

Tacit Knowledge and Hermeneutics: Critical Reflections on Automated Text Interpretation with Large Language Models

Jens Dörpinghaus^{*†‡}, Michael Tiemann^{*†}

^{*} Federal Institute for Vocational Education and Training (BIBB), Bonn, Germany

[†] University of Koblenz, Germany

[‡] University of Pretoria, South Africa

Email: jens.doerpinghaus@bibb.de, tiemann@bibb.de

Abstract—This study explores the interplay between tacit knowledge and hermeneutical interpretation within the domain of large language models (LLMs). In this study, we examine the question of whether automated systems can approximate interpretive practices that depend on implicit understanding, situated experience, and contextual embedding. To this end, we draw on philosophical and psychological accounts of tacit knowledge and traditions of hermeneutics. Utilizing a comparative experimental design, multiple LLMs are prompted to interpret texts from the domains of philosophy, theology, literature, and conversation. We find considerable variability and constrained reproducibility in the interpretive outcomes. These findings suggest that the core dimensions of tacit knowledge, particularly embodied inference and context-sensitive understanding, remain inaccessible to current generative models. The paper posits that hermeneutics cannot be reduced to formal pattern recognition and calls for refined theoretical and empirical criteria for evaluating machine-based interpretation.

I. INTRODUCTION

ALL interpretation presupposes forms of knowledge that are not fully articulable. Discussions within the field of philosophy concerning understanding have repeatedly emphasized that explicit rules and formal representations are contingent upon background competencies that remain implicit. This dimension is commonly described as tacit knowledge, as evidenced by the works of Polanyi and Turner, see [1], [2]. Tacit knowledge is defined as skills, perceptions, and interpretive abilities that guide action and judgment without being fully codified. The catch that tacit knowledge is not explicable has been shown [1], [3], [4] for general and also specific contexts like knowledge in work. Tacit knowledge is still a central part of interpretation.

Hermeneutics, understood as the theory of interpretation and understanding, presupposes precisely such background competencies. Nevertheless, Tomkins and Eatough assert that the field of hermeneutics comprises two primary branches: one branch focuses on the activities of interpretation, while the other branch focuses on the philosophy of understanding [5]. Interpretation, therefore, does not operate solely through explicit procedures but rather through situated engagement with language, tradition, and experience. Gadamer's concept

of hermeneutics as “a theory of the real experience that thinking is” [6] offers a compelling framework for understanding this phenomenon. The fields of classical and contemporary hermeneutics share the emphasis on the idea that meaning emerges from within interpretive horizons that are shaped by history, practice, and community. Ricoeur asserts that this phenomenon of sense-making “means being distrustful of whatever appears to us in consciousness, whether this is explicit knowledge or more tacit intuition, for both modes are potentially manifestations of false, not true, consciousness.” [5].

In recent years, substantial language models have been introduced as systems with the capacity to comprehend and interpret text. This prompts a fundamental question, which the present study seeks to explore. Do systems that operate on statistical pattern recognition have the potential to approximate hermeneutical practices in general, and more specifically, whether such systems show use of tacit knowledge. The central research question guiding this study is as follows: To what extent can LLMs approximate hermeneutical interpretation, and do their outputs reflect forms of tacit knowledge or merely formal linguistic correlations?

The present approach is predicated on methodological reflection in combination with empirical experimentation. The objective of this investigation is to ascertain whether the reiterated prompting of LLMs results in the formation of stable interpretive structures. Furthermore, it is to be determined whether this stability could be construed as a computational analogue of tacit knowledge.

II. THEORETICAL BACKGROUND AND LITERATURE REVIEW

A. Tacit Knowledge and Hermeneutics

The concept of tacit knowledge was a central theme in the work of Michael Polanyi, who posited that “there are things that we know but cannot tell” [7]. The acquisition and application of knowledge are contingent upon a certain degree of background awareness that resists complete formalization. This insight carries profound implications for the field of

hermeneutics, as interpretation entails this dynamic interplay between articulated meaning and implicit understanding. Collins's work on tacit knowledge posits the existence of three distinct dimensions: relational, somatic, and collective [8]. Relational tacit knowledge is defined as knowledge that, in principle, could be made explicit but remains unarticulated due to constraints related to the organization and dissemination of knowledge. Somatic tacit knowledge is rooted in the physical capacities of the body and brain, while collective tacit knowledge resides at the societal level. This collective tacit knowledge involves personal experiences and shared practices, such as language, that are not yet fully explicable.

In contemporary philosophical and social scientific debates, tacit knowledge and hermeneutics are closely related, as both challenge the ideal of fully explicit, formalizable knowledge. In this context, tacit knowledge is defined as skills, intuitions, and understandings that are challenging to articulate yet are fundamental to cognition, action, and scientific practice [8]. This insight has been further developed in fields such as the sociology of knowledge and organizational studies, where tacit knowledge is regarded as context-dependent and resistant to complete codification. The philosophical tradition, originating in the work of thinkers such as Friedrich Schleiermacher and Wilhelm Dilthey, and later transformed by Martin Heidegger and Hans-Georg Gadamer, has undergone a shift in focus from method to the conditions of understanding itself. In Gadamer's philosophical hermeneutics, understanding is not a purely objective procedure but an interpretive, historically situated process shaped by prior assumptions, language, and social context [9]. The concept of meaning, therefore, is not merely extracted but rather emerges through a dialogical interaction between the interpreter and the subject matter. The field of hermeneutics underscores the notion that interpretation is inherently situated, implying that it is contingent on the context in which it is employed. The concept of meaning is derived from the dynamic interaction between the text and its interpreter. Classical theological hermeneutics, narrative analysis, and sociological reconstruction theories posit that interpretation necessitates engagement with structures that are not immediately observable in the text.

