Risk Communication in Healthcare: The Management of Misunderstandings

Monica Consolandi[1,*,†], Simone Magnolini[1,*,†] and Mauro Dragoni[1]

[1]Fondazione Bruno Kessler, Trento, Italy

Abstract

Risk communication represents a nuanced discourse within the healthcare sector, characterized by the topics' sensitivity and the potential for misunderstandings between healthcare providers and patients. This delicacy stems from the complexity of effectively conveying information about risks. Consequently, a primary obstacle lies in fostering healthcare providers' understanding of implicit communication nuances inherent in pre-operative risk discussions. This study aims to address this gap in the literature by examining the topic through the lens of the philosophy of language, specifically utilizing pragmatic analysis tools to elucidate implicit understandings in doctor-patient interactions. We employ this approach to scrutinize instances of risk evaluation preceding cardiac surgery. Through empirical analysis of gathered data, we illustrate the inadequacies of current state-of-the-art models in accurately identifying misunderstandings within healthcare dialogues. In conclusion, we propose avenues for future research in this domain, underscoring the importance of further exploration into improving risk communication in healthcare settings.

Keywords

Risk Communication, Misunderstanding, Dialogues, Digital Health

1. Introduction

Risk communication represents a highly contentious and extensively discussed topic within the realm of doctor-patient interactions. [1]. Although medicine operates on the principles of evidence-based practice [2, 3], it remains inherently imprecise, characterized by uncertainty and probabilistic reasoning [4, 5, 6, 7]. In formal doctor-patient interactions, physicians are tasked with imparting information to patients regarding treatment success rates, potential complications, and associated side effects. This communication inherently entails uncertainty, as medical information primarily relies on probabilities rather than certainties. Medicine inherently involves risks, such as the possibility of treatment ineffectiveness, surgical complications, and post-operative recovery complications. [8, 9]. The risk ratio must always be considered when assessing the patient's best interests, with greater emphasis placed on communicating risks that carry higher significance. For instance, when comparing treatment options for a rash versus a surgical procedure, understanding the risk balance in the latter is paramount. Various efforts have been made to categorize uncertainty based on medical risks. Citing [10] claim that uncertainty derives from source, "incomplete information, inadequate understanding, or undifferentiated alternatives of equal attractiveness", or issue, "the particular outcomes, situation, or alternatives to which a given uncertainty applies" [10, 11]. Weinfurt adds to Miller and Joffes' categories of uncertainty, "the causal agent and the validity of surrogate endpoint", "uncertainty about generalizability", meaning "using data from past studies to predict what will happen in a new" one; the "estimation error" associated with statistics; and the fact that any estimation is "characteristic of a population, not of an individual" [12]. Discussing benefits in phase 1 oncology trials, Miller and Joffe conclude "in order to facilitate informed consent, it is vital that clinical investigators communicate meaningfully to patients regarding the probability, magnitude, and duration of potential benefits and risks, along with their attendant uncertainties" [13]. As highlighted by Weinfurt, this matter is not straightforward, primarily because it necessitates physicians to possess a comprehensive understanding of both the benefits and risks associated with medical interventions, a

HC@AIxIA 2024: 3rd AIxIA Workshop on Artificial Intelligence For Healthcare, November 25–28, 2024, Bolzano, Italy

*Corresponding author.

†These authors contributed equally.

