studies, the lesson is that the social fabric of narrative does carry generic tendencies – an ensemble novel may weave dense connections across many figures, a biography may center on a single dominant protagonist, while adventure stories often divide into rival factions or teams. Yet these tendencies are fluid and historically contingent, making character networks more valuable as comparative instruments than as absolute classifiers. They encourage us to notice patterns that might otherwise escape attention, but they cannot on their own provide a stable taxonomy of genre. A further and promising direction lies in attending not only to the presence of ties but also to the roles that characters play within them (Valls-Vargas, Zhu, Ontañón 2014; Soares de Lima et al., 2025). Theories of narrative have long emphasized distinctions between protagonists, antagonists, helpers, and foils, or between major and minor figures. If network analysis could incorporate such role-based distinctions – marking who drives the action, who mediates, who obstructs – it would move closer to the interpretive categories familiar to literary scholars. Piper, Xu and Ruths (2024) annotated five types of character interactions to understand if genre and audience influence social structure in narrative. Character networks might enrich genre theory by showing not simply how many characters interact, but how their functions within the narrative structure differ across genres. The potential lies in bridging the quantitative mapping of relations with the qualitative insights of narratology, allowing us to see genre as a configuration of social roles as much as of formal or thematic conventions.

3.4 Paratexts

Paratexts – covers, titles, jacket copy – are thresholds that mediate between text and world. When AI systems analyze them to assign genres, they operationalize precisely this mediating function rather than the poetics of the works themselves. Early computer-vision studies showed that visual regularities in covers are legible enough for machines to “judge a book by its cover”, i.e., to predict genre from layout, imagery, and color schemes alone, even while acknowledging the intrinsic ambiguity of genre boundaries (Iwana et al., 2016). Subsequent benchmarking on large style datasets confirmed both the promise and the limits of image-only genre identification: state-of-the-art models do learn salient visual cues, but accuracy plateaus and errors cluster around hybrid or adjacent categories (Lucieri et al. 2020). A clear trend has therefore been toward multimodality: combining visual features with textual elements available on the cover (detected title and straplines) significantly outperform image-only baselines and better cope with cross-cultural design variation and fuzzy taxonomies. In this direction, Rasheed et al. (2023) have proposed improved multimodal architectures that integrate convolutional, textual, and metadata layers, demonstrating higher accuracy in multi-label recognition and, importantly, showing how composite paratexts can be read together to capture more fluid, overlapping genre assignments. Fine-grained analyses now go further, showing that not only the semantics of

cover words but their design – font style, position, and color – systematically correlates with genre signals (e.g., bold condensed type high on the page for mystery or thriller, pastel chromas and centered scripts for romance), with measurable gains when such design features are modeled alongside word meaning (Haraguchi, Iwana, and Uchida 2024). Parallel work on back-cover discourse and blurbs constructs large corpora of promotional paratexts and demonstrates that these brief, highly codified texts support reliable genre classification, precisely because they distill market-facing expectations into portable lexical and rhetorical patterns (Aly, 2018). Recent library-science research consolidates these strands, fusing title and covers analysis to deliver markedly higher balanced accuracy on public datasets, while explicitly confronting hybrid genres and cataloging pragmatics (Yang and Zhang 2025).

From a literary-theoretical perspective, these results do not collapse genre into mere marketing; rather, they show that paratexts participate in the genre contract as socially negotiated cues. Genette's reflection on paratexts as liminal, illocutionary frames helps explain why AI excels on them: models learn the institutionalized markers that publishers deploy to guide readers – typographic atmospheres, emblematic icons, and formulaic blurbs – markers that belong as much to the market and editorial world as to the texts (Genette 1997). Yet this centrality of paratexts raises conceptual questions. If genre is traditionally thought as a principle of textual organization – linked to narrative forms, stylistic conventions, and rhetorical expectations – then the reliance on paratextual prediction risks shifting the focus from immanent structures to external signals. In other words, what is being captured is less the genre of the work than the genre of its positioning. The generic characterization of paratextual elements not only demonstrates the marketable nature of the notion of genre but also dialogues with the idea of literary genre as a social construct, which, as we have seen above, has been a recurring theme in second half of the twentieth century theory.