The nexus of tacit knowledge and hermeneutics is characterized by a shared rejection of the possibility of fully explicit knowledge. According to scholars in the field, hermeneutic interpretation is contingent on pre-reflective, background competencies that function as tacit knowledge. These background competencies could be found in knowing how to formulate or how to do specific tasks and they are part of cultural and personal habits. These competencies enable individuals to make sense of texts, practices, and social interactions [10], [11]. From this perspective, tacit knowledge constitutes a condition of possibility for interpretation, as it underpins the practical mastery of language and the implicit rules governing understanding. In contrast, hermeneutic theory offers a structured approach for examining the manner in which these tacit dimensions are interwoven within historical traditions and social practices. This theoretical framework facilitates

the analysis of the processes through which meaning is constructed and interpreted. For a more detailed exposition of this topic, see [12]. Empirical and theoretical work in the field of knowledge studies further substantiates this correlation by demonstrating that the generation and dissemination of knowledge, particularly within scientific and organizational settings, are contingent on interpretive processes that cannot be reduced to explicit information alone [13], [14]. A synthesis of these perspectives reveals that knowledge is inherently situated, embodied, and socially constructed, necessitating interpretive engagement rather than mere formal representation.

Measuring tacit knowledge is not easy as it is by definition inexplicated. There are scholars, especially in Psychology, who devised ways to measure tacit knowledge nonetheless. They focus on the process of learning [15], [16] or whether this kind of knowledge can be shown [17], [18]. They have in common that tacit knowledge is indirectly sought: One may learn a pattern without realising and thus complete a task increasingly faster (relying on that pattern) or reconstruct how additional activities helped increase productivity or efficiency in certain tasks. We also try to show the presence of tacit knowledge indirectly. In a very broad view one might say that without tacit knowledge, the LLMs would not be able to give an interpretation at all. But this is also harder to test. So we take a narrower approach by giving texts for interpretation of which we know that tacit knowledge is needed to even understand them correctly. One such type of text are political jokes. If the LLM cannot even determine that the text in question is a political joke, then we cannot assume that tacit knowledge had been used.

B. AI and Hermeneutics

A considerable number of researchers have posited the utilization of hermeneutics in the context of AI-generated texts [19]. Conversely, alternative approaches have advocated for the implementation of hermeneutics as a tool to enhance the efficacy of prompts [20]. Nevertheless, there is a conspicuous absence of hermeneutics for AI-generated texts, as acknowledged in extant research [21], [22], [23]. While some researchers are generally positive on the application of extant theory, e.g., that of Heidegger or Lacan, to LLMs [24], there is no consensus on this issue [25]. In general, researchers from non-technical domains appear to exaggerate the features and capabilities of LLMs and AI in general. For a detailed discussion of this phenomenon, see [26]. As other researchers have noted, the importance of domain knowledge in various fields is paramount. One notable example is the use of large language models (LLMs) in confessional religious education, particularly in the context of Bible teaching [27]. However, given that the central objective of our research is to ascertain the capacity of LLMs to execute hermeneutic tasks, we will henceforth direct our attention to extant literature on the subject.

According to the extant literature, there is only one preprint paper that addresses this question directly [28], which is critical of the hermeneutic and understanding capacities of

LLMs. Alternate research endeavors concentrate on technical limitations or the appropriate application of LLMs. One example is the conceptualization of LLMs as “AI-interns” as proposed by [29]. Another paper [30] explores the theological and ethical inclinations of LLMs in the interpretation of texts from the Hebrew Bible. The study identifies a bias that aligns with contemporary socio-political and environmental concerns. Another journal article proposes that religious questions should serve as a benchmark for LLMs; however, it does not provide any technical considerations and is limited to the theory of the four senses of Scripture [31]. A number of scholars have chosen to focus their research on hermeneutics within the medical domain. For instance, Hans-Georg Gadamer [32] posits that Large Language Models (LLMs) are incapable of assuming the role of patient interaction, as they lack the capacity to incorporate moral practices. It is generally acknowledged that the human element of interaction and experience represents a significant challenge in the application of LLMs for hermeneutical purposes. Despite the proposition that machine relationships can complement human-human interactions, as posited by [33], this assertion fails to address the underlying hermeneutical inquiries.

Tacit knowledge is not merely personal intuition; rather, it is embedded in practices and communities. The manifestation of this concept is evident through the execution of actions, the application of sound judgment, and the adept execution of tasks. In hermeneutical traditions, analogous assumptions emerge when interpreters rely on tradition, pre-understanding, and contextual familiarity. However, machine learning systems function in a distinct manner. These models approximate patterns from extensive corpora, yet they lack experiential knowledge and embodied participation. A significant body of research has been dedicated to exploring the capacity of LLMs to effectively model or mimic persona [34]. Furthermore, it was determined that persona variables account for less than 10% of the variance in LLMs.

Therefore, while statistical generalization can mimic aspects of interpretive behavior, it remains unclear whether such processes constitute tacit knowledge or merely simulate its effects. The tension between tacit knowledge and formal computation becomes central when evaluating LLM-based interpretation.

C. Objective Hermeneutics

Objective Hermeneutics, sometimes also labeled structural hermeneutics, derives from cultural anthropology. [35] discusses its connections to theories by Berger and Luckmann, Mead, and Chomsky. The idea is that how entities experience the world they live in is reflected in their speech. On an individual level this would relate to an expression of one person’s perception of reality, its ordering and the connections between entities. While these expressions are singular to one person in almost all cases they are in fact structures expressing themselves through these persons. Sociological theory also distinguishes broader levels of groups of persons or organizations (meso level) up to the level of states (macro level). It is important to note that entities on these levels (like firms

or unions on the meso level or governments on the macro level) also are such expressions of structures. In each case a structure expresses its respective views through text. This can be any kind of textualised or textualisable expression, from pharaoh’s hieroglyphs to frescos to speech-messages or poems. In this sense, every expression becomes text and every text carries meaning in the senso of information on underlying assumptions and perceptions. These assumptions and perceptions are, according to objective hermeneutics, structures or carriers of meaning. Objective hermeneutics is therefore a knowledge hermeneutics.