mconsolandi@fbk.eu (M. Consolandi); magnolini@fbk.eu (S. Magnolini); dragoni@fbk.eu (M. Dragoni)

task complicated by inherent uncertainties. Additionally, patients must comprehend the information disclosed by their doctor (see [14] study about this topic). This second point holds particular significance within our analytical framework because it encompasses not only the content of communication but also the manner in which it is conveyed. Effectively communicating risks to patients demands specific skills, given the sensitivity of the moment; patients are confronted with potential risks, including the risk of mortality. This aspect is of considerable interest to psychologists, as it delves into the emotional dimensions of the interaction: communicating properly means not being traumatizing [15], and good communication also helps build a relationship of trust between the doctor, the patient [16, 17], and the family [18]. Some psychological studies show a direct connection between the communication of risk and the level of trust in the doctor-patient relationship [19].Some articles talk about risk communication in a broader discourse, particularly as part of obtaining informed consent [20], patient's rights [21], and the physician's duty to tell the truth [22]. These issues have also been investigated by philosophers, specifically in clinical ethics and bioethics [23].There are also relevant applications of the philosophy of language analysis to this topic [24, 25]. As previously mentioned, a philosophical analysis of language delves into the ramifications of verbal expressions. Such analysis could aid in discerning the types of implicature that prove beneficial versus those that may be detrimental [26]. Applied to doctor-patient interactions during risk explanation, this analysis has the potential to enhance speakers' awareness, thereby improving doctor-patient communication overall.

In this paper, we aim to present the challenge of detecting misunderstandings during risk communication dialogues in healthcare. This research topic has generally received inadequate attention in the literature, primarily due to limited available resources and the inherent complexity of the task. [27, 28, 29]. Firstly, we will discuss such a challenge from a philosophical perspective by analyzing the implicit dimension of dialogues and trust. Then, we will detail some use cases extracted from the dataset we used to perform a preliminary classification of the text with respect to the different types of misunderstandings we identified in the analyzed dialogues. The set of possible types of misunderstanding used to annotate our data followed the codebook described in [30] that specifically addresses the topic of risk communication in healthcare. In conclusion, we assess the capabilities of state-of-the-art Large Language Models (LLMs) such as GPT-4 developed by OpenAI in recognizing and analyzing misunderstandings through a qualitative examination of a single dialogue.

2. The Implicit Dimension

The implicit dimension of discourse harbors rich meaning, and philosophers of language are dedicated to elucidating it. Grice is the first to call the implicit dimensions of the discourse 'implicatures' [31]. According to the Stanford Encyclopedia of Philosophy, "implicature denotes either (i) the act of meaning or implying one thing by saying something else, or (ii) the object of that act". When posing the question "Do you know what time it is?" one anticipates responses indicating the actual time, such as "Noon" or "It's late", rather than a simple "Yes". Implicitly, this question may convey various meanings: firstly, it may signify the asker's lack of knowledge regarding the current time and their query to ascertain it from the respondent; alternatively, it could imply a desire to gauge whether one is punctual, with a response of "No" suggesting tardiness. Despite the question's uniform wording, it carries nuanced implications depending on context and interpretation. As evidenced by this example, Gricean implicatures play a functional role: it is not solely the explicit content of our utterances but also the implicit information conveyed through what is left unsaid that contributes to the effectiveness of a conversation. We can distinguish two primary types of implicatures: (i) those that are inevitable and (ii) those stemming from the speaker's deliberate choices. Both types are context-dependent and indispensable.

The first type (i) comprises implicatures essential for facilitating effective communication. For instance, when a doctor offers a patient "some water," it is understood without explicit specification that the water will be served in a suitable vessel and in a drinkable state. The omission of such details is expected and contributes to the smooth functioning of the conversation. This phenomenon is termed a "convention" or "conventional implicature" between the interlocutors. Nonetheless, we have to keep in

mind that these conventions change from one culture to another, i.e., they work in proper circumstances and do not in others (see the concept of felicity in [32]). This could generate difficult or even amusing circumstances: e.g. if you nod in Greece to show your enthusiasm for having that glass of water, you are actually refusing it (see the concept of infelicity in [32]). There are also more specific conventions in communication. For instance, when a patient asks her doctor about "Luke," it is understood without explicit clarification that she is referring to the doctor's son, husband's child, or the brother of the doctor's daughter. It is evident why communication would become impractical if every detail needed explicit explanation.