3.5 Narratology

Narratology has long provided literary theory with conceptual tools to analyze genre through structures of story and discourse, from Genette's categories of focalization to Propp's functions and Bakhtin's dialogism. These categories can formalize, making it possible to test them across large corpora. Recent work shows how AI can detect narrative point of view and shifts in focalization, continuing early experiments such as Wiebe (1994) algorithm for assigning perspective at the sentence level, and extending them with the capacities of LLMs (Hicke et al., 2024; Piper and Bagga, 2024). Similar advances have been made in quotation attribution and the mapping of dialogic relations: by automating the detection of speech boundaries and speakers, AI enables us to quantify dialogue density and heteroglossia as possible genre markers (Muzny, Algee-Hewitt and Jurafsky, 2017). Temporal organization, too, has become accessible to algorithmic modeling: systems can trace event sequencing, flashbacks, and tense-aspect patterns, thus making visible the different chronologies that characterize genres (Piper, So, and Bamman, 2021). Another development is causal modeling: by extracting chains of cause and effect from plots, AI allows us to distinguish genres in terms of their logic of events, contrasting the deductive chains typical of crime fiction with the looser causalities of experimental narratives (Mostafazadeh et al., 2016).

Alongside these conceptual applications, new benchmarks make narratological analysis operational. The AIStorySimilarity project, for example, employs LLMs to extract elements such as characters, plot, setting, and themes, and then computes semantic distances between stories, thus clustering narratives according to structural affinities rather than

predefined categories (Chun 2024). These developments suggest that when narratology is operationalized, genre emerges less as a static taxonomy than as a constellation of features – voice, temporality, causality, character roles – that can be identified, compared, and clustered at scale. This opens the way to redefining literary genre as a dynamic and evolving set of patterns, not only traceable across vast corpora but also visible in the interactions between human and machine storytelling.

3.6 Language, rhetoric, style

In literary studies, linguistic, rhetorical, and stylistic analysis has always been central to the classification of genres, but the application of AI makes it possible to operationalize and scale these features in new ways. One productive line of research lies in rhetorical structure and discourse modes: through rhetorical structure theory (RST) and discourse parsers, AI models can automatically detect coherence relations such as narration, contrast, or explanation, revealing how different genres privilege particular rhetorical strategies (Taboada and Mann, 2006). Closely related is the analysis of figurative language, where large datasets of metaphors and similes are automatically extracted to demonstrate how certain genres are characterized by distinct figurative densities (Shutova 2010). Prosodic features have also been reintroduced into genre theory thanks to automatic scansion and rhyme detection, which allow us to map verse structures and hybrid poetic forms at scale (Agirrezabal et al., 2016). Beyond poetry, AI has been applied to epistemic stance and modality, quantifying hedges, assertions, and markers of unreliability that genres mobilize in distinctive ways – for instance, the factual stance of historiography versus the modalized uncertainty of speculative fiction (Pyatkin et al., 2021). Information-theoretic measures such as surprisal and entropy provide another powerful tool: by modeling predictability and complexity across texts, researchers can show how genres differ in pacing and informational density (Linzen and Jaeger, 2016; Degaetano-Ortlieb and Teich, 2017). Even punctuation and paragraphs rhythms have been reframed as stylistic signals: computational stylistics has demonstrated that the distribution of commas, dashes, and paragraph breaks constitutes a low-level stylistic fingerprint of genre (Darmon, 2021; Mendhakar, 2022). Similarly, profiling methods such as Profiling-UD exploit grammatical and morpho-syntactic bundles to discriminate between registers and genres across languages (Brunato et al. 2020).

