Journal of Digital Terminology and Lexicography - ISSN:3103-3601
Vol. 2, No. 2 (2026) DOI: 10.25430/pupj.jdtl.1794322612

Terminologization of Anthropomorphic Metaphors in the Italian Discourse around AI

Rossella Resi

Department of Information Engineering, University of Padova, Via Gradenigo 6/B, Padova, 35131, Italy.

Corresponding author(s). E-mail(s): rossella.resi@unipd.it

Abstract

This paper investigates anthropomorphic language in Italian discourse on artificial intelligence (AI), asking whether and how lexical expressions attributing human traits to AI systems undergo terminologization, the process by which a metaphorical expression is conventionalized as a technical term, with a corresponding loss of perceived figurativity. After building a targeted web corpus through a semi-automated bootstrapping procedure (BootCaT, 2004), a dataset of 162 anthropomorphic expressions was extracted from 68 accurately selected Italian sources (2023–2026) and annotated across four levels: semantic category (Inie, Zukerman, and Bender (2026)'s eight-category taxonomy), grammatical level, distancing markers (quotation marks, negation, hedges), and source register. Five hypotheses tested frequency, diachronic trajectory, grammatical level, register injection, and semantic category as proxies for terminologization. Results show a structural split rather than a uniform trend: nominal role-metaphors terminologize silently, as predicted, showing low distance marking despite high frequency, while verbal cognition-metaphors remain heavily marked and negated regardless of frequency, sustained by an ongoing public debate on machine cognition. Contrary to expectation, distancing markers increased corpus-wide between 2023 and 2026, driven by implicit rather than explicit hedging, suggesting that critical metalinguistic discourse can itself act as a brake on terminologization. At the individual-lemma level, there are cases where the three tested proxies behave as expected: marking falls with frequency, over time, and as terms enter institutional use. The two findings are not contradictory: individual terms quietly terminologize even as the broader critical discourse keeps the metaphorical status visible, slowing terminologization as a corpus-wide phenomenon.

Keywords: Terminologization, Anthropomorphism, Metaphor, Artificial Intelligence, Italian, Specialized Discourse

1 Introduction

Metaphor is a common strategy for creating specialized terms. One of the most frequent metaphors in the domain of technology is anthropomorphism.[1] The attribution of human traits, human capacities, or mental states to objects that do not in fact possess them, pervades specialized discourse on artificial intelligence (AI),[2] yet the linguistic status of these metaphors is not always clear. Verbs like *to learn* with *model*

[1] In this article, the term anthropomorphism refers specifically to the attribution of human traits, human capacities and mental states to non-sentient objects through linguistic expressions, hence the use of anthropomorphic language, while acknowledging that this represents only one of several possible forms of anthropomorphism associated with technology.

[2] The term artificial intelligence (AI) is used here in a deliberately broad sense to encompass a wide range of computational technologies that have shaped its evolution. This includes machine learning, deep neural networks, and more recent developments like transformers, large language models, and multimodal generative models amongst others.

Copyright ©2026 The Author(s)

as subject or nouns like *assistant* to refer to *chatbot* can either be a figurative shorthand with epistemological and pedagogical functions or conventionalized technical terms, and it is often unclear which of the two applies. This paper investigates the process by which a metaphorical expression becomes conventionalized as a specialized term, hence undergoing the process of terminologization, entailing loss of perceived figurativity. While the anthropomorphizing tendency of AI-related language has been extensively documented for English, the phenomenon has not previously been examined in Italian at the level of terminology. This paper addresses this gap through a corpus-based study that asks not only how anthropomorphic language is categorized in Italian, but whether, and under what conditions, it can be said to have become terminologized.

The paper is organized as follows. Section 2 reviews metaphor in terminology, introduces the concepts of terminologization and determinologization, and examines anthropomorphic language in AI discourse, including the taxonomy of Inie et al. (2026). Section 3 discusses how metaphorical meaning relates to terminologization and proposes four criteria for measuring it: frequency, diachronic trajectory, register injection, and semantic category. Section 4 turns to Italian, treating anthropomorphism as a cross-linguistic phenomenon shaped by language-specific factors, and proposes distancing markers as a further, Italian-specific proxy for terminologization. Section 5 presents the corpus, the annotation scheme, and the five hypotheses tested. Section 6 reports the results for semantic category, distancing markers, frequency, and register injection, their grammatical realization and diachronic breakdown. Section 7 summarizes the findings and outlines directions for further research.

2 Metaphor, Terminology, and Anthropomorphism in AI Discourse

2.1 The Use of Metaphor in Specialized Discourse

Specialized discourse relies pervasively on metaphorical language. Temmerman (2000) argues that metaphor is not an accidental or imprecise feature of specialized language but one of the central cognitive mechanisms by which experts create categories and coin terms. Understanding a new scientific or technical concept, on this view, typically involves constructing a mental model. This model is often built by analogy with an already familiar experience, and it can later be stretched and adapted as the phenomenon it describes becomes better understood. Metaphor therefore operates not merely at the level of naming, but also at the levels of conceptualization and categorization. It involves the transfer of meaning from a source domain, where a word or term is already established, to a target domain, where it acquires an additional domain-specific meaning. For instance, the general-language adjective *friendly* is used in the expression *friendly bacteria* to characterize bacteria as beneficial. Similarly, *heartbeat*, originally associated with the physiological domain, can be metaphorically extended to the economic domain in expressions such as *economic heartbeat*, where it evokes the rhythm or vitality of economic activity.

Terms such as *friendly bacteria* and *economic heartbeat* illustrate how metaphorical expressions drawn from known experience become integrated and normalized within technical vocabularies across disciplines, independent of domain. This conceptual function is inseparable from a pedagogical and epistemological one. Metaphors act as cognitive tools: they support analogical reasoning, and facilitate communication between experts, particularly in emerging fields or when addressing novel phenomena. They also make complex or abstract concepts more accessible to lay audiences (Gentner & Jeziorski, 1993; Pascolini, 2004; Temmerman, 2000).

This positive assessment of metaphor's role in terminology is not uncontroversial, and the contrast is instructive for framing the present study. The Vienna School, or General Theory of Terminology (Wüster, 1979), historically adopted a cautious stance toward metaphor. Figurative naming produces polysemy and threatens the univocity ideal of the one-term-one-concept principle that classical terminology theory takes as a normative goal. Temmerman (2000)'s sociocognitive approach explicitly rejects this stance, arguing that categorization in the mind does not respect the clear-cut boundaries the Vienna School presupposes; metaphor-driven polysemy is instead a predictable and functional consequence of how specialized concepts are actually formed and understood.

This productive view of metaphor has to be qualified by the risk of polysemy it also introduces: an excess of figurative naming may hinder rather than support understanding. A term may lose its original conceptual boundaries as it is extended to cover a widening range of referents. Metaphor in terminology

therefore functions as a double-edged instrument: the same associative productivity that allows experts to name new phenomena efficiently is also what allows a term's reference to drift beyond its original conceptual scope. This duality is especially salient in domains characterized by rapid innovation, where a stable terminology is lacking and metaphor is called upon to do quick naming work.

Within this broader landscape of metaphor-driven term formation, technology as a domain shows a marked tendency specifically toward anthropomorphic metaphors: the mapping of human properties, human capabilities, and mental states onto non-human systems (Epley, Waytz, & Cacioppo, 2007; Salles, Evers, & Farisco, 2020). Such metaphors are rarely isolated coinages; they generate entire families of related terms, so that a single source-domain logic continues to shape how new concepts are named as a field develops (Temmerman, 2000). This is not merely a stylistic preference but a cognitive-structural mechanism of term formation. Consistent with the sociocognitive approach outlined above, this mechanism becomes increasingly important as AI, a field that draws extensively on concepts from human cognition, enters both scientific discourse (proving terminologization) and public discourse (possibly causing determinologization) (See Section 2.2).