It poses very interesting features that make it seem to lend itself to automated analysis. One example is how meaning is reconstructed through comparison of textual expression with other possible expressions. In a sentence a person utters the structure is expressed through choices of words, particular changes or deviations from grammatical rules or the like. In a way, what LLMs do is the reverse: Based on statistical knowledge about which token will be most likely to follow in specific contexts it constructs texts. The statistical knowledge needed for this is comparable to knowledge about normality layers (see below). It seems conceivable that by showing which out of all possible alternative expressions (i.e. expressions with different probabilities) is chosen a structure can be made visible. The problem is that this requires huge computing and memory capacities. A human interpreter would also rely on these two kinds of knowledge for a reconstruction, albeit she would use the knowledge about alternative expressions to exclude those that do not fit the structure: What is sought is the one alternative which fits all expressions of a structure. This is a genuinely creative and implicit process, which is categorised as “abductive reasoning” [36, 145]. Should a LLM be able to reconstruct meaning this would not only indicate tacit knowledge but also indicate actual generative capabilities.

The structures that can be found by applying objective hermeneutics can be communicated intersubjectively and are answers to objective problems of action [35, p. 12]. For every structure there are numerous situations that pose problems, or crises. In fact, every single action is at first a situation of crises. The structure devises some solution and thus tests it by application. If this solution works repeatedly, it may become a new structure. Structures can also change and be adapted to different situations. Even if only for singular entities (which can comprise of many individuals), these structures are seen as rules and procedures of an algorithmic nature. Individual socialization is thus seen as a continuous process of handling crises [37].

The way to re-construct structures within objective hermeneutics is thus interesting in two ways in view of its application to LLMs. The first interesting aspect is a “layer of normality” (“Normalitätsfolie”, [38, p. 136]). This is one of the starting points for reconstructing a structure and refers to ideal types of normal behavior. Statistically speaking these are means, as in the share of people with a certain qualification degree, or the share of people watching a specific film. Since

every action at one point called for a decision of the structures carrier ('Should I train as a plumber or try to get a degree?', 'Should I go watch this movie?') we can already learn something about a structure if we compare it to this layer of normality. Since LLMs are trained on data including statistical information on qualifications, for example, we could expect them to be able to construct a typical structure. The differences to this would then be a starting point for reconstruction.

The second interesting aspect refers to the logic of cognition. In objective hermeneutics researchers explicitly use abductive reasoning. This can be seen as an inference to the best explanation while at the same time only having assumptions about the laws and rules from which to infer. This is seen as a creative process and might as such be particularly challenging for LLMs.

The actual analysis is carried out on texts which are sequentially interpreted. Starting with the first token (which can be any kind of expression from taking a deep breath to saying "Okay" as in "Okay let's start", sequences are rebuild. This reconstruction only closes a sequence if there is a preliminary structure that could explain the expression of this sequence. This preliminary structure is then tested against all following sequences. That way, at first, a vast number of possible structures is being kept and tested, which will successively be combed out until the one structure remains which can explain all sequences. This is also one reason for this hermeneutic being labeled 'objective': Every researcher should, for any given structure, reconstruct this exact structure, no matter which expressions (sequences) are being evaluated.

Sequences relate to critical situations. The way a person introduces herself, starts or ends a conversation, chooses a meal from a menu — all of these are situations of crises where action is required. Since some actions have worked well for this person, their structure will routinely act on these. This also means that sequences, as expressions of structures, can refer to places, people, interactions, and generally any kind of situation a (social) being can find himself in.

For the inclusion of objective hermeneutics in this experiment this leads to some assumptions: If the LLMs are capable of applying an objective hermeneutics approach, they should all come to the same conclusion. Even though generally structures are reconstructed, these all refer to specific situations that find expressions in text. Thus, characters, places, times, etc. of these situations can also be reconstructed and should be comparable between different analysts.

D. Tacit Knowledge and Hermeneutical Understanding

The concept of tacit knowledge, most prominently articulated by [7], serves as a foundational starting point for the evaluation of interpretive capacities in artificial systems. Polanyi's well-known formulation that "we know more than we can tell" expresses a structural asymmetry between explicit articulation and practical knowing. In this sense, knowledge is not limited to propositional content. This concept encompasses subsidiary awareness, embodied orientation, and skilled performance. Tacit knowledge functions as a background condition for any

explicit statement. As previously mentioned, Polanyi differentiates between focal and subsidiary awareness. In the process of understanding, individuals focus intently on a particular object while relying on a network of implicit cues, habits, and contextual sensitivities. These subsidiary elements are not themselves thematized; however, they are indispensable for coherent perception and judgment. Tacit knowledge, therefore, is not merely an accidental residue of incomplete articulation; rather, it serves as the enabling ground of explicit meaning. Attempts to fully formalize such knowledge risk destroying the integrative structure that makes understanding possible in the first place.

Hermeneutics shares this structural insight. Interpretive traditions in philosophy, theology, and the social sciences posit the notion that meaning cannot be derived from explicit textual features alone. The concept of understanding is predicated on a pre-reflective horizon that is shaped by historical situatedness, linguistic competence, and lived experience. Interpretation unfolds within what has often been described as a hermeneutical circle, where parts and wholes mutually inform one another. This circularity is not a logical flaw but rather an epistemic structure grounded in tacit orientation. The relationship between tacit knowledge and hermeneutics becomes particularly evident when considering the role of judgment. Interpretation entails the selection of pertinent aspects, the differentiation of significant from insignificant elements, and the recognition of patterns that cannot be reduced to mechanical rule application. Polanyi's (1958) seminal work posits that the very essence of scientific discovery is contingent upon individual engagement and the implicit integration of diverse competencies. If this assertion is valid for the natural sciences, it is all the more applicable to hermeneutical disciplines, wherein meaning, intention, and context are of paramount importance.