Together, these strands suggest that linguistic, rhetorical, and stylistic analysis powered by AI enables a rethinking of genre as a complex bundle of formal features. If narratological models tend to describe genre through story structure, these approaches demonstrate that genre also emerges through language itself, its rhythms, patterns, and stylistic signatures. In this sense, AI does not merely automate classical stylistics but helps redefine genre as a multi-layered construct, rooted as much in linguistic texture as in narrative form.

4. From application to theory

The framework outlined in the previous section reveals a wide range of approaches to literary genre within AI studies. Beyond the individual results of applied studies, the key question is how this body of research can contribute to redefining the notion of literary genre and orienting contemporary genre theory.

4.1 Empirical and pragmatic

What clearly emerges is that the use of AI in relation to the concept of literary genre decisively accelerates what Anceschi has described as a “reclaiming of the empirical” in contemporary genre theory (Anceschi, 1997, p. 71). AI contributes to this regaining through the predominant use in current systems of a bottom-up, descriptive approach that starts from the objects of study in order to highlight patterns. Due to its very nature, machine learning constitutes a methodology of empirical analysis. Yet the empirical component is expressed not only in the chosen technique but also in the characteristics of the texts selected for analysis, as we have seen in the previous section. Studies reveal a renewed attention to formal, narratological, and linguistic aspects, which prove to be extremely multifaceted.

Alongside the analysis of formal aspects – largely shared with classical genre theory – computational analysis also focuses on categories more closely tied to the pragmatic dimension, that is, to the relationship established between texts, literary genres, and the surrounding world, represented by readers and society⁷. This includes attention to the emotional dimension of literary genre, but also the social relations expressed through character networks and the market-oriented paratextual dimension. We can therefore say that AI contributes to the reclaiming of the empirical both in its textual and in its pragmatic dimension, revealing a convergence with the formalist and structuralist tradition as well as with more recent approaches from cultural studies⁸. For instance, the use of sentiment analysis for defining genre falls within the affective turn in literary studies, while the attention to cultural diversity recalls a long tradition ranging from postcolonial studies to transcultural studies.

A few more words should be added on paratexts – an aspect usually not considered in genre theory, and a perfect example of AI’s ability to expand the boundaries of the notion toward the pragmatic dimension. AI-based paratextual analysis pushes literary theory to confront a dual horizon of genre: on the one hand, the internal horizon of textual form; on the other, the external horizon of editorial framing. What emerges is not merely a technical tool for classification but an epistemological challenge: genre becomes visible not only as a literary category but also as an institutional performance enacted at the book’s threshold. This obliges us to acknowledge that genre contracts are never purely formal, but always negotiated across multiple layers of discourse – from the narratological to the commercial, from the stylistic to the paratextual. A constructive response is to treat paratext-based classification as modeling a paratextual genre horizon – the ensemble of cues that circulate around a book at its entry into the marketplace – rather than as a verdict on intrinsic form. In practice, scholarship and curation should articulate the relation between this horizon and textual evidence: use paratext-trained systems to surface the industry’s expectations (how a book is positioned), then read against or with those expectations in the work’s rhetoric and narrative organization. That dual register preserves genre’s status as both an immanent pattern of discourse and an institutional pact inscribed, conspicuously, on the book’s skin.

4.2 The heterogeneous and totalizing nature of genre

⁷ The distinction is well rooted in literary theory and philosophy, beginning with Kant (1798), who distinguished between empirical knowledge – arising from observation and experience – and pragmatic reason – concerned with action and consequences. In the literary field, particularly relevant is Jauss (1970), who emphasizes the pragmatic dimension of literature – its reception, function, and use in society – as distinct from purely empirical textual analysis.

⁸ The term “empirical” also recalls empirical literary studies, that focus on reader response and are also connected with computational methodologies (Kuiken and Jacobs, 2021).