2.2 Terminologization and Determinologization

Metaphorical term formation is recognized as one of the term-formation methods that use existing forms rather than create new ones. More specifically, it involves "assigning new, frequently analogous or metaphoric meanings[,] to existing terms in more or less related fields or [general-purpose-language] words" (Infoterm, 2005). This mechanism can be understood as one possible manifestation of the broader process of *terminologization*, whereby "a general language [...] word or lexical unit is being used more and more as a term [...] in a specific domain [...] or subject [...]" (Costa et al., 2022). Costa et al. (2022)'s own phrasing "used more and more as a term" already signals that terminologization is gradual rather than binary: a lexical expression does not switch abruptly from general to specialized use, but moves along a continuum of increasing termhood.[3]

The reverse process, known as *determinologization*, involves the movement of a term from specialized discourse into general language (Infoterm, 2005). A term in this case "does not refer to a concept anymore and[,] it is no longer part of a concept system within a given domain," losing its identity and specificity as it comes to be used as a lexical expression of general language (Costa et al., 2022), through what is characterized as "a rather unconscious process" (Humbert-Droz, 2023). The risk of entering general language is that of stretching the boundaries of the concept in question (as well as of related concepts) beyond their original conceptual scope. The word *translation*, for instance, once denoted unambiguously a human activity. As terms like (as well as the practice of) *machine translation* became general knowledge, the boundaries of the concept of *translation* blurred to the point that it has now become necessary to specify *human translator* and reconstitute the distinction the plain noun no longer reliably marks. Anthropomorphic metaphors in AI discourse, examined in the following Section 2.3, are a case in which these two movements are visible.

2.3 Anthropomorphic Language in AI

Anthropomorphism can be defined as the "tendency to imbue the real or imagined behavior of nonhuman agents with humanlike characteristics, motivations, intentions, or emotions" (Epley et al., 2007). Its consequences for human–computer interaction are not new: the famous *ELIZA effect* describes the long-documented human tendency to attribute understanding and empathy to computational systems that merely simulate dialog (Nass & Moon, 2000; Weizenbaum, 1966).

In recent years, anthropomorphism has intensified markedly in both scientific and popular discourse on AI (Bunz & Braghieri, 2022; Ibrahim & Cheng, 2025). One key driver is that many of the most consequential advances in machine learning have been directly or indirectly shaped by theories from neuroscience, cognitive psychology, and social epistemology. As a result, terms such as *memory*, *learning*, and *neural network*, together with more recent coinages such as *thinking model* (introduced into scientific discourse in January 2025), and verbs such as *to understand*, *to think*, *to reason*, or *to hallucinate*, suggest intentionality and mental states that these systems do not, in fact, possess. Whether such language has to be

[3] Termhood is the "degree to which a lexical unit is recognized as a term" (International Organization for Standardization, 2023) or "the degree that a linguistic unit is related to (or [...] represents) domain-specific concepts" (Kageura & Umino, 1996). This gradualness is precisely what makes the process resistant to introspective judgment alone, and motivates the search for measurable proxies capable of approximating the degree of terminologization empirically.

taken metaphorically or literally is sometimes genuinely ambiguous even to expert users (Watson, 2019). Beyond direct analogy with cognitive science, Epley et al. (2007)'s three-factor theory of anthropomorphism further identifies independent mechanisms driving this trend: the accessibility and applicability of anthropocentric knowledge, the motivation to explain and predict the behavior of other nonhuman entities (Waytz et al., 2010), and the desire for social contact and affiliation.

The scale of this phenomenon, combined with a general lack of explicit acknowledgment of the metaphorical status of such terms, has raised substantive concerns about its social implications. Anthropomorphic language may inadvertently shape public and professional perceptions of AI by fostering misinformation about the actual capabilities of systems (Lipton, 2018), promoting overreliance on automated outputs (Vasconcelos et al., 2023), decreasing accountability (Kempton, Parmiggiani, & Vassilakopoulou, 2023; Tiku, 2022), and fostering emotional attachment to non-sentient artifacts (Waytz et al., 2010). What distinguishes the present moment from earlier instances of this phenomenon is that the pervasive presence of AI across public discourse means that every level of the domain's vertical stratification, from lay to expert usage, is now permeated by anthropomorphic language. Human-centered concepts are used with the additional risk of stretching concepts' boundaries and favoring semantic or conceptual shift that Costa et al. (2022) and Meyer and Mackintosh (2000) discuss under determinologization (Section 2.2). So the term *intelligence* has expanded its scope to encompass both *human* and *artificial intelligence*, making it necessary to specify the type of intelligence being referred to rather than using the term alone when discussing humans.

At the expert end, human-centered concepts remain cognitively accessible also to specialists (Shneiderman, 2020), and anthropomorphism risks obscuring lines of responsibility, displacing accountability from developers and corporations onto the systems themselves (Tiku, 2022; Watson, 2019), a concern well illustrated for example by the public controversy surrounding claims of sentience in Google's LaMDA system (Tiku, 2022). These concerns have generated a growing body of literature calling for clearer distinctions between metaphorical and literal uses of humanizing terms, and hence for greater linguistic transparency in describing AI systems (Rahwan et al., 2019; Shanahan, 2023; Watson, 2019; Ye & Li, 2024).

2.3.1 Categorization of Anthropomorphism

Because anthropomorphic language in AI is neither uniform nor binary, several recent studies have sought to categorize it systematically. Ryazanov, Öhman, and Björklund (2025) distinguish established anthropomorphizing terms in the field (e.g., *machine learning*, *pattern recognition*) from task-based anthropomorphism, which explains a system's functionality in human terms, and high anthropomorphism, which implies genuine cognitive abilities or feelings; the latter two categories are further subdivided according to the strength and source-domain of the anthropomorphic expression. Working from a complementary, more fine-grained perspective, Inie et al. (2026), building on earlier work (Inie, Druga, Zukerman, & Bender, 2024), construct a taxonomy of eight categories of anthropomorphic language: `Cognizer` (attributions of cognitive capacities such as understanding or believing), `Products of cognition` (outcomes attainable only through cognitive activity, such as skills, capabilities, or biases), `Emotion` (e.g., struggling, empathizing, caring), `Communication` (verbs of dialogic exchange, such as asking or answering), `Agent` (intentional, independent action, such as collecting, monitoring, or choosing), `Human role analogy` (naming a system through a human social role, e.g., *tutor*, *assistant*, "AI doctor"), `Names and pronouns` (proper names and animate pronouns such as *I*, *we*, *he*, *she*, *they*), and `Biological metaphors` (biological vocabulary applied to computational processes, such as *neural net*, *neuron*, *synapse*).

Taken together, these two taxonomies show that anthropomorphic language in AI discourse is internally stratified by strength and source domain, and that it ranges along a terminologization continuum from clearly figurative to deeply conventionalized. Bowdle and Gentner (2005) argue that terminologization shifts the process from comparison to categorization. However, it is questionable whether the figurative sense remains stored within a fully *terminologized* expression. In other words, what happens to the perception of an anthropomorphic metaphor as it moves along this continuum toward terminologization?

3 The Perception of Metaphorical Meaning

Framing the question mentioned at the end of the previous section requires investigating the traditional idea that a metaphor, once conventionalized, ceases to be metaphorical and becomes an arbitrary term. Lakoff and Johnson (1980) argue instead that a conceptual metaphor does not die upon becoming conventional: it becomes *invisible*. It ceases to be consciously perceived as a figurative projection from one domain to another, while the underlying conceptual mapping between source and target domains remains fully operative and continues to constrain how the target concept is reasoned about. New terms from the same source domain keep being coined, in fact, which shows that the metaphor remains productive within the same semantic area. Criteria for determining the liveness of metaphorical meaning in general language are said not to exist, since metaphorical activation is a property of use rather than of the lexeme itself (Müller, 2008). It is nonetheless worth asking whether such criteria can be identified within the more circumscribed boundaries of a specialized domain. The working assumption here is that once a metaphorical expression has terminologized, it has acquired a stable conceptual referent within the target domain and may therefore cease to be perceived as metaphorical: it designates a concept that is defined without explicit reference to its source domain. Conversely, as long as a metaphor remains perceptible as such within a specialized domain, that is, remains *visible*, its users retain some capacity to consciously evaluate its appropriateness, contest its implications, and reject inferences that do not hold. The richness of the derivational cluster built around the same source domain may be evidence that the terminologization process is taking place but it is not concluded. During the process of terminologization a metaphor may be able to import an inferential logic from the source domain. Whether the metaphorical meaning persists even in a fully terminologized expression (e.g., the term *prompt*) is still controversial.

Anthropomorphic metaphor in AI discourse can be understood as instantiating this general mechanism through a specific process of terminologization. An anthropomorphic expression is first deployed as a comprehension tool, attributing human traits to explain or make a technical phenomenon accessible (e.g., the system *learns*, the algorithm *decides*). Through habitual use, repeated over time by a community of experts, the expression ceases to be an occasional choice and becomes the standard way of referring to that phenomenon (the model *hallucinates*). Precisely because it is used this way so consistently, a linguistic bias feeds into conceptual attribution: speakers begin to actually conceptualize the phenomenon in these terms, not merely to talk about it that way, a shift from lexical anthropomorphism to a conceptual anthropomorphism. At this point, terminologization proper occurs: the expression becomes a term, and its metaphorical origin fades out from users' awareness, no longer felt as an image but as the label of the process it denotes (*deep learning*, *artificial intelligence*). During terminologization, the term itself and the source domain may become the base for new, related cognitive metaphors, like *thinking model* and, as concept produced by sisterhood, *non-thinking model*, which restart the cycle. The relationship between linguistic use and conceptual effect is therefore circular rather than linear: linguistic use produces a conceptual effect, and that conceptual effect in turn reinforces and generates further linguistic use, such that no single stage can meaningfully be identified as logically prior. This conceptual framing, rather than isolation of linguistic use, raises the concerns discussed in Section 2.3 about the use of anthropomorphic language.