Tacit knowledge is also socially embedded. This discipline is characterized by a multifaceted development, encompassing apprenticeship, tradition, and communal practice. The development of interpretive competence in disciplines such as theology, literature, or sociology is often facilitated by engagement in interpretive communities. Such competencies cannot be transmitted solely through explicit rulebooks; they emerge through exposure, imitation, and correction within a shared horizon of meaning. In this sense, hermeneutical expertise can be considered analogous to craft knowledge rather than algorithmic procedure. Large language models function according to a fundamentally different logic. These models are trained on extensive corpora of text and approximate probabilistic distributions of linguistic sequences. While they are capable of reproducing stylistic and thematic regularities, their architecture does not provide a mechanism for subsidiary awareness in Polanyi's sense. The model does not attend from implicit integrations toward focal objects. Instead, it generates outputs through statistical weighting across learned parameters.

It could be argued that sufficiently complex statistical systems have the capacity to approximate tacit integration. However, Polanyi's account places significant emphasis on the

concepts of embodiment and commitment. Tacit knowing is characterized by its association with risk, responsibility, and personal participation. This concept presupposes the existence of an individual who is endowed with knowledge and inhabits a structured framework, relying on it while striving to attain coherent meaning. Conversely, artificial systems are devoid of such existential embedding. The subjects' "knowledge" is not rooted in experiential engagement; rather, it is based on numerical parameter space. This distinction has direct consequences for hermeneutical evaluation. If interpretation necessitates tacit integration of context, tradition, and embodied judgment, then models that operate exclusively on textual correlation encounter structural limitations. Despite the resemblance of outputs to interpretive acts, their generative basis differs fundamentally from the integrative processes described by Polanyi.

Concomitantly, it is imperative to circumvent a simplistic dichotomy between human profundity and machine superficiality. The notion of tacit knowledge does not inherently entail mystical ineffability. The phenomenon operates through patterns that may, at least partially, manifest in behavior. The fundamental question is thus both empirical and philosophical in nature: The question of whether LLM outputs demonstrate stable integrative behavior that can functionally approximate tacit inference, or whether they merely reflect fluctuating statistical tendencies, remains a subject of active research. The experimental design of this study is situated within the theoretical framework outlined above.

Consequently, the evaluation of LLMs transcends the realm of mere technical benchmarks, delving into a philosophical inquiry into the very essence of understanding itself. In the event that tacit knowledge is considered to comprise the hidden infrastructure of hermeneutics, then any assertion that machines "interpret" must necessarily address the question of whether such infrastructure can be simulated, approximated, or fundamentally replaced by statistical computation. The subsequent discussion will entail an examination of the experimental setup and the potential tests that could be conducted.

E. Remark on anthropomorphism

[39] proves remarkably well and concise how expecting and seeing human-like features in LLMs both helps to further hype them and thus exaggerate their actual capabilities. It could be argued that trying to find tacit knowledge in LLMs reactions to interpretative tasks extends this anthropomorphic fallacy. In fact it should be noted that the way we discuss "'answers'" or "'interpretations'" by LLMs can be seen as exactly this type of fallacy. We do, though, intend to look deeper and try to show if there is any reason for assuming that LLMs have human-like capabilities with regard to interpretation. Our text would be the same but only less legible if we quoted and unquoted accordingly. But the underlying idea of "'applying'" tacit knowledge actually only works under the assumption of a knowledge carrying or even generating consciousness doing the interpretation. Which cannot, in any sense be said about LLMs – consciousness, application, generation are concepts

only applicable to biological intelligence and its mimicry by statistical models is not yet accomplished. What we could show in answering our research question thus could be that LLMs are capable of tacit knowledge and thus maybe give reason for anthropomorphising them, if and only if we did find hints for tacit knowledge in their interpretations.

III. METHODOLOGICAL APPROACH, EXPERIMENTAL SETUP AND TEXTS

A. Overview

In order to investigate the relationship between tacit knowledge and machine interpretation, we have analyzed hermeneutical tasks that require structural reconstruction, segmentation, and identification of interpretive actors. Hermeneutical approaches, including narrative analysis and structural reconstruction, operate under the assumption that texts harbor implicit structures that must be inferred. The objective reconstruction of meaning frequently relies on abductive reasoning and sequential interpretation. These processes bear a resemblance to inference mechanisms; however, they are grounded in interpretive expertise.

In our experimental setting, we operationalize hermeneutical tasks through two measurable indicators around the detection of conjectures or readings of a text, e.g., what kind of category it is. Is it a joke? In the following subsections, the rationale behind the necessity of implicit knowledge for specific literary works will be examined. The subsequent inquiry pertains to the quantity of conjectures that can be formulated for a given text. It is hypothesized that if Large Language Models (LLMs) approximate hermeneutical reasoning and tacit knowledge, repeated runs with identical prompts should yield relatively stable structures in a model and among different models. The presence of high variability would suggest a tendency to be influenced by prompt formulation rather than by the strength of interpretive grounding.

It is further hypothesized that models trained on data distributions with similar characteristics should produce comparable interpretations under identical conditions. The experimental design involves the repeated solicitation of interpretive tasks from multiple Large Language Models (LLMs) on four texts selected from diverse genres. The texts encompass a variety of genres, including poetry, theology, philosophy, and conversation.

B. Political jokes

Tacit knowledge is widely recognized as a fundamental prerequisite for understanding puns and jokes [40]. Humor frequently depends on shared background knowledge, implicit assumptions, and culturally specific references that are not explicitly stated but must be inferred by the audience.

As has been previously indicated, "jokes may be ways of telling stories about people or events or more abstract states of affairs (for example, stereotypical national or cultural or gender dispositions) without speaking directly upon the matter, for reasons of either avoiding, or to the contrary bringing, insult" [41], [42]. In this sense, jokes can be considered as

indirect communicative acts, with meanings encoded beneath their surface structure.