The second kind of implicature is more intricate, as it does not arise from spontaneous simplification but rather from deliberate choices made by the speakers. This type of implicature depends on factors such as obviousness, complexity, and temporal context. For example, when a dermatologist instructs a patient to "bring me your exams tomorrow," the doctor does not specify the time, location, or nature of the exams. The implicatures of the dermatologist's request imply that the exams should be brought on the following day, during the appointment, and pertain to skin pathology. Unlike conventional implicatures, these are case-specific and termed "conversational implicatures" by Grice. The primary distinction between the two types lies in the subtlety of the second type's implications compared to the clarity of the first. In both cases, what is explicitly stated intertwines with what is left unsaid. Crafting accurate representations of implicatures related to speakers' choices demands more refined analytical tools. Despite Blečić's suggestion [33], we do not think it is possible that "both doctors and patients should avoid conversational implicatures", as implicatures are an unavoidable piece of meaning in communication, while agreeing with her when saying that "both parties could benefit from the ability to detect them". The philosophy of language endeavors to systematize interactions by elucidating implicit meanings, thereby exposing gaps and misunderstandings for correction. By elucidating the implicatures inherent in doctor-patient communication, we can pinpoint specific issues stemming from missed or misunderstood meanings, thus constructing a comprehensive map of the interaction.

Grice posits that the cornerstone of all conversations lies in the cooperative principle, which emphasizes the importance of mutual recognition of a shared goal in facilitating understanding during dialogue. This principle aligns closely with Wittgenstein's perspectives on the subject. Wittgenstein [34] tells us that every speaker has a set of rules for speaking, which he calls a "language game". When engaging in conversation, we participate in a language game, which inherently evokes a sense of unease because each interlocutor brings their own distinct language game to the interaction. Fully comprehending the language game of the other person is impossible, as many of its rules remain unspoken. Similar to the way none of us explicitly articulate the grammatical rules we adhere to while speaking, the same applies to other rules elucidated by Wittgenstein. We can imagine that a third language game has to emerge [35]: this is the set in which two or more different speakers can meet, sharing common rules in a workable compromise. In the context of doctor-patient communication, there exist two primary systems of rules: those governing the behavior and expectations of the doctor, and those governing those of the patient. These systems are intended to converge within the figurative space of the doctor's office. It is imperative for both speakers to be open to mutually accepting the rules inherent in the language games being played during the interaction. Combining Grice's cooperative principle with Wittgenstein's insights, we assert that effective communication necessitates reciprocal trust, where both parties share and accept the rules governing the interaction. In the context of doctor-patient interaction, this foundation of shared rules and reciprocal trust forms the basis of a robust therapeutic alliance. In pursuit of this goal, the speakers have to follow four axioms to make the communication effective, the so-called Gricean maxims. They are: (i) maxim of quality: be truthful; (ii) maxim of quantity: make your contribution as informative as is required, and not more so; (iii) maxim of relation: be relevant; (iv) maxim of manner: be clear, avoiding obscurity, ambiguity, prolixity, and disorder. There are particular cases in which the speakers violate these maxims not to flout the cooperative principle, for example when using metaphors, irony, and hyperbole. I believe that, even in such cases, having in mind the Gricean maxims is helpful to fulfill the interlocutor's expectations and keep in mind the thread of the discussion. We thus obtain a linguistic universe ruled by (i) what is said, and (ii) the implicatures, as even if you can't see them, they are meaningful. If we combine this with the four maxims, we can say that trust is the foundation

of communication, not only regarding the activity of the speakers (maxims) but also the inactivity of the speakers (implicatures). The implicit dimension of communication encompasses the unspoken elements that contribute significantly to the overall meaning of a conversation. Despite its absence in verbalization, it holds substantial meaning and contributes to the holistic understanding of the discourse. It is evident that if the implicit dimension of communication were characterized by mistrust rather than trust, it would pose significant challenges to effective communication. Consider inquiring with your housemate about whether they've locked the apartment. If they respond with a simple "Yes" and you trust them, the conversation typically concludes. However, if doubts linger due to their reputation for being moody or unreliable, you might proceed with further questioning: "Alright, but did you also double-lock it? And did you secure the door with the latch? Have you activated the alarm?"