If terminologization systematically reduces the perceptibility of a metaphor's meaning, then the degree of terminologization becomes a plausible index of the extent to which a term's anthropomorphic meaning remains perceptible to language users. Searching for objective criteria capable of operationalizing the degree of terminologization independently of individual introspective judgment therefore becomes important. Among the plausible criteria for this purpose are *frequency*, on the assumption that increased and more evenly distributed frequency across registers indexes high degree of terminologization; *diachronic trajectory*, tracking stability and occurrences over time; *register injection*, in particular whether a term is present in institutional text despite circulating primarily in dissemination or commercial texts; and *semantic category*, assuming that some categories may trigger terminologization quicker than others. Further, language-specific criteria are also possible, as discussed in Section 4. This set of criteria provides the theoretical and methodological basis on which the perception of anthropomorphic metaphors in AI terminology can be investigated empirically, and on which the present study's corpus-based analysis is grounded.

4 Anthropomorphic Terminology in the Italian Discourse around AI

The mechanisms outlined above are not specific to English. Cognitive linguistics research emphasizes that anthropomorphism is a cross-linguistic phenomenon (Epley et al., 2007; Shanahan, 2023). How anthropomorphism manifests, however, can differ considerably across linguistic and cultural contexts, particularly when it functions as a naming device within specialized discourse rather than as a purely rhetorical or literary device (Temmerman, 2000; ten Hacken & Panocová, 2024).

Recent studies have described the systematic use of anthropomorphic language in Italian popular-science communication about AI, functioning simultaneously as a heuristic tool for scientific exploration and as a communication device for effective transmission of knowledge across disciplines (Calaresu, 2025; Pascolini, 2004; Stanič, 2025). Stanič (2025) analyzes the Italian journalistic discourse on AI through conceptual metaphor theory from a cognitive-linguistic point of view. On the methodological side, Vernillo (2024) develops a procedure for the systematic identification of linguistic metaphors in Italian. Complementing this literature, related scholarship has examined more broadly how language frames the perceived capacities of AI systems (Petricini, 2026). However, no study has systematically investigated anthropomorphism in Italian AI discourse at the level of terminology. The present study seeks to address this gap by identifying and categorizing anthropomorphic expressions, examining the extent to which they have become terminologized, and assessing the degree to which they retain figurative meaning.

Anthropomorphism in Italian specialized discourse in the domain of AI is grammatically distributed primarily across nominal and verbal categories. Adjectives are typically absorbed into multi-word terms rather than functioning as independent anthropomorphizing markers (e.g., *modello pensante*, 'thinking model'). A further language-specific factor is that, unlike English, every Italian noun carries grammatical gender. This property has been proposed as a further facilitating condition for personification, as an instance of anthropomorphism (Calaresu, 2025), since gender assignment can align a technical noun with human-like categorial expectations independently of any deliberate figurative intent. Italian anthropomorphic language is also permeated by distancing markers, such as quotation marks, italics, or metalinguistic qualifiers like *una sorta di* ('a sort of'). These mark an expression as an extended or figurative use, as opposed to appearing unmarked as ordinary terminology (Resi, 2026). This offers a further corpus-based diagnostic for investigating the degree of terminologization. A final complicating factor specific to Italian terminology within AI discourse is its considerable reliance on English borrowing, whose motivations have been analyzed from several angles for other domains by Resi and ten Hacken (in press). For the purposes of terminologization, this borrowing pattern is theoretically significant. An anglicism's semantic transparency, and with it its capacity to be perceived as figurative, may differ systematically from that of a native equivalent. As a result, the same underlying conceptual metaphor can carry a different degree of perceived anthropomorphism depending on whether it surfaces through a borrowed or a native Italian term. Less metaphorical perception may speed up the process of terminologization by promoting the perception of the technicality of a borrowing.

Taken together, the cross-linguistic productivity of anthropomorphic metaphor, Italian's specific grammatical resources for personification, the differential terminologization behavior of nominal versus verbal anthropomorphism, the massive presence of distancing markers, and the pervasive, sociolinguistically motivated presence of English borrowings jointly indicate that anthropomorphism in Italian AI discourse cannot simply be assumed to mirror its English-language counterpart. Combined with the terminologization criteria established in Section 3, and with the absence of any existing linguistic-level study of the phenomenon in Italian, these considerations motivate the empirical, corpus-based investigation of this paper.

5 Methodology

5.1 The Italian Corpus

One of the primary challenges in building a corpus for the Italian language lies in the scarcity of scientific literature specifically focused on AI that is written in Italian. Although AI is often mentioned in academic

publications in relation to other disciplines, the existence of a consistent domain-specific scientific discourse on AI in Italian remains limited. The corpus was constructed to approximate scientific Italian as closely as possible within the domain, while acknowledging a clear domain-related loss with respect to *Theoriesprache* (Ischreyt, 1965) in Italian.

To build such a targeted corpus of AI-related discourse in Italian, a semi-automated bootstrapping approach using seed terms and tuples was adopted at first, supported by the BootCaT toolkit (BootCaT, 2004). The BootCaT method leverages algorithmic randomness to automatically construct domain-specific web corpora: starting from a small user-provided list of domain seeds, it randomly shuffles them into combinations (tuples), which are submitted as search-engine queries to harvest relevant text. The process is iterative and, at each step, the retrieved websites can be manually inspected or ignored to narrow the scope of the next query.

Using randomly combined tuples to discover text offers the advantage of avoiding overfocusing. Randomly shuffling seeds creates unique semantic combinations that force search engines to index a wider array of sources than a more directly targeted query would, bridging linguistic contexts that a manually compiled list would not otherwise reach. This randomness generates an objective, diverse, and well-distributed sample of how language is naturally used within a domain constrained by the seeds and tuples supplied and by the manual screening that was carried out at each round of querying. Since random queries cannot fully control for search biases related to search-engine optimization practices on the part of indexed websites, nor for undocumented adjustments to indexing algorithms on the part of search engines themselves, the risk of bias is minimized but not removed.

The actual composition of a web corpus can only be established with certainty once the corpus is complete, and its adequacy, focus, and quality must therefore be assessed through post hoc evaluation (Baroni, Bernardini, Ferraresi, & Zanchetta, 2009). Accordingly, the final list of retrieved URLs was manually checked to enhance corpus quality, consistent with recognized concerns about indexing bias in web-derived samples (Henzinger, Heydon, Mitzenmacher, & Najork, 1999). The final corpus contains 82 texts from 68 sources, spanning three broad registers: institutional and technical outlets (e.g., professional and research organizations, university research centers, educational and encyclopedic portals, public and governmental bodies, academic theses), corporate sources (company websites and blogs), and journalistic and popular-science publications (national, regional, and online newspapers, popular-science and current-affairs magazines, press agencies, and professional or personal blogs). The final list of URLs was stored in a separate file without web scraping or preliminary automated text extraction, so that a dedicated script could load the URLs directly from this file, open them, parse their content, and proceed with the anthropomorphism analysis described below (Resi, 2026).

The AnthroScore metric was used to identify 104 sentences with a presumed high level of anthropomorphism (Resi, 2026). The procedure worked in both directions: for nominal entities, sentences were identified by masking the noun and estimating the probability that the masked slot would be filled by a human versus a non-human pronoun, with the log-ratio between the two probabilities providing the anthropomorphism score. Second, when the target entity was a verb, the same masking and substitution procedure was applied to the subject of the verbal phrase. The result of this procedure is a dataset of 162 examples (since 58 of the 104 sentences included more than one anthropomorphic expression) of lexical expressions belonging to the domain of AI, for which anthropomorphism was computed.[4]

5.2 Annotation

The dataset was annotated on four levels, combining a linguistically motivated taxonomy with independent variables relevant to the terminologization hypotheses (see Section 5.3). Annotation was conducted primarily by hand; automated or LLM-assisted procedures were explored for selected levels, but all automatically generated labels were manually reviewed before inclusion in the final dataset.