The analysis of political jokes reveals a particular challenge concerning the processes of identity construction and the activation of tacit knowledge. Referencing [43], the intertextual nature of political jokes is a salient factor in this analysis. The interpretation of such discourse is influenced by a multifaceted array of factors, including but not limited to linguistic competence, familiarity with political contexts, shared experiences, and implicit norms that are characteristic of the given society. In particular, within the confines of restrictive or authoritarian environments, such as the former GDR, political jokes function as a form of subtle critique, resistance, or social cohesion. However, these jokes remain opaque to individuals lacking the requisite contextual understanding.

C. GDR texts

A political text's implicit meanings are often accessible only to individuals with specific forms of tacit knowledge. In the context of the former German Democratic Republic (GDR), this phenomenon has been the subject of extensive research [44], [45]. As posited by [46] and [47], discourses on language are inextricably linked with political structures and communicative practices. It is found in political statements but also lyrics and literature [48].

As demonstrated in the *Gesetzblatt der DDR*, an illustrative example is available for reference. As stated in the Legal Gazette of the GDR (*Gesetzblatt der DDR*) on May 26, 1962:

“Lehrberufe, die auf Grund ihrer wirtschaftlichen Bedeutung und des geforderten hohen theoretischen Niveaus besondere Anforderungen an die Lehrlinge stellen, sind ausschließlich Abiturienten vorbehalten.” [Apprenticeships that place special demands on apprentices due to their economic importance and the high level of theoretical knowledge required are reserved exclusively for high school graduates.]

Nevertheless, the formal statement is accompanied by recommendations that appear contradictory. For instance, employment opportunities for high school students are found in agriculture and handicrafts (Landwirt, Hauer, Werkzeugmacher) or in the restaurant industry (Koch, Kellner).

Such formulations elicit interpretations that extend beyond their literal meanings. While these interpretations are inherently subjective, they illustrate how individuals who were socialized within the GDR might decipher latent political intentions. For instance, the assertion that Abiturienten should pursue careers as cooks or waiters is a subject of concern, particularly in light of the fact that only approximately 10 percent of each graduating class was permitted to obtain the Abitur, as documented by [49]. The rarity of this qualification suggests that such recommendations were not arbitrary. It is plausible that individuals in these positions were expected, whether voluntarily or under pressure, to demonstrate ideological conformity.

The interpretative framework is further reinforced by additional examples. The prioritization of occupations such as

miner (Hauer) can be linked to the ideological elevation of industrial labor within socialist society. In a similar vein, occupations such as precision mechanic or fine optician exhibited a strong correlation with industries that contributed significantly to the nation's foreign exchange earnings through exports. Watchmaking centers such as Glashütte and Ruhla, as well as optical enterprises like Carl Zeiss Jena, primarily produced goods for international markets, including the Federal Republic of Germany. These sectors demanded highly qualified professionals and were of strategic economic importance.

The interpretation of such texts is contingent upon contextual and experiential knowledge. The inferences concerning these documents are not explicitly stated, but rather, they emerge through the interaction between textual formulation, historical context, and the reader's tacit understanding, which is shaped by life in the GDR.

D. Religious texts

Tacit knowledge is also of crucial importance in the interpretation of religious texts, particularly in the context of the Old Testament [50], [51]. Such texts frequently depend on implicit theological assumptions, symbolic language, and historically embedded interpretive traditions that cannot be fully accessed through explicit textual analysis alone.

As stated in the research, “Polanyi and many of his interpreters recognized the implications of tacit knowing for restoring the Augustinian principle that faith precedes understanding” [52]. This perspective underscores the notion that comprehension within a religious framework does not stem solely from analytical processes, but rather is firmly rooted in preexisting commitments or convictions. In a similar vein, [1] posits the following: “We may envisage then a cosmic field which called forth all these centres by offering them a short-lived, limited, hazardous opportunity for making some progress of their own towards an unthinkable consummation. And that is also, I believe, how a Christian is placed when worshipping God”. In this sense, religious knowing is participatory and experiential rather than solely propositional.

This line of thought is further developed in the observation that “it is striking to me that if you take Polanyi's belief in the Pauline principle of justification by faith in Romans 1:17, you could restate it as follows: The one who is righteous shall live by tacit knowing of God's grace” [52]. In this context, tacit knowledge is conceptualized as an integral component of faith itself.

In light of the aforementioned theoretical framework, a rigorous evaluation of four prominent language models is conducted on the interpretive tasks delineated above. The texts selected for analysis encompass a diverse array of literary genres, including classical literature, theological discourse, philosophical argumentation, and concise conversational sentences.

IV. EXPERIMENTAL RESULTS

In this study, a range of LLM models, encompassing both larger and smaller scales, were examined through the uti-