One conclusion that emerges from many of the studies cited is the composite nature of the notion of genre. Taken individually, none of the elements seem strong enough to define a genre on its own. About the emotional component, for instance, Kim et al. admit that “the prototypical emotion development of a given genre appears to be a sufficient, but not necessary condition for membership in a genre” (Kim, Pado, Klinger, 2017, p. 25). The same point also emerges – indeed even more strongly – in the case of character networks and paratexts. We may therefore say that what arises from AI-based analysis is a notion of genre as an entity composed of heterogeneous elements⁹. This makes the combinatory possibilities potentially infinite. Crystallization is provided by tradition, which acts as a catalyst for these elements within an otherwise limitless range of possible choices. It is here, precisely, that the conventional dimension of genre comes into play. As mentioned in paragraph 2, the formalization of genres cannot be entirely disregarded, because without it the very concept of genre could not exist, and texts would dissolve into random combinations impossible to cluster. Within a contemporary notion of literary genre that is, and must remain, descriptive, the rule-based component of genre thus finds its specific role without lapsing into normativity.

Another consequence of the convergence of heterogeneous elements into the definition of genre is a totalizing dimension of genre in relation to other concepts of literary theory. By encompassing other theoretical concepts, genre functions as a kind of meta-concept that is at once a principle of generation and of ordering. In this regard, AI analysis seems to confirm some intuition from literary scholars, that envisioned genre as a collective framework for literary study. For instance, Guillén regards genre as an organizational concept that encompasses and orders other categories (historical generations, textual structures, sociocultural attitudes), an idea that can relate to the capacity of machine learning to categorize phenomena. Guillén defines genres as “a conceptual model that guides literary creation and interpretation” (Guillén, 1985, p. 146) and a “synthesizing-concept, in which the predominant features of a plurality of works, authorities, and canons are brought together and reconciled” (Guillén, 1985, p. 149). In these words, we can glimpse – though not explicitly stated – the idea of genre as a totalizing concept within literary theory, a notion that also appears in the studies we have examined on the application of AI to genres.

4.3 Old chestnuts and new directions

AI applications to textual analysis confirm some well-established insights of contemporary genre theory, such as the historical evolution of genres. Underwood (2016) addresses the question of the duration and coherence of literary genres by employing quantitative methods and distant reading. He argues that not all genres share the same lifespan: some prove to be long-lasting and coherent, while others are more unstable and fragmented. Machine learning techniques make it possible to empirically measure the strength and persistence of textual similarities. By analyzing lexical and stylistic transformations across large corpora, one can observe how the formal and pragmatic features of a genre shift historically: for example, the growing emphasis on readability in twentieth-century fiction indicates a tendency toward stylistic simplification (Feldkamp, 2023). Pretrained language models segmented by historical periods can reveal lexical, morphological, and syntactic changes that are distinctive of specific cultural epochs (Fittschen et al., 2025). In this sense,

⁹ This recalls Wittgenstein’s (1953) concept of “family resemblance,” often applied to literary genres to highlight that they lack rigid boundaries or a fixed set of necessary features, and are instead recognized through partial and variable similarities (Fishelov, 1991; Henny-Krahmer, 2024).

temporality becomes not only context but also method: genre emerges as a fluid aggregate, a dynamic distribution of traits that continually redefines itself in relation to cultural history. The use of diachronic embeddings and semantic divergence metrics makes it possible to map shifts in meanings, metaphors, and tropes that delineate the contours of genres (Kutuzov et al., 2018).