Level 1: Semantic category, following the eight-category taxonomy proposed by Inie et al. (2026) (`Cognizer`, `Agent`, `Human role analogy`, `Emotion`, `Biological metaphors`, `Communication`, `Names and pronouns`, `Products of cognition`). An initial attempt at automatic annotation with GPT-5.6 Sol was discarded: the model tended to flatten distinctions, forcing marginal cases into existing categories, for example, generating a hallucinated instance of `Products of cognition` while assigning

[4]The complete dataset is available as Github repository at https://github.com/RossellaResi/AnthroTerm.

Table 1 One example per category, annotated across Levels 1–4

ID	Expression (gloss)	Category (L1)*	Grammar (L2)	Distancing Marker (L3)	Register (L4)
1a	*l'IA impara a pensare come noi* 'AI learns to think like us'	`Cognizer`	Verbal	No	Journalistic/ opinion
16d	*l'IA agentica ragiona, pianifica e agisce; diventerà un partner collaborativo* 'agentic AI reasons, plans, and acts; it will become a collaborative partner'	`Agent`	Verbal	No	Corporate/ marketing
9	*l'IA agisce come assistente nella diagnosi medica* 'AI acts as an assistant in medical diagnosis'	`Human role analogy`	Nominal	No	Journalistic/ opinion
29a	*l'IA non prova emozioni ma le riconosce, interpreta e simula* 'AI does not feel emotions but recognizes, interprets, and simulates them'	`Emotion`	Verbal	Partial	Corporate/ marketing
41	*analogie tra il cervello e l'IA nella gestione dei big data* 'analogies between the brain and AI in big data management'	`Biological metaphors`	Nominal	No	Journalistic/ opinion
64b	ChatGPT fa un'osservazione acuta; l'IA *"ammette"* 'ChatGPT makes an insightful observation; AI "*admits*" (in quotation marks)'	`Communication`	Verbal	Yes	Journalistic/ opinion
83a	*Claude ha insistito di essere un essere umano e ha tentato di licenziare dei dipendenti* 'Claude insisted it was a human being and attempted to fire employees'	`Names and pronouns`	Verbal	No	Institutional/ technical

*`Products of cognition` is an empty category in the Italian dataset; see Section 6.1.

two clearly unrelated examples to arbitrary categories; a manual check of the first ten rows found four mislabeled examples. To avoid the classification bias introduced by this automated procedure, Level 1 annotation was ultimately performed manually for the complete dataset and, allowing multi-labeling for sentences, a total of 173 labels were applied (see Table 1 for some examples).

Level 2: Grammatical level. Each example was first classified according to its primary grammatical realization: verbal (the attribution is carried by a verb) or nominal (the attribution is carried by a noun). This annotation was initially generated via a GPT-5.6 Sol query and subsequently verified manually (see Table 1 for some examples).

Level 3: Distancing markers (Resi, 2026). Each example was annotated for the presence and type of distancing marker, following a three-way scheme: *Yes* (an explicit formal marker, such as quotation marks, or an explicit simile such as *è come*, 'it is like'); *No* (flat, unhedged usage with neither typographic nor content-based caution); and *Partial*, capturing cases in which the text signals awareness of a term's non-literal status without resorting to quotation marks. The *Partial* category covers two distinct mechanisms: metalinguistic comment, in which the writer steps outside the claim to comment on the term itself, signaling awareness of its contested or approximate use without quotation marks; and systematic negation without quotes, in which a cognitive or emotional verb is used but explicitly denied, so that the writer is not claiming the AI system does something, without this qualifying as a full *Yes* in the absence of quotation marks. Each example was additionally coded for the presence of grammatical negation of the anthropomorphic predicate (e.g., *non capisce*, 'does not understand'), independently of the distancing-marker label, to allow computation of an effective negation rate per semantic category. Table 2 illustrates the three-way distancing-marker scheme with representative examples from the dataset.

Level 4: Source register. Each example was additionally labeled with its year of publication and, for its source, a register label. Automatic detection via GPT-5.6 Sol query produced 43 distinct source-type labels,

Table 2 Examples of distancing markers (Level 3)

ID	Expression (gloss)	Label	Distancing Marker
1a	*l'IA impara a pensare come noi* (title) 'AI learns to think like us'	No	unhedged, no distancing marker
3	*i modelli "imparano" come sembra il ragionamento umano* 'models "learn" in a way that resembles human reasoning'	Yes	quotation marks around the verb
53	*l'IA è come un bambino molto piccolo che impara a bere da una bottiglia* 'AI is like a very young child learning to drink from a bottle'	Yes	explicit simile (*è come*)
56a	*l'IA non capisce di cosa parla, ma sa come parlarne* 'AI does not understand what it is talking about, but knows how to talk about it'	Partial	metalinguistic comment
23c	*l'intelligenza generativa non capisce, non ragiona, non sa quello che dice* 'generative intelligence does not understand, does not reason, does not know what it is saying'	Partial	negation, no quotation marks

which were manually grouped into three macro-registers, *institutional/technical*, *journalistic/opinion*, and *corporate/marketing*, to provide an independent variable for the terminologization analysis (see Table 1 for some examples).

5.3 Hypotheses

In order to investigate the process of terminologization, five initial hypotheses were formulated. These were then subsequently refined and expanded in light of the first empirical findings. Each initial hypothesis tests whether a measurable proxy can be associated with a specific variable as an indicator of terminologization.

H1 (Frequency). The degree of terminologization is operationalized as the token frequency of the lemma in the corpus; awareness of the metaphor is operationalized as the presence versus absence of a distancing marker (*Yes*/*Partial* vs. *No*). The prediction is a negative correlation: higher-frequency lemmas should carry fewer distancing markers, since distancing markers such as quotation marks, explicit negation, and hedges downplay the literal reading of the expression and thus neutralize, rather than reinforce, its anthropomorphic effect (Resi, 2026), acting as a direct proxy for awareness of the metaphor.

H2 (Diachronic trajectory). Building on H1, distancing markers are predicted to decrease over time, based on the assumption that terminologization is a diachronic process that develops progressively.

H3 (Grammatical level). Verbs are predicted to constitute a harder case of terminologization than nouns, owing to their more dynamic conceptualization and their dependence on entity–verb co-occurrence. Nouns are expected to terminologize faster because terminology is predominantly nominal. Italian grammatical gender facilitates personification of nouns specifically and, on the level of single terms, English borrowings, largely unavailable to verbs, favor technical reading. This motivates analyzing verbs and nouns separately throughout the study.

H4 (Register injection). Institutional/technical sources are predicted to show lower distancing marking than journalistic or corporate sources. This prediction extends to the diachronic dimension: if institutional texts show lower marking while journalistic and critical registers remain high, this will constitute strong and clean evidence of terminologization, indicating that the term has settled in the register where technical terms properly reside, independently of any year-to-year drift in corpus composition.

H5 (Semantic category). It is predicted that the degree of terminologization varies systematically with semantic category, mediated by the grammatical realization each category typically takes. Nominal role-metaphors (`Human role analogy`, `Biological metaphors`) are predicted to show low distancing marking despite high frequency. Verbal cognition-metaphors, notably `Cognizer`, are predicted to retain high marking and negation despite high frequency. The prediction is therefore not a uniform correlation between category and marking, but a structural split: nominal-dominant categories should pattern with

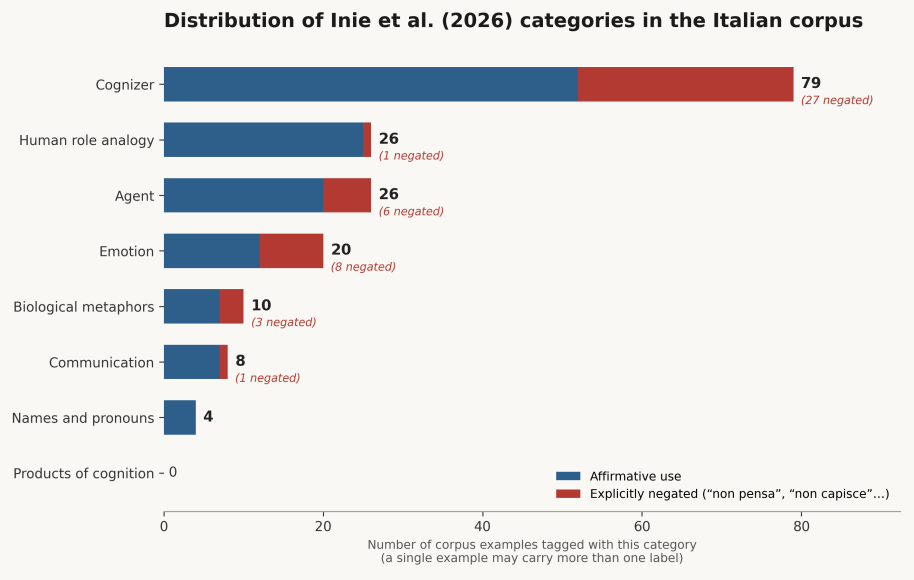


Fig. 1 Distribution of corpus examples across the Inie et al. (2026) categories

H1's frequency-driven terminologization, while verbal-dominant categories should systematically violate it.