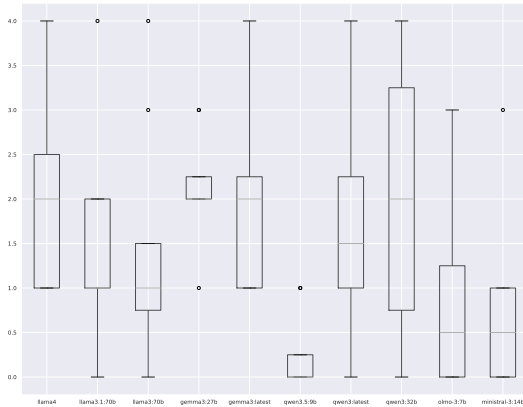


Fig. 3: Quantity of conjectures for political jokes

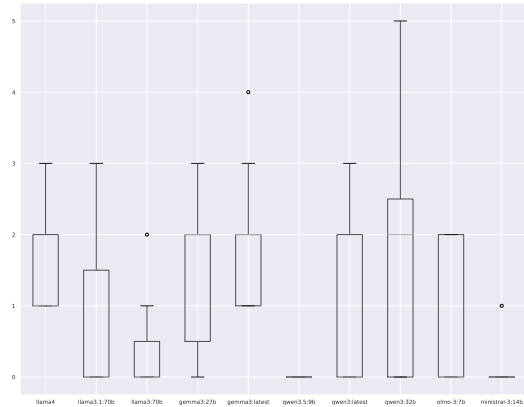


Fig. 5: Quantity of conjectures for biblical texts

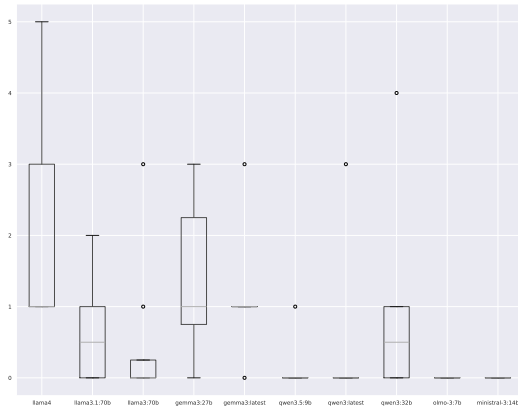


Fig. 4: Quantity of conjectures for GDR texts

This facilitates a comparison of the behavior of models with different numbers of parameters. It also allows for an analysis of how these models behave in different versions and how consistent their findings are. The source code, data, and prompts are publicly available at <https://zenodo.org/records/19566962>.

The empirical findings suggest that current LLMs do not exhibit stable hermeneutical behaviour. The outputs of these models are contingent on prompt formulation, model architecture, and language variation. As illustrated in Figures 1 and 2, the categories are distributed across the texts. It is evident that no two models propose equivalent values; nevertheless, frequently, models from the same category (for example, llama models) demonstrate a distribution that is rather similar. While the labels assigned broadly overlap, it should be noted that not all of them are associated with humour (e.g. German Literature, World War II, Religion, Shopping, Quiz, etc.). The findings concerning GDR texts are also distributed in a

dispersed manner; only the Minstral text is consistent with a political interpretation. With regard to biblical texts, a similarly intriguing distribution is evident. Whilst the contextualisation exhibited by llama3:70b and llama3.1:70b is deemed to be relatively consistent, it is evident that other religious labels are also present, including those which are designated as ‘German Proverb’ and ‘personal conflict’.

From the perspective of tacit knowledge theory, this raises important conceptual issues. Tacit knowledge is rooted in embodied practice and contextual participation. In contrast, LLMs are based on statistical regularities that have been learned from text corpora. This assertion is further substantiated by the findings of the present study. Despite the fact that their outputs occasionally bear a resemblance to interpretative acts, this similarity may be superficial. The observed variability indicates an absence of stable interpretive commitments.

As illustrated in Figures 3-5, the number of conjectures identified by the LLM in various texts is demonstrated. It is interesting to note that the results vary not only between models, but also between different texts. It is evident that the argument advanced by a LLM that there are no conjectures, yet that a context category is returned, is fallacious. From a hermeneutical perspective, interpretation is predicated on engagement with meaning as something that resists arbitrary segmentation. Machine-generated segmentations appear to reflect probabilistic clustering rather than understanding.

Nevertheless, it can be argued that if tacit knowledge manifests through reliable performance in interpretive tasks, then stability in outputs could approximate a computational shadow of tacit inference. The findings of this study suggest that there is a paucity of evidence to support the hypothesis that there is stability. The fundamental question remains: does interpretation necessitate consciousness and participation, or is formal pattern alignment sufficient?

V. CONCLUSION AND OUTLOOK

In this study, we sought to ascertain the extent to which LLMs possess the capacity to engage in hermeneutical analysis. The present study evaluated the consistency, interpretive depth and methodological coherence of LLM outputs by testing multiple texts on tacit understanding.

The findings indicate that current LLMs exhibit significant variability and lack reproducibility in the application of hermeneutical principles, which necessitate tacit knowledge. The findings suggest that contemporary models exhibit considerable variability in these hermeneutical tasks. Their performance appears to be sensitive not only to linguistic framing and prompt structure, as demonstrated in our previous study [53]. Across a range of models, there was inconsistency in the interpretation of tasks requiring tacit knowledge, both within and across iterations. This variability suggests that LLMs are neither applying coherent hermeneutical frameworks nor capable of acquiring tacit knowledge, but rather relying on surface-level linguistic patterns that diverge significantly depending on language, prompt design, and model architecture.

Nonetheless, this is incongruent with the assertions made in recent literature, which has purported that LLMs “can analyze nearly any textual statement.” [54] The experimental findings of this study appear to be in accordance with the assertions of researchers who propose that LLMs exhibit deficiencies in their ability to execute even fundamental logic-based tasks, such as counting and identifying general substructures in graphs (see [55], [56]) Therefore, the question of whether LLMs could offer any technical assistance with textual analysis remains a subject of debate. For instance, as [57] demonstrate in their examination, the potential for such assistance is unclear.

It is concluded that hermeneutics cannot be reduced to LLM-based text processing. In the philosophical tradition, tacit knowledge is defined as involving embodied competence and contextual immersion. These elements are not currently replicable by generative models. However, it is important to acknowledge the limitations of the present study. Firstly, there is no clear way in research to measure tacit knowledge. Consequently, we were compelled to devise an experimental approach that would facilitate the identification of pertinent questions. Secondly, the selection of LLMs was limited, and further research could be carried out on other commercial LLMs. Thirdly, the implementation of additional prompting strategies could facilitate a more profound comprehension of LLMs’ behavioural tendencies. The interplay between the person querying and their tacit knowledge, as well as the possible interpretations, could also be a subject for further research.

However, it is also recommended that future research should concentrate on the development of evaluation frameworks that measure interpretive coherence rather than superficial accuracy.