In addition to confirming well-established tendencies in genre theory, AI can reveal more recent developments of genre. The empirical turn fostered by AI expands genre theory beyond the strictly literary field, pushing it toward a genuinely cross-media horizon where texts, images, sounds, and platforms converge in shaping generic expectations. This is coherent with the contemporary transmedia landscape, described by Jenkins as convergence culture. Media's modal affordances actively participate in genre formation: social-media 'sharing' practices, for instance, evolve into genres in their own right, shaped by the capacities of specific media forms (Gruber, 2019). Yet this enlargement also exposes a paradox, that is linked to the double nature of AI as both a creative and an analytical tool: empirical pattern-detection, presented as descriptive and open-ended, may quietly reassert normativity, since algorithms that cluster works into genres simultaneously influence production and circulation by establishing the patterns to which future works and readers are expected to conform. This feedback loop becomes evident in the field of reception, where the empirical traces of reader behavior – reviews, recommendation systems, or social media classifications – are themselves absorbed into algorithmic taxonomies, reinforcing some genre's boundaries while marginalizing others. Indeed, recent computational studies of AI-generated writing emphasize that reader judgments vary significantly depending on their profile: surface-focused readers prioritize clarity and richness, while holistic readers value thematic and rhetorical complexity (Marco, Gonzalo and Fresno, 2025). In this scenario, genre ceases to be a stable category grounded in tradition or form and increasingly functions as a dynamic site of negotiation between cultural production, computational modeling, and reader practices, with AI operating both as a descriptive instrument and as a prescriptive agent within the cultural ecosystem.

5. Conclusion

The comparison between studies that employ AI for genre analysis and the theories of genre developed during the twentieth-century shows that AI confirms the main tendencies of literary theory, particularly the empirical character of genre, its social dimension, its heterogeneous and mediated nature, and its prominent role within literary studies. At the same time, the emergence of specific tendencies – such as the focus on emotional patterns, cultural diversity, and paratexts – indicates that AI is also capable of opening new pathways for the future development of genre theory. It also emerges the usefulness of integrating applied studies from digital humanities within the broader framework of literary theory. Although this operation can at times prove difficult due to differences in approach and terminology, it is in fact highly productive, as it allows literary theory to engage with a wide body of empirical research that not only confirms already consolidated notions but also provides new and valuable contributions.

In conclusion, it is necessary to reflect on the inclusion of AI in the study of literary genres. The contribution of AI to genre theory does not lie solely in the provision of new empirical tools, but in the reconfiguration of what counts as empirical evidence in the first place. Machine learning systems do not simply extend the scholar's capacity to detect patterns: they introduce a different epistemology, attentive to correlations and micro-traces that escape human critical gaze. As a result, genre emerges less as a humanly constructed

taxonomy than as a hybrid object negotiated across two forms of cognition, human and machinic. The empirical is thus redeemed not by being restored to its traditional role of raw data for interpretation, but by becoming an arena of dialogue where human hermeneutics and algorithmic pattern-recognition meet, clash, and sometimes converge. Genre theory, in the age of AI, is called to inhabit this liminal space: to recognize that the categories it works can be shaped by the interplay of interpretive intentionality and machine-driven discovery, and that this hybrid empiricism challenges the very distinction between creation, analysis, and classification.

Bibliography

- Agirrezabal M., Arrieta B., Astigarraga A., Hulden M. (2016), *ZeuScansion: A tool for scansion of English poetry*, in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, European Language Resources Association, pp. 1334-1338.
- Aly R. (2018), *Hierarchical Writing Genre Classification with Neural Networks*, bachelor's thesis, Universität Hamburg.
- Anceschi L. (1997), *Progetto di una sistematica dell'arte*, Milano, Mursia, 1997.
- Arielli E. (2024), *Human Perception and The Artificial Gaze*, in Manovich L., Arielli E., *Artificial Aesthetics*, pp. 95-118. <https://manovich.net/index.php/projects/artificial-aesthetics>.
- Bajohr, H. (2024), *Writing at a distance: Notes on authorship and artificial intelligence*, "German Studies Review", 47, 2, pp. 315-337.
- Brunato D., Cimino A., Dell'Orletta F., Venturi G., Montemagni S. (2020), *Profiling-UD: a Tool for Linguistic Profiling of Texts*. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, Marseille, European Language Resources Association, pp. 7145-7151.
- Cabezas-Clavijo Á., Magadán-Díaz M., Rivas-García J., Sidorenko-Bautista P. (2024), *This Book is Written by ChatGPT: A Quantitative Analysis of ChatGPT Authorships Through Amazon.com*, "Pub Res" 40, pp. 147-163.
- Chun J. (2024), *AIStorySimilarity: Quantifying story similarity using narrative for search, IP infringement, and guided creativity*. In *Proceedings of the 28th Conference on Computational Natural Language Learning*, Association for Computational Linguistics, pp. 161-177.
- Darmon A., Bazzi M., Howison S. D., Porter M. A. (2021), *Pull out all the stops: Textual analysis via punctuation sequences*, "European Journal of Applied Mathematics", 32, 6, pp. 1069-1105.
- Degaetano-Ortlieb S., Teich E. (2017), *Modeling intra-textual variation with entropy and surprisal: topical vs. stylistic patterns*. In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Vancouver, Association for Computational Linguistics, pp. 68-77.
- Elson, D. K., Dames N., McKeown K. (2010), *Extracting Social Networks from Literary Fiction*, In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, ACL, pp. 138-147.
- Farina M., Pedrycz W., Lavazza A. (2024), *Towards a mixed human-machine creativity*, "Journal of Cultural Cognitive Science", 8, pp. 151-165.
- Farrell H., Gopnik A., Shalizi C., Evans J. (2025), *Large AI models are cultural and social technologies*, "Science", 387, pp. 1153-1156.