6 Data Analysis

6.1 Distribution of Examples across the Inie et al. Categories

The Inie et al. (2026) taxonomy proved broadly applicable to the Italian corpus, although not every occurrence could be readily accommodated within one of the eight categories. Multi-labeling occurred for 11 occurrences. In two examples, the human, rather than the AI system, was the target of the relevant cognitive or affective process (*proviamo empatia verso le intelligenze artificiali*, 'we feel empathy towards artificial intelligences'; *rischio di amnesia digitale per chi delega la memoria all'IA*, 'risk of digital amnesia for those who delegate memory to AI'). These two examples were retained in the dataset as evidence that the anthropomorphic framework can run in either direction. Although too few to support generalization, this reversed directionality may warrant separate investigation as a potential cross-linguistic difference.

`Cognizer` accounts for roughly half of all category assignments (79/173), confirming that anthropomorphic language about AI in Italian is overwhelmingly cognitive (*capire* 'to understand', *pensare* 'to think', *imparare* 'to learn', *ragionare* 'to reason', *decidere* 'to decide', *sapere* 'to know'). This pattern holds across frequency, register, and year. `Human role analogy` ($n = 26$, 1 negated) and `Agent` ($n = 26$, 6 negated) form the next tier. Human-role nouns such as *assistente* 'assistant', *collega* 'colleague', *partner*, and *tutor* are poorly marked, whereas the cognitive-verb category is simultaneously the largest and the most heavily negated (27 of 79 instances). The same asymmetry holds for distancing markers more broadly: `Human role analogy` is poorly marked, while `Cognizer` shows the highest incidence of distancing markers of any category. `Names and pronouns` ($n = 4$) is marginal, and `Products of cognition` is an empty category in the Italian dataset (see Figure 1).

Inie et al. (2026) do not report a category-frequency breakdown for English, so a direct cross-linguistic comparison of category distribution is not possible. What can be stated with more confidence is structural: a taxonomy developed for English maps onto the Italian material without requiring new categories for seven of its eight slots. This indicates that the underlying anthropomorphic frames are stable across these two languages.

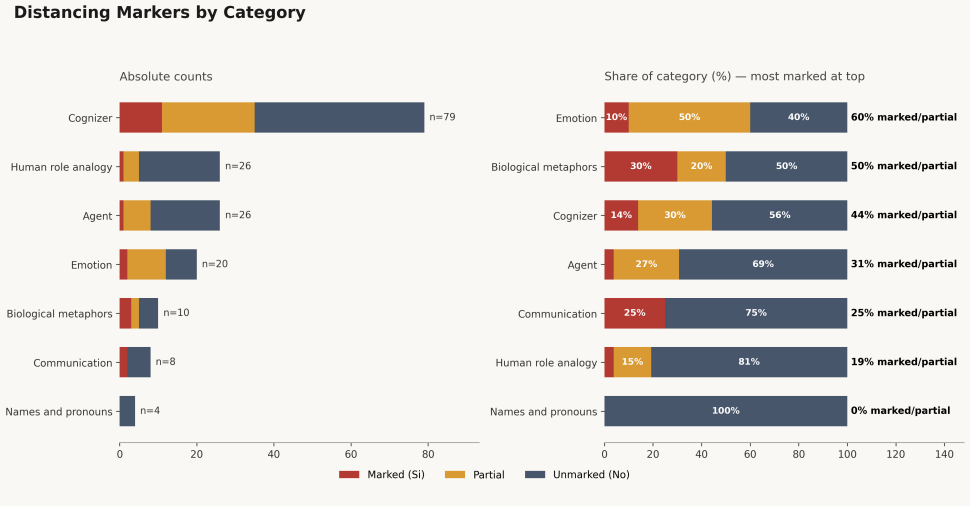


Fig. 2 Distancing markers by Inie et al. (2026) category

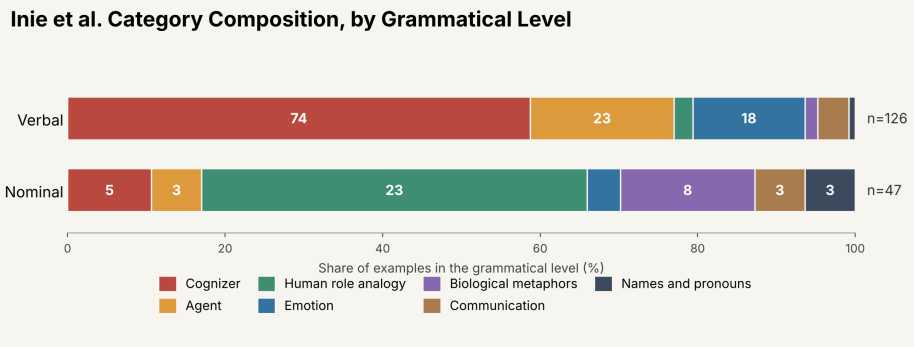


Fig. 3 Category distribution by Grammatical level

6.2 Distancing Markers

Distancing markers were first of all examined by semantic category and by grammatical level to test the structural asymmetry between nominal and verbal anthropomorphic expressions.

By category (see Figure 2), the ranking of marking rates does not follow raw frequency. `Emotion` (60% marked) and `Biological metaphors` (50%) are the most hedged categories despite mid-range frequency. `Cognizer` sits at 44%, despite being by far the largest category ($n = 79$); `Human role analogy` is marked in only 19% of cases ($n = 26$) despite being high-frequency. This mirrors the asymmetry found at the lemma level: nominal role-metaphors such as *collega* 'colleague' or *assistente* 'assistant' pass largely without marking, while attributing emotion or cognition to a system continues to draw quotation marks, hedges, or explicit negation. Emotion attribution appears to be the most cautious category in current Italian discourse around AI, even more so than cognition itself.

By grammatical level, as illustrated in Figure 3 and Figure 4, the split is structurally clean. Verbal forms ($n = 126$) are overwhelmingly `Cognizer` (74) and `Agent` (23), as well as cognition and autonomy predicates (*pensa* 'thinks', *impara* 'learns', *decide* 'decides', *agisce* 'acts'). Nominal forms ($n = 47$) lean toward `Human role analogy` (23) and `Biological metaphors` (8): *assistente*, *collega*, *cervello* 'brain'. Verbal-dominant categories, `Cognizer` above all, and to some extent `Emotion`, remain in a state of contested resistance: marking and negation stay high despite, and indeed because of, high frequency, since their referent, a machine *thinking* or *understanding*, remains the object of an ongoing public and epistemic debate that overrides the standard lexical-entrenchment mechanism predicted by H1. Nominal-dominant categories such as `Human role analogy` and `Biological metaphors` follow a predicted trajectory: high frequency and less marking, consistent with a metaphor that has become invisible. The grammatical gender system of the Italian language may further facilitate this tendency for nominal categories specifically

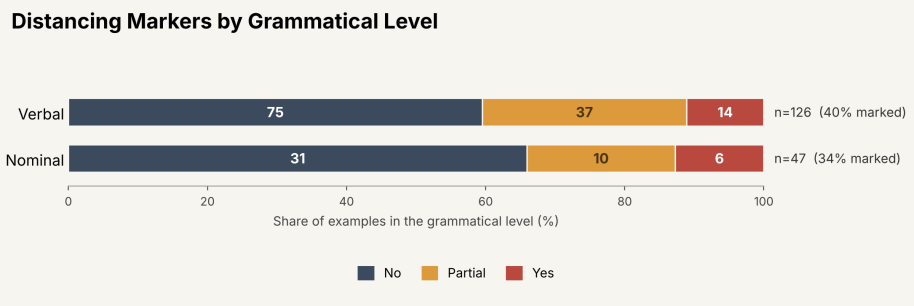


Fig. 4 Distancing markers by Grammatical level

Table 3 Examples of diachronic marking change for `Communication` (*conversare, conversazione, agente conversazionale*) and `Human role analogy` (*tutor*)

ID	Expression (gloss)	Source	Year	Marker Level	Distancing Marker
Communication (*conversare* 'to converse', *conversazione* 'conversation', *agente conversazionale* 'conversational agent')					
38	*con Alexa o Siri non è possibile avere una "conversazione"* 'with Alexa or Siri it is not possible to have a "conversation"'	SimplyScience	2023	Yes	quotation marks + explicit negation
36	*il chatbot conversa con l'utente interpretandone i bisogni* 'the chatbot converses with the user, interpreting their needs'	Il Messaggero	2025	No	unhedged, conventionalized term
82	*il chatbot [è] definito come agente conversazionale che dialoga con il suo utente* 'the chatbot [is] defined as a conversational agent that dialogues with its user'	Senato della Repubblica (MinervaWeb glossary)	2026	No	institutional glossary definition; strong terminologization
Human role analogy (*tutor*)					
74	*il Ministero sperimenta l'IA come una sorta di tutor nelle scuole* 'the Ministry is testing AI as a sort of tutor in schools'	Today.it	2024	Partial	explicit hedge (*una sorta di*)
75	*il tutor virtuale è la voce che sblocca altre voci 'the virtual tutor is the voice that unlocks other voices'*	Orizzonte Scuola Notizie	2025	No	unhedged, full technical label

(see Section 4). It seems therefore that authors reflect explicitly on a term mainly when a cognition verb, rather than a role noun, is at stake.