Applying this understanding of implicit communication and trust to doctor-patient interactions highlights inherent challenges. For instance, envision a scenario where a doctor lacks trust in a patient's adherence to medication: "Have you taken your pills? I mean, every day? And, all five of them? After dinner, as I suggested?" Such an interaction, resembling an interrogation rather than a clinical consultation, underscores the critical role of trust and the complications stemming from its absence. This issue is particularly pronounced during moments of risk communication.

3. Risk Communication

Indeed, effective communication hinges on reciprocal trust, a principle that holds true for doctor-patient interactions and, more broadly, for the doctor-patient relationship [36, 37]. As previously noted, trust plays a vital role in establishing a robust therapeutic alliance. Drawing from the Gricean maxims, we can conceptualize the therapeutic alliance as a virtuous circle where trust and truth reinforce each other. The meaning of the noun trust ("the belief that you can trust someone or something") becomes clearer when looking at the definition of the verb "to trust": "to believe that someone is good and honest and will not harm you, or that something is safe and reliable; to hope and expect that something is true" (according to the Cambridge Dictionary). Trust is usually a three-sided relation: "X trusts Y to do Z". This is translatable in the clinical setting to "P trusts D to do X" and "D trusts P in doing X". Likewise for the meaning of the noun truth ("the quality of being true") with the definitions of its respective adjective true: "sincere or loyal, and likely to continue to be so in difficult situations" and "having all the characteristics necessary to be accurately described as something". As we can notice from the italicized words, there are two main common points in the two definitions: (i) true, sincere, or loyal; and (ii) accurately described. The (i) first cluster is related to the reliability of the relationship, which has to be based on truth to be trustworthy, and vice versa. The (ii) second part is focused on the necessity to describe, i.e. to have arguments about the object of the relationship (e.g. in the clinical setting, a particular treatment). This definition highlights a challenge directly associated with risk communication: it is often neither feasible nor beneficial to provide precise descriptions when discussing open-ended subjects such as risk. On the other hand, certain kinds of omissions could be problematic, too, even though "the desire to establish trust can conflict with the imperative to disclose the whole truth" [38]. We can see how difficult it is to make the virtuous circle work when the truth has to cope with expectations ("...will not harm you..."), and with uncertain communications.

As Wittgenstein emphasized, prior to engaging in communication, both the doctor and the patient possess individual language games, representing distinct systems of rules. The doctor, typically adorned in a white coat and uniform, operates as a professional, documenting patient information in a medical record. On the other hand, the patient typically presents as unwell, often without a uniform, and their symptoms may not always be outwardly apparent. Consequently, the two primary systems of rules in contact are those of a professional healthcare provider and an unwell individual. This juxtaposition carries significant implications: the doctor bears professional responsibility for the patient's life as defined by legal and ethical standards, while the patient retains personal responsibility for their own well-being, relying on the doctor to safeguard it. In this context there are two levels of communication involved: (i) the professional level and (ii) the personal level. The professional level pertains to the