- Feldkamp P., Bizzoni Y., Lassen I. M., Rosendahl Thomsen M., Nielbo K. (2023), *Readability and Complexity: Diachronic Evolution of Literary Language Across 9000 Novels*, In *Proceedings of the Joint 3rd International Conference on Natural Language Processing for Digital Humanities and 8th International Workshop on Computational Linguistics for Uralic Languages*, Tokyo, Association for Computational Linguistics, pp. 235-247.
- David Fishelov (1991), *Genre theory and family resemblance—revisited*, “Poetics”, 20, 2, pp. 123-138.
- Fittschen E., Li S., Lippincott T., Choshsem L., Messner C. (2025). *Pretraining Language Models for Diachronic Linguistic Change Discovery*. arXiv.2504.05523
- Frye, N. (1957), *Anatomy of Criticism: Four Essays*, Princeton, Princeton University Press.
- Galib A., Prima M., Debi S., Mahadi M., Ahmed N., Rhythm E., Amit A., Rasel A. (2023), *Genre Classification: A Machine Learning Based Comparative Study of Classical Bengali Literature*, In *26th International Conference on Computer and Information Technology*, IEEE, pp. 1-6.
- Genette G. (1997), *Paratexts: Thresholds of Interpretation*, Cambridge, Cambridge University Press.
- Guillén, C. (1985), *Entre lo uno e lo diverso*, Barcelona, Editorial Critica.
- Gruber, H. (2019), *Genres, media, and recontextualization practices. Re-considering basic concepts of genre theory in the age of social media*, “Internet Pragmatics”, 2, 1, pp. 54-82.
- Haraguchi D., Iwana B.K., Uchida S. (2024). *What Text Design Characterizes Book Genres?*, in Sfikas G., Retsinas G. (eds), *Document Analysis Systems*. Cham, Springer, https://doi.org/10.1007/978-3-031-70442-0_10
- Henny-Krahmer U. (2024), *Family Resemblance in Genre Stylistics: A Case Study with Nineteenth-Century Spanish-American Novels*, in Hesselbach R., Calvo Tello J., Henny-Krahmer U., Schöch C., Schlör D. (eds.), *Digital Stylistics in Romance Studies and Beyond*, Heidelberg, Heidelberg University Publishing, pp. 149-171.
- Hettinger L., Becker M., Reger I., Jannidis F., Hotho A. (2015), *Genre Classification on German Novels*, in *Proceedings of the 26th International Workshop on Database and Expert Systems Applications*, Valencia: IEEE, pp. 249-253.
- Hicke R., Bizzoni Y., Feldkamp P., Kristensen-McLachlan R. (2024), *Says Who? Effective Zero-Shot Annotation of Focalization*, arXiv:2409.11390.
- Iwana B., Rizvi S., Ahmed S., Dengel A., Uchida S (2016), *Judging a Book by Its Cover*, arXiv:1610.09204
- Jauss, H. R. (1982), *Toward an Aesthetic of Reception*. Minneapolis, University of Minnesota Press.
- Jenkins H. (2006), *Convergence Culture: Where Old and New Media Collide*, New York, NYU Press.
- Kant, I. (2006), *Anthropology from a Pragmatic Point of View*, Cambridge, Cambridge University Press.
- Kim E., Padó S., Klinger R. (2017), *Investigating the Relationship between Literary Genres and Emotional Plot Development*, in *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Vancouver, Association for Computational Linguistics, pp. 17-26.
- Kotzé E., Senekal B. (2024), *Afrikaans Literary Genre Recognition using Embeddings and Pre-Trained Multilingual Language Models*, in *International Conference on Artificial Intelligence, Computer, Data Sciences and Applications*, Seychelles, IEEE, pp. 439-444.