Coming back to the remark on anthropomorphism we did not find hints for LLMs being capable of tacit knowledge and

thus actual generative processes. This could be seen as a hint though on how LLMs ‘make’ sense: It is by their users interpreting their results. While in this study we explicitly tested the results for weather they are sensible, users would usually rather make sense and interpret the outcome of a prompt themselves instead of routinely check how other models fared and even running the same prompt multiple times for comparison. Against this background it is perceivable how said superficial accuracy can be regarded as correct interpretations. For this reason we would also strongly advise for the development and broad implementation of AI literacy and a less humanifying approach of interaction with such models.

ACKNOWLEDGMENTS

During the preparation of this manuscript, the authors used DeepL to check the language and grammar. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

DATA AVAILABILITY

The source code, data, and prompts are publicly available at <https://zenodo.org/records/19566962>.

REFERENCES

- [1] M. Polanyi, *Personal knowledge*. University of Chicago Press, 1958.
- [2] S. Turner, “Polanyi and tacit knowledge,” in *The Routledge Handbook of Philosophy and Implicit Cognition*. Routledge, 2022, pp. 182–190.
- [3] M. Tiemann, *Wissensintensität von Berufen und ihre Entlohnung. Klassifikation und empirische Ergebnisse*. BeltzJuventa, 2015, pp. 281–299.
- [4] M. Lejeune, “Tacit knowledge: Revisiting the epistemology of knowledge,” *McGill Journal of Education*, vol. 46, pp. 91–105, 2011.
- [5] L. Tomkins and V. Eatough, “Hermeneutics: interpretation, understanding and sense-making,” *SAGE handbook of qualitative business and management research methods*, pp. 185–200, 2018.
- [6] H.-G. Gadamer, “Destruktion und deconstruction,” *Dialogue and Deconstruction: The Gadamer-Derrida Encounter*, pp. 102–113, 1989.
- [7] M. Polanyi, “Tacit knowing,” *Philosophy Today*, vol. 6, no. 4, p. 239, 1962.
- [8] H. Collins, *Tacit and explicit knowledge*. University of Chicago press, 2019.
- [9] H. G. Gadamer, *Wahrheit und methode Grundzüge einer philosophischen Hermeneutik*, 1960.
- [10] H. M. Collins, “Tacit knowledge, trust and the q of sapphire,” *Social studies of science*, vol. 31, no. 1, pp. 71–85, 2001.
- [11] B. Trubody, “When tacit is not tacit enough: A heideggerian critique of collins’ “tacit” knowledge,” *Research in Hermeneutics, Phenomenology, and Practical Philosophy*, vol. 5, no. 2, pp. 315–335, 2013.
- [12] D. Ingram, “Hermeneutics and truth,” *Journal of the British Society for Phenomenology*, vol. 15, no. 1, pp. 62–78, 1984.
- [13] J. Loenhoff, “Tacit knowledge: Shared and embodied,” transcript Bielefeld, 2015, pp. 21–40.
- [14] I. Nonaka, “The knowledge-creating company,” in *The economic impact of knowledge*. Routledge, 2009, pp. 175–187.
- [15] T. Johnstone and D. R. Shanks, “Abstractionist and processing accounts of implicit learning,” *Cognitive Psychology*, vol. 42, pp. 61–112, 2001.
- [16] A. S. Reber, “Implicit learning of artificial grammars,” *Journal of Verbal Learning and Verbal Behavior*, vol. 6, pp. 855–863, 1967.
- [17] G. S. Insch, N. McIntyre, and D. Dawley, “Tacit knowledge: A refinement and empirical test of the academic tacit knowledge scale,” *The journal of psychology*, vol. 142, no. 6, pp. 561–580, 2008.
- [18] V.-W. Mitchell, W. S. Harvey, and G. Wood, “Where does all the ‘know how’ go? the role of tacit knowledge in research impact,” *Higher Education Research & Development*, vol. 41, no. 5, pp. 1664–1678, 2022.
- [19] R. Maciag, “The hermeneutics of artificial text,” in *Computer Sciences & Mathematics Forum*, vol. 8, no. 1. MDPI, 2024, p. 94.