The relative proportion of distancing markers per year was examined as an indicator of how terminologization within the domain changes over time. H2 predicts a decrease in marking over time, as terminologization progresses. Some lemmas are consistent with this prediction. For the semantic category `Communication` (*conversare* 'to converse', *agente conversazionale* 'conversational agent') and `Human role analogy` (*tutor*), for example (see Table 3), more disclosed uses, such as a 2023 popular-science source explicitly denying that a *chatbot* can have a *conversazione* 'conversation', or a 2024 source hedging *una sorta di tutor* 'a sort of tutor', give way to later, unmarked uses of the same term as a plain technical label, such as a 2025 institutional glossary entry defining a *chatbot* as an *agente conversazionale* 'conversational agent' with no distancing marker at all.

At the aggregate level, however, the dataset shows an opposite trajectory (see Figure 5). Distancing markers show a marked increase overall between 2023 and 2026, driven primarily by the *Partial* category rather than by explicit (*Yes*) marking, which remains essentially flat (17% → 7% → 11% → 10%). Unmarked (*No*) occurrences fall from 83% in 2023 to 34% in 2026, with roughly half of that decline occurring in a single year (73% unmarked in 2025, 34% in 2026), and the entire shift is carried by the *Partial* category (0% →13% → 16% → 56%).

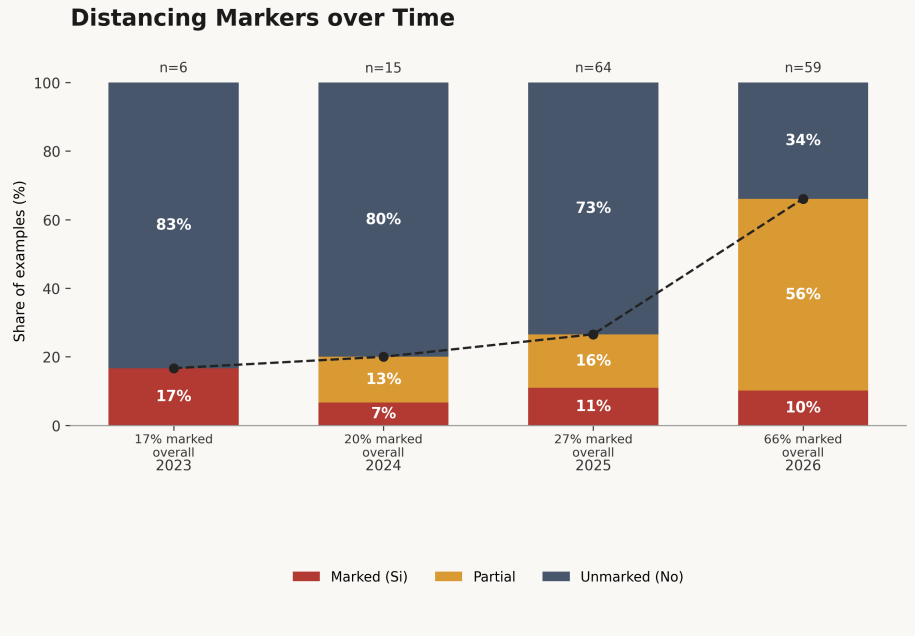


Fig. 5 Distancing markers over time — relative frequency

Two explanations for this counter-hypothesis increase were considered. The first is a genre-composition effect: the 2025–2026 corpus contains substantially more explicitly metalinguistic or critical texts about anthropomorphism than earlier years ($n = 6$ in 2023 versus $n = 43$ in 2026); under this account, the rise in marking reflects which sources are sampled rather than a change in how any single use is perceived. The second is a genuine stiffening of Italian public debate on AI over the same period, under which the rise reflects an active counter-movement against silent terminologization at the level of individual lemmas. What drives the shift is not the marked/unmarked (*Yes*/*No*) dichotomy but the rise of *Partial*, since in 2023–2024 the overwhelming majority of instances were full, unqualified uses carrying no cautionary marker, whereas by 2026 the majority instead carry an implicit caution conveyed through content, such as hedging or negation, rather than through explicit typographic marking. As a robustness check, the same 2025–2026 rise in cautionary marking appears when the data are split by verbs and nouns: both show an increase, verbs more markedly. This favors a genuine stiffening of the debate over a shift confined to nominal role-metaphors (see Figure 6). The way the corpus was built, through tuples and seeds designed to minimize bias toward any specific genre, lends some further support to this second explanation.

6.3 Frequency

H1 predicts that higher-frequency lemmas should show a lower marking rate than singletons. Three anthropomorphic terms borrowed from or modeled on English illustrate this at the level of individual lemmas: *deep learning*, *prompt*, and *thinking model*. The first two have long been established in Italian AI discourse; *thinking model* entered scientific discourse only in January 2025. In the dataset, *deep learning* and *prompt* occur six and five times respectively, whereas *thinking model* occurs only once. The three expressions also differ in marking: *deep learning* and *prompt* carry no typographic distancing, while *thinking model* is italicized and glossed with an explanatory reformulation, *modelli "pensanti"* 'thinking models'. The first two show a higher degree of terminological entrenchment; *thinking model*'s explicit marking and gloss point instead to the cautious presentation of a less established term in Italian AI discourse (Resi, 2026).

However, at the corpus-aggregate level, the opposite holds, and the difference is statistically significant. Using 160 usable examples (162 minus 2 reverse-direction rows) across 50 atomic lemmas, singleton lemmas are marked 27% of the time; the rate remains flat through the 2–3 and 4–6 occurrence bands; lemmas occurring seven or more times rise to 51% marking. A Mann-Whitney test confirms the difference is not attributable to noise: marked examples come from lemmas with a median corpus frequency of 11, against

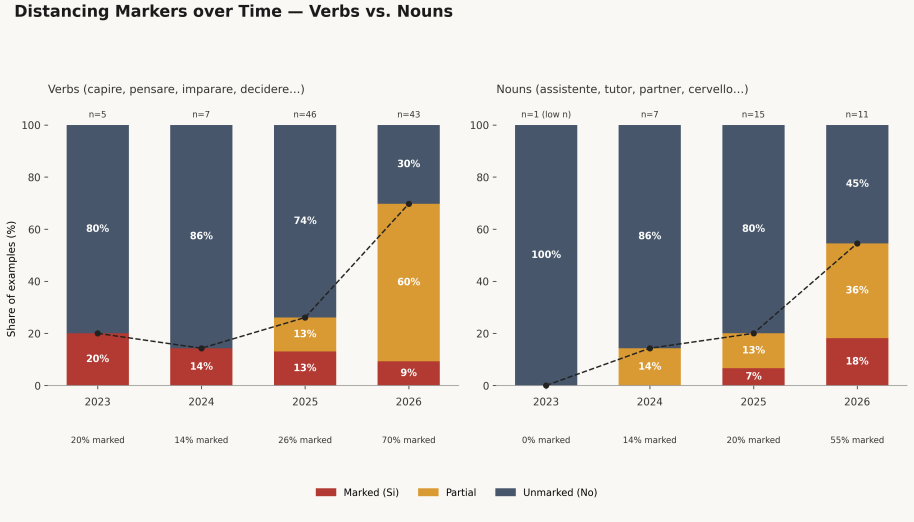


Fig. 6 Distancing markers over time — verbs vs. nouns

Table 4 Distancing markers by source register ($n = 160$)

Register*	n	Yes	Partial	No	% marked
Institutional/technical	20	3	6	11	45.0%
Journalistic/opinion	117	13	36	68	41.9%
Corporate/marketing	23	2	3	18	21.7%

*Institutional/technical: glossaries, institutional portals, research centres, theses, encyclopaedias. Journalistic/opinion: newspapers, magazines, professional blogs. Corporate/marketing: company blogs.