- Kuiken D., Jacobs A. (eds.) (2021), *Handbook of Empirical Literary Studies*, Berlin, De Gruyter.
- Kutuzov A., Øvrelid L., Szymanski T., Velldal E. (2018). *Diachronic word embeddings and semantic shifts: a survey*, in *Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, Association for Computational Linguistics, pp. 1384-1397.
- Kuzman T., Ljubešić N. (2025), *Automatic genre identification: a survey*, “Lang Resources & Evaluation”, 59, pp. 537-570.
- Labatut V., Bost X. (2019), *Extraction and Analysis of Fictional Character Networks: A Survey*, “ACM Computing Surveys”, 52, 5, pp. 1-40.
- Linzen T., Jaeger T. F. (2016), *Uncertainty and Expectation in Sentence Processing: Evidence from Subcategorization Distributions*, “Cognitive Science”, 40, 6, pp. 1382-1411.
- Lucieri, A., Sabir H., Siddiqui S., Rizvi S., Iwana B., Uchida S., Dengel A., Ahmed S. (2020), *Benchmarking Deep Learning Models for Classification of Book Covers*, “SN Computer Science”, 1, pp. 1-16.
- Marco G., Gonzalo J., Fresno V. (2025), *The Reader is the Metric: How Textual Features and Reader Profiles Explain Conflicting Evaluations of AI Creative Writing*, in *Findings of the Association for Computational Linguistics*, Vienna, Austria. Association for Computational Linguistics, pp. 25432-25449.
- McCorduck P. (2004), *Machines who think*, A.K. Peters, New York.
- McDermid J. A., Jia Y., Porter Z., Habli I. (2021), *Artificial intelligence explainability: the technical and ethical dimensions*, “Philosophical Transactions of the Royal Society A”, 379 (2207), 20200363.
- Mendhakar A. (2022), *Linguistic Profiling of Text Genres: An Exploration of Fictional vs. Non-Fictional Texts*, “Information”, 13, 8, pp. 1-17.
- Mengliev D. B., Barakhnin V. B., Saidov B. R., Atakhanov M., Eshkulov M. O., Ibragimov B. B. (2024), *A Computational Approach to Recognizing Poetry Genres in Uzbek Texts*, in *2024 IEEE International Multi-Conference on Engineering, Computer and Information Sciences*, Novosibirsk, IEEE, 2024, pp. 319-322.
- Miller, C. R. (2015), *Genre as social action (1984), revisited 30 years later (2014)*, “Letras & Letras”, 31, 3, pp. 56-72.
- Moretti, F. (2011), *Network Theory, Plot Analysis*, Stanford, Stanford Literary Lab.
- Mostafazadeh N., Chambers N., He X., Parikh D., Batra D., Vanderwende L., Kohli P., Allen J. (2016), *A corpus and cloze evaluation for deeper understanding of commonsense stories*, in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, Association for Computational Linguistics, pp. 839-849.
- Muzny G., Algee-Hewitt M., Jurafsky D. (2017), *Dialogism in the novel: A computational model of the dialogic nature of narration and quotations*, “Digital Scholarship in the Humanities”, 32, 2, pp. 31–52.
- Pérez y Pérez R., Sharples M. (2024), *Three computer-based models of storytelling: BRUTUS, MINSTREL and MEXICA*, “Knowledge-Based Systems”, 17,1, pp. 15-29.
- Piper A., So R., Bamman D. (2021), *Narrative Theory for Computational Narrative Understanding*, in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Punta Cana, Association for Computational Linguistics, pp. 298-311.
- Piper A., Xu M., Ruths D. (2024a), *The Social Lives of Literary Characters: Combining citizen science and language models to understand narrative social networks*, in

- Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities*, Miami, Association for Computational Linguistics, pp. 472-482.
- Piper A., Bagga S. (2024), *Using Large Language Models for Understanding Narrative Discourse*, in *Proceedings of the 6th Workshop on Narrative Understanding*, Miami, Association for Computational Linguistics, pp. 37-46.
- Pyatkin V., Sadde S., Rubinstein A., Portner P., Tsarfaty R. (2021), *The Possible, the Plausible, and the Desirable: Event-Based Modality Detection for Language Processing*, in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, Association for Computational Linguistics, pp. 953-965.
- Rahul A., Agarwal D., Vijay D. (2021), *Genre Classification using Character Networks*, in *2021 5th International Conference on Intelligent Computing and Control Systems*, Madurai, IEEE, pp. 216-222.
- Rasheed A., Umar A.I., Shirazi S.H., Khan Z., Shahzad M. (2023), *Cover-based multiple book genre recognition using an improved multimodal network*, "International Journal on Document Analysis and Recognition", 26, 1, pp. 65-88.
- Samothrakis S., Fasli M. (2015), *Emotional Sentence Annotation Helps Predict Fiction Genre*, "PLOS ONE", 10, 11, pp. 1-10.
- Soares de Lima E., Casanova M. A., Feijó B., Furtado A. L. (2025), *Characterizing the investigative methods of fictional detectives with large language models*, arXiv:2505.07601
- Shutova, E. (2010), *Models of metaphor in NLP*, in *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 688-697.
- Taboada M., Mann W. C. (2006), *Rhetorical Structure Theory: Looking back and moving ahead*, "Discourse Studies", 8, 3, pp. 423-459.
- Todorov T. (1976). *The Origin of Genres*, "New Literary History", 8,1, pp. 159-170.
- Underwood T. (2017), *The Life Cycles of Genres*, "Journal of Cultural Analytics", 2, 2, pp. 1-25.
- Valls-Vargas J., Zhu J., Ontañón S. (2014), *Towards automatic identification of narrative structure in classical tragedies*, in *Interactive Storytelling*, Cham, Springer, pp. 78-89.
- Wiebe, J. (1994), *Tracking point of view in narrative*, "Computational Linguistics", 20, 2, pp. 233-287.
- Winko, S. (2023), *Literature and emotion*, in Schiewer, G., Altarriba, J., Ng B., *Language and Emotions*, vol. 3, Berlin-Boston, De Gruyter Mouton, pp. 1417-1436.
- Wittgenstein, L. (1953), *Philosophical Investigations*, Oxford, Blackwell.
- Yako M. (2021), *Emotional Content in Novels for Literary Genre Prediction and Impact of Feature Selection on Text Classification Models*, Master Thesis, University of Uppsala.
- Yang X., Zhang Z. (2025), *Enhancing Book Genre Classification with BERT and InceptionV3: A Deep Learning Approach for Libraries*, "PeerJ Computer Science", pp. 1-20.
- Zhu M., Wang G., Li C., Wang H., Zhang B. (2023), *Artificial Intelligence Classification Model for Modern Chinese Poetry in Education*, "Sustainability", 15, 6, pp. 1-19.