- [20] L. Henrickson and A. Meroño-Peñuela, "Prompting meaning: a hermeneutic approach to optimising prompt engineering with chatgpt," *AI & SOCIETY*, pp. 1–16, 2023.
- [21] L. Tiantian and M. Carl, "Toward practical hermeneutics of fourth paradigm ai for science," *«Technology and language»*, vol. 5, no. 1, pp. 89–105, 2024.
- [22] K. Petković, "The truth and ai: Towards a new hermeneutic of the academic subject?" *Sociologija*, vol. 66, no. 4, pp. 581–611, 2024.
- [23] P. Pinell, "Does artificial intelligence speak our language?: A gadamerian assessment of generative language models," *Political Research Quarterly*, vol. 77, no. 3, pp. 713–728, 2024.
- [24] M. Heimann and A.-F. Hübener, "Circling the void: Using heidegger and lacan to think about large language models," *Cognitive Systems Research*, p. 101349, 2025.
- [25] A. Diržytė, "Large language models and the enhancement of human cognition: Some theoretical insights," *Filosofija. Sociologija*, vol. 36, no. 1, 2025.
- [26] M. Barnard and W. Otte, "Is the machine surpassing humans? large language models, structuralism, and liturgical ritual: A position paper," *International Journal of Practical Theology*, vol. 28, no. 2, pp. 289–306, 2024.
- [27] M. Chrostowski and A. J. Najda, "Chatgpt as a modern tool for bible teaching in confessional religious education: a german view," *Journal of Religious Education*, pp. 1–20, 2024.
- [28] R. Demichelis, "The hermeneutic turn of ai: Is the machine capable of interpreting?" *arXiv preprint arXiv:2411.12517*, 2024.
- [29] M. Gonzalez, "Beyond replacement: A hermeneutic phenomenological approach to human-ai collaboration," in *2024 IEEE International Symposium on Technology and Society (ISTAS)*. IEEE, 2024, pp. 1–4.
- [30] A. Elrod, "Uncovering theological and ethical biases in llms: An integrated hermeneutical approach employing texts from the hebrew bible," *HIPHIL Novum*, vol. 9, no. 1, pp. 2–45, 2024.
- [31] M. Trepczyński, "Religion, theology, and philosophical skills of llm-powered chatbots," *Disputatio Philosophica: International Journal on Philosophy and Religion*, vol. 25, no. 1, pp. 19–36, 2023.
- [32] K. Karches, "Hermeneutics as impediment to ai in medicine," *Theoretical Medicine and Bioethics*, pp. 1–19, 2025.
- [33] E. Krapp, R. Neuhaus, M. Hassenzahl, and M. Laschke, "In a quasi-social relationship with chatgpt. an autoethnography on engaging with prompt-engineered llm personas," in *Proceedings of the 13th Nordic Conference on Human-Computer Interaction*, 2024, pp. 1–16.
- [34] A. Li, H. Chen, H. Namkoong, and T. Peng, "Llm generated persona is a promise with a catch," *arXiv preprint arXiv:2503.16527*, 2025.
- [35] U. Oevermann, "Zur analyse der struktur von sozialen deutungsmustern," 1973. [Online]. Available: <https://publikationen.ub.uni-frankfurt.de/opus4/frontdoor/deliver/index/docId/4951/file/Struktur-von-Deutungsmuster-1973.pdf>
- [36] J. Reichertz, "Empirische sozialforschung und soziologische theorie," in *Handbuch Methoden der empirischen Sozialforschung*. Springer, 2022, pp. 69–86.
- [37] U. Oevermann, "Sozialisation als prozess der krisenbewältigung," in *Sozialisationstheorie interdisziplinär – Aktuelle Perspektiven*, D. Geulen and H. Veith, Eds. Lucius, 2004, pp. 155–181.
- [38] J. Reichertz, "Die objektive hermeneutik – darstellung und kritik," in *Qualitative Forschung*, E. König and P. Zedler, Eds. Beltz Verlag, 2002, pp. 123–156.
- [39] A. Placani, "Anthropomorphism in ai: hype and fallacy," *AI and Ethics*, vol. 4, pp. 691–698, 2024.
- [40] D. Aarons, "Puns and tacit linguistic knowledge," in *The Routledge handbook of language and humor*. Routledge, 2017, pp. 80–94.
- [41] R. E. Day and L. Ma, "Rethinking unsaid information: Jokes and ideology," in *Proceedings of the 2011 iConference*, 2011, pp. 63–67.
- [42] J. Newirth, "Jokes and their relation to the unconscious: Humor as a fundamental emotional experience," *Psychoanalytic Dialogues*, vol. 16, no. 5, pp. 557–571, 2006.
- [43] V. Tsakona, "Intertextuality and/in political jokes," *Lingua*, vol. 203, pp. 1–15, 2018.
- [44] D. Bathrick, *The Powers of Speech: the Politics of Culture in the GDR*. U of Nebraska Press, 1995.
- [45] U. Altschüler and I. Kemble, "Ist die ddr-sprache abgewickelt?(has gdr-language been wound up?)," *Journal of Area Studies*, vol. 1, no. 2, pp. 12–24, 1993.
- [46] U. Fix, *Sprache, Sprachgebrauch und Diskurse in der DDR: ausgewählte Aufsätze*. Frank & Timme GmbH, 2013, vol. 15.
- [47] W. Thierse, "'spricht, damit ich dich sehe'. beobachtungen zum verhältnis von sprache und politik in der ddr-vergangenheit," in *Deutsch als Verkehrssprache in Europa*. de Gruyter, 1993, pp. 114–126.
- [48] A. Borgwardt, "Politik und literatur in der ddr," in *Im Umgang mit der Macht: Herrschaft und Selbstbehauptung in einem autoritären politischen System*. Springer, 2002, pp. 77–136.
- [49] I. Miethe, "„die universität dem volke!“ entwicklungsphasen der arbeiter-und-bauern-fakultäten (abf) der ddr," *Beiheft zum Report: Erwachsenenbildung und Demokratie*, vol. 26, no. 1, pp. 215–224, 2003.
- [50] T. Krüger, "Tacit knowledge in the old testament," *Archaeology of Mind in the Hebrew Bible/Archäologie alttestamentlichen Denkens*, p. 347, 2023.
- [51] T. A. Demmons, "Tacit and tactile knowledge of god: Toward a theology of revelation for persons with intellectual disabilities," *Journal of Religion, Disability & Health*, vol. 11, no. 4, pp. 5–21, 2008.
- [52] R. L. Gelwick, "From tacit knowing to a theory of faith1," *The Polanyi Society Periodical*, p. 10, 2014.
- [53] J. Dörpinghaus and M. Tiemann, "Do llms dream of antique hermeneutics? critical remarks on automated text interpretation," in *2025 20th Conference on Computer Science and Intelligence Systems (FedCSIS)*. IEEE, 2025, pp. 687–692.
- [54] P. Törnberg, "How to use llms for text analysis," *arXiv preprint arXiv:2307.13106*, 2023.
- [55] T. Fu, R. Ferrando, J. Conde, C. Arriaga, and P. Reviriego, "Why do large language models (llms) struggle to count letters?" *arXiv preprint arXiv:2412.18626*, 2024.
- [56] L. Nguyen and Y. Yan, "Evaluating the structural awareness of large language models on graphs: Can they count substructures?" *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*, 2024.
- [57] R. H. Tai, L. R. Bentley, X. Xia, J. M. Sitt, S. C. Fankhauser, A. M. Chicas-Mosier, and B. G. Monteith, "An examination of the use of large language models to aid analysis of textual data," *International Journal of Qualitative Methods*, vol. 23, p. 16094069241231168, 2024.