4 for unmarked examples ($U = 3646$, $p = 0.004$). The lemmas driving this pattern are the canonical verbs of debate, *capire* (24×), *provare emozioni* 'to feel emotions' (15×), *imparare* (13×), *pensare* (11×), *decidere* (11×), with marking rates of 36–67%. Low-frequency human-role nouns such as *assistente*, by contrast, sit at 0% marked despite being genuine anthropomorphic instances.

This finding is robust: repeated lemmas predict *more* marking, not less. The most plausible explanation is that the same verbs (*capire*, *pensare*, *provare emozioni*, *decidere*) are frequent precisely because they are central to the debate on anthropomorphism. Being contested drives both their repetition and their hedging. Raw corpus frequency is therefore ruled out as a general terminologization proxy. The next section instead separates genre from frequency, with a register-stratified analysis.

6.4 Register injection

H4 predicts lower marking in institutional/technical sources than in journalistic ones. In institutional sources such as the Senato della Repubblica's Minerva, *chatbot* is used as a fully established term accompanied by a definition and without any distancing marker, in contrast with journalistic and blog sources from the same period (2025–2026), which continue to signal distance from the same lemma through quotation marks or hedging devices.

At the dataset level, however, H4 does not hold. As Table 4 shows, institutional/technical sources have a marking rate of 45.0%, journalistic sources 41.9%, a difference that is not statistically significant ($p = 0.81$). Corporate/marketing sources show a markedly lower rate (21.7%; versus journalistic, $p = 0.10$, also not significant), consistent with promotional language avoiding hedging for reasons of hype rather than genuine terminologization. Genuinely institutional sources thus do not show the near-invisibility of the anthropomorphic term that would be expected of a fully terminologized technical register: the diachronic shift toward more hedged usage documented in Section 6.2 appears to reflect a broader change in Italian public discourse on AI, rather than one confined to journalism.

7 Conclusions

The convergence between Inie et al. (2026)'s categories and the Italian categories found in the dataset points to a pattern of anthropomorphic conceptualization in AI discourse, even as its specific linguistic realizations in the two languages vary. Transparency of anthropomorphic terms depends on the balance between a metaphor's residual figurative visibility and its degree of terminologization. Italian specialized discourse's extensive use of distancing markers was taken as proxy for this balance, alongside frequency, diachronic trajectory, and register injection. The dataset analysis allows the following conclusions.

Terminologization is not uniform across the dataset: it proceeds along two distinct tracks. `Human role analogy` (*assistente* 'assistant', *collega* 'colleague') seems to terminologize as H1 predicts: high frequency, low marking, the metaphor remains invisible. `Cognizer` (*capire* 'to understand', *pensare* 'to think') does not. Despite being the most frequent category, it remains marked and/or negated. Rather than standardizing through use, it is held open by an extra-linguistic driver: the live public debate on machine capabilities, expressed through human cognition metaphors, overrides ordinary psycholinguistic entrenchment. These two categories align with the noun/verb dichotomy since the majority of `Human role analogy` is produced through nominal expressions while most `Cognizer` occurrences are verbs. Noun metaphors may terminologize more easily for two reasons: Italian grammar favors personification, and terminology itself is predominantly nominal. A further related factor is borrowing from English, which affects nouns but not verbs: borrowing can obscure the source domain, reduce perceived figurativity, and so make distancing markers less necessary. *Deep learning* for instance, is now firmly established as a term precisely because it does not activate the semantic transparency that the Italian *apprendimento* 'learning' or *imparare* 'to learn' would. By weakening the connection to the source domain, borrowed terms may terminologize faster than their Italian equivalents. Verbs, by contrast, rely on dynamic conceptualization. Distinguishing a general-language verb from a specialized term requires examining the established entity-verb relation, which sets the relevant functional semantic boundaries, a step that matters most while a meaning extension from general to specialized use is still in progress. This dynamic conceptualization can slow down the definition of the concept boundaries that terminologization needs. And because the metaphor stays more evident, verbs are more easily contested in the current Italian debate on AI anthropomorphism, a debate that may in turn slow their terminologization even further. Among all variables tested (H1 frequency, H2 diachronic trajectory, H3 grammatical level, H4 register injection, H5 semantic category), only semantic category produces a clean separation in the direction a terminologization proxy would predict. Semantic category is therefore the real axis along which terminologization operates in this corpus. It bridges the two annotation levels, specifically semantic category (Level 1) and grammatical level (Level 2), which converge on a single structural pattern: nominal role-metaphors terminologize through ordinary, silent, frequency-driven entrenchment, while verbal cognition-metaphors are held open by the live public debate on machine consciousness.

verbal ↔ `Cognizer` ↔ contested
nominal ↔ `Human role analogy` ↔ conventionalized

This reframes the central question of the research. The question is not whether terminologization of anthropomorphic terms is occurring in Italian AI discourse, but which track a given term follows, and it is Inie et al. (2026)'s categories, via their typical grammatical realization, that predict the answer.

Diverging evidence (lemma level vs. corpus-aggregate level): The analysis of distancing markers as a proxy for terminologization supports two distinct conclusions, one at the lemma level and one at the corpus-aggregate level. At the lemma level, where a single term is investigated over time, the dataset offers three pieces of evidence:

1. For H1 (frequency): marking declines as a specific term increases in frequency as it is the case with *prompt* with respect to *thinking model*;
2. For H2 (diachronic trajectory): marking decreases over time like it is the case with *tutor* and *conversare* 'to converse'.
3. For H4 (register injection) marking declines as terms settle into institutional use, as it is the case with *chatbot*.

At the corpus-aggregate level, however, the three hypotheses are contested and need reinterpretation. In fact, frequency correlates positively, not negatively, with marking; the diachronic trajectory runs opposite to what was expected; and institutional register is marked as much as journalistic register. Marking increases from 2023 to 2026, driven almost entirely by the *Partial* category. This means that fewer corpus occurrences are now entirely unmarked. The register as a whole has grown more cautious through implicit marking, especially from 2025 onward.

Two explanations can account for this divergence: (1) a genre-composition shift: the 2025–2026 sample might simply contain substantially more explicitly metalinguistic or critical text than earlier years, so that rising marking reflects which sources were sampled rather than a real change in how any single use is perceived. Because the sampling method selected text through tuples and seeds designed to minimize genre bias, this explanation is less likely; (2) a genuine stiffening of the debate: a real counter-movement against terminologization of anthropomorphism is under way, visible as rising distancing markers. This fits the fact that the rise is carried by *Partial* marking (implicit, content-based caution) rather than by explicit *Yes* marking. The aggregate trend should therefore be reported with explicit caution, as possible evidence of a brake on the terminologization of anthropomorphic metaphors, driven by increasing metalinguistic reflection, and not as a claim that individually unmarked terms are becoming marked again.

Reconciling the two levels: It is important to keep the aggregate-level finding separate from the lemma-level analysis. Some individual terms have indeed become unmarked, evidence of their own terminologization. But an overall tendency toward terminologization of anthropomorphic metaphors as a phenomenon has been held back by critical discourse about the language itself. Authors who never explicitly mention terminologization, or the standardization of a term, still felt the need to distance themselves from this way of talking about AI. Within this debate, the same terminology is used both to reflect critically and to be negated. So even as individual terms (mainly nominal) quietly terminologize, an explicit "anti-anthropomorphism" genre (blog posts, opinion pieces) has grown in parallel and appears to slow the overall terminologization of anthropomorphic metaphor by keeping the metaphorical status of candidate terms visible.

Limitations and directions for future research: A further diagnostic would include isolating grammatical level using domain-matched verb/noun pairs drawn from the same conceptual metaphor, e.g. *assistere* (verb) vs. *assistente* (noun), or *pensare* (verb) vs. *pensatore*/*pensiero* (noun), and comparing distancing rates within each pair. Attesting the presence of clusters for Italian anthropomorphic expression (e.g., *apprendere* 'to learn' surfacing also as *apprendimento* 'learning', *sistema che apprende* 'system that learns', *capacità di apprendimento* 'learning capacity') can also represent a complementary indicator that the terminologization process is taking place. These observations call for further work. Moreover, it would be interesting to investigate whether the categorization of anthropomorphism can be extended cross-linguistically, and which proxies for terminologization are relevant in other languages.

Declaration on Generative AI

During the preparation of this work, GPT-5.6 Sol was used in order to check grammar and spelling as well as to check reference consistency. GPT-5.6 and Copilot were also used to support the Python script for automatic annotation tests and compute the dataset. After using this service, the author carefully reviewed and edited the output as needed and takes full responsibility for the content of the paper.

References

Baroni, M., Bernardini, S., Ferraresi, A., Zanchetta, E. (2009). The WaCky wide web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, *43*(3), 209–226, https://doi.org/10.1007/s10579-009-9081-4

BootCaT (2004). *BootCaT: Simple utilities to bootstrap corpora and terms from the web.* https://bootcat.dipintra.it/?section=home.

Bowdle, B.F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216, https://doi.org/10.1037/0033-295X.112.1.193

Bunz, M., & Braghieri, M. (2022). The AI doctor will see you now: Assessing the framing of AI in news coverage. *AI & Society*, *37*, 9–22, https://doi.org/10.1007/s00146-021-01145-9

Calaresu, E. (2025). Significati non letterali nel discorso scientifico e accademico. Metafore, metonimie e personificazione di referenti inanimati. *Italiano LinguaDue*, *17*(2), 143-166, https://doi.org/10.54103/2037-3597/30540

Costa, R., Ramos, M., Salgado, A., Carvalho, S., Almeida, B., Silva, R. (2022). Neoterm or neologism? A closer look at the determinologisation process. A. Klosa-Kückelhaus & I. Kernerman (Eds.), *Lexicography of Coronavirus-related neologisms* (pp. 237–260). De Gruyter. https://doi.org/10.1515/9783110798081-012

Epley, N., Waytz, A., Cacioppo, J.T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886, https://doi.org/10.1037/0033-295X.114.4.864

Gentner, D., & Jeziorski, M. (1993). The shift from metaphor to analogy in Western science. A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 447–480). Cambridge University Press. https://doi.org/10.1017/CBO9781139173865.022

Henzinger, M.R., Heydon, A., Mitzenmacher, M., Najork, M. (1999). Measuring index quality using random walks on the Web. *Computer Networks*, *31*(11-16), 1291–1303, https://doi.org/10.1016/S1389-1286(99)00016-X

Humbert-Droz, J. (2023). When terms become neologisms: A contribution to the study of neology from the perspective of determinologisation. *CLINA: Revista Interdisciplinaria de Traducción Interpretación Y Comunicación Intercultural*, *9*(1), 135–158, https://doi.org/10.14201/clina202391135158

Ibrahim, L., & Cheng, M. (2025). Thinking beyond the anthropomorphic paradigm benefits LLM research. *arXiv* https://doi.org/10.48550/arXiv.2502.09192

Infoterm (2005). *Guidelines for terminology policies: Formulating and implementing terminology policy in language communities.* UNESCO.

Inie, N., Druga, S., Zukerman, P., Bender, E.M. (2024). From "AI" to probabilistic automation: How does anthropomorphization of technical systems descriptions influence trust? *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)* (pp. 2322–2347). Association for Computing Machinery. https://doi.org/10.1145/3630106.3659040

Inie, N., Zukerman, P., Bender, E.M. (2026). De-anthropomorphizing "AI": From wishful mnemonics to accurate nomenclature. *First Monday*, *31*(2) https://doi.org/10.5210/fm.v31i2.14366

International Organization for Standardization (2023). *Management of terminology resources — Terminology databases — Part 3: Content (ISO Standard No. 26162-3:2023).* Retrieved from https://www.iso.org/standard/80464.html

Ischreyt, H. (1965). *Studien zum verhältnis von sprache und technik: Institutionelle sprachlenkung in der terminologie der technik* (Vol. 4). Schwann.

Kageura, K., & Umino, B. (1996). Methods of automatic term recognition: A review. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, *3*(2), 259–289, https://doi.org/10.1075/term.3.2.03kag

Kempton, A.M., Parmiggiani, E., Vassilakopoulou, P. (2023). Accountability in managing artificial intelligence: State of the art and a way forward for information systems research. *Proceedings of the 31st European Conference on Information Systems (ECIS 2023).* Retrieved from https://aisel.aisnet.org/ecis2023_rp/361

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* University of Chicago Press.

Lipton, Z.C. (2018). The mythos of model interpretability. *Communications of the ACM*, *61*(10), 36–43, https://doi.org/10.1145/3233231

Meyer, I., & Mackintosh, K. (2000). When terms move into our everyday lives: An overview of determinologization. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, *6*(1), 111–138, https://doi.org/10.1075/term.6.1.07mey

Müller, C. (2008). *Metaphors dead and alive, sleeping and waking: A dynamic view.* University of Chicago Press. https://doi.org/10.7208/9780226548265

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103, https://doi.org/10.1111/0022-4537.00153

Pascolini, A. (2004). Metafore e comunicazione scientifica. *JCOM: Journal of Science Communication*, *3*(1), 1–18, https://doi.org/10.22323/2.03010301

Petricini, T. (2026). The power of language: Framing AI as an assistant, collaborator, or transformative force in cultural discourse. *AI & Society*, *41*(2), 1005–1017, https://doi.org/10.1007/s00146-025-02586-2

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., … Wellman, M. (2019). Machine behaviour. *Nature*, *568*, 477–486, https://doi.org/10.1038/s41586-019-1138-y

Resi, R. (2026). Measuring anthropomorphism in Italian specialized language on AI. F. Vezzani, G.M. Di Nunzio, L. Grčić, A. Ostroški Anić, & V. Bonato (Eds.), *Proceedings of the 5th International Conference on Multilingual Digital Terminology Today (MDTT 2026) Zadar, Croatia, June 25-26, 2026.* CEUR-WS.ORG. Retrieved from https://ceur-ws.org/Vol-4219/paper5.pdf

Resi, R., & ten Hacken, P. (in press). Perdita di domini: Is Italian losing domains? M. Thelen, M.-C. Wermuth, & L. Van Vaerenbergh (Eds.), *Domain loss and gain in contemporary terminology studies.* Taylor & Francis.

Ryazanov, I., Öhman, C., Björklund, J. (2025). How ChatGPT changed the media's narratives on AI: A semi-automated narrative analysis through frame semantics. *Minds and Machines*, *35*(2), 1-24, https://doi.org/10.1007/s11023-024-09705-w

Salles, A., Evers, K., Farisco, M. (2020). Anthropomorphism in AI. *AJOB Neuroscience*, *11*(2), 88–95, https://doi.org/10.1080/21507740.2020.1740350

Shanahan, M. (2023). Talking about large language models. *arXiv* https://doi.org/10.48550/arXiv.2212.03551

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *arXiv* https://doi.org/10.48550/arXiv.2002.04087

Stanič, D. (2025). The metaphorical conceptualisation of artificial intelligence: Analysis of contemporary Italian journalistic discourse. *Linguistica*, *65*(1), 185–203, https://doi.org/10.4312/linguistica.65.1.185-203

Temmerman, R. (2000). *Towards new ways of terminology description: The sociocognitive approach.* John Benjamins Publishing Company. https://doi.org/10.1075/tlrp.3

ten Hacken, P., & Panocová, R. (2024). *Word formation as a naming device.* Edinburgh University Press. https://doi.org/10.3366/edinburgh/9781474487016.001.0001

Tiku, N. (2022). *The Google engineer who thinks the company's AI has come to life.* The Washington

Post. Retrieved from https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Gerstenberg, T., Bernstein, M.S., Krishna, R. (2023). Explanations can reduce overreliance on AI systems during decision-making. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW1), 1–38, https://doi.org/10.1145/3579605

Vernillo, P. (2024). MIP-IT: Identificazione delle metafore linguistiche in italiano. Una proposta preliminare tra questioni teoriche e problemi pratici. *Annali Del Dipartimento Di Studi Letterari, Linguistici E Comparati. Sezione Linguistica*, *12*, 345-382, https://doi.org/10.6093/2281-6585/10707

Watson, D. (2019). The rhetoric and reality of anthropomorphism in artificial intelligence. *Minds and Machines*, *29*(3), 417–440, https://doi.org/10.1007/s11023-019-09506-6

Waytz, A., Morewedge, C.K., Epley, N., Monteleone, G., Gao, J.-H., Cacioppo, J.T. (2010). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435, https://doi.org/10.1037/a0020240

Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45, https://doi.org/10.1145/365153.365168

Wüster, E. (1979). *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Springer.

Ye, Z., & Li, J. (2024). Artificial intelligence through the lens of metaphor: Analyzing the EU AIA. *International Journal of Digital Law and Governance*, *1*(2), 361–381, https://doi.org/10.1515/ijdlg-2024-0016