doctor's role as a representative of their medical profession, embodying expertise within the interaction. Here, the doctor bears the responsibility to communicate information regarding the pathology, quality of life, treatments, and associated risks to the patient. Conversely, the personal level encompasses the emotional aspects and personal histories of both parties involved. This is the domain where fears, doubts, and human vulnerability manifest. These two levels are interconnected: on the professional level, the patient assumes a more passive role as the recipient of information, while the doctor takes on an active role as the provider. However, on the personal level, both the patient and the doctor are active participants, engaging with their respective emotional experiences and personal narratives. Indeed, there exists a double asymmetry within the doctor-patient relationship: firstly, the patient is never elevated to the professional level, remaining primarily in the role of recipient of medical expertise; secondly, while both the doctor and patient inhabit the personal level of the relationship, their roles within this sphere are inherently distinct. The first asymmetry is well-documented and arises from the discrepancy in knowledge between the doctor and the patient: the doctor, as the specialist, possesses expertise in the medical field and an understanding of the patient's body that exceeds that of the patient. The second asymmetry underscores the dual role of the doctor: they must maintain their professional identity, adhering to the boundaries established within the doctor-patient relationship, while simultaneously acknowledging their personal involvement and humanity. This juxtaposition is mirrored in the legal realm: instances where professional responsibilities of doctors lead to criminal offenses can result in criminal liability, as exemplified by the Italian law "Gelli-Bianco" n.24/2017. This brief examination underscores a fundamental aspect of the complexity inherent in doctor-patient interactions: it emphasizes that risk communication does not occur in a neutral environment.

3.1. Case Analysis

We present below five cases extracted from our dataset (see Section 4) and we analyze them in more detail aiming to show and discuss how misunderstanding may occur in risk communication and which consequences they may have [39].

CASE 1: DNR order The scenario depicts an anesthesiologist, a woman in her forties, engaging in a discussion with a man in his seventies regarding the prospect of surgery for his aortic aneurysm.

D: They told me you decided with the surgeon.
P: Well, 'decided'... I can only give my willingness; it's you who decides.
D: I can tell you that this surgery is very risky in a small percentage of cases... the problem is in this minimum percentage of cases there would be no indication to take you to the ICU. This means that we put in a tube that we would unlikely take out. When you have self-suspended the therapy... you are a delicate patient in many ways... what I want you to understand is that your risk is to lose a lot of autonomy... It's up to you.

Indeed, the doctor's communication contains implicit information that is not explicitly stated. The crucial piece of information implied here is that "There would be no indication to take you to the ICU," which essentially suggests that in the event of complications, the patient would not undergo resuscitation efforts. This implies that the patient needs to be aware that the mortality risk associated with the surgery is higher than usual for him due to this decision. The doctor chooses an expression that is evocative, "There would be no indication..." instead of saying sharply "You won't be taken to the ICU". But why the decision, and who made it? The doctor uses two key phrases: "You have self-suspended the therapy", and "You are a delicate patient in many ways". The situation presents a curious contrast between the two statements: the first appears to be a reprimand, while the second focuses on patient care. This discrepancy suggests that there is much left unsaid, resulting in potentially confusing communication when conveying a message. The physician seems to be attempting to communicate

something indirectly, without being forthright. Additionally, it's worth noting that the patient in question is delicate and has halted their therapy on their own accord. Patient compliance is crucial for the success of therapy, and when compliance is interrupted, treatment outcomes suffer. Moreover, a patient who is non-compliant poses a challenge to the therapeutic alliance, which relies heavily on reciprocal trust between the doctor and the patient. [40], and a non-compliant patient seems not to be a trustworthy person. The physician here is saying without saying explicitly "How can we still trust you?", and this strong concern is at the same time strengthened and softened by the following: "You're a delicate patient". Again, what does this mean? How do we have to unfold the meaning of the word 'delicate' in this context? Unfortunately, there is no way to reach a satisfying interpretation solely on the basis of the interaction. The solution is in the patient's medical record, in which we find that he has advanced lung cancer. Going back to the opening, we read again "You decided with the surgeon" that now turns out not to be an assertion but rather a questioning of this decision. The combination of two implicit pieces of information, the lack of the patient's compliance and his severe disease makes the "loss of a lot of autonomy" extremely likely, and the prospect of the patient's death closer than ever. The physician's final remark exposes a tension in the clinical decision-making process that extends beyond solely considering the surgical risk: she says "It's up to you" in direct opposition to the patient's "It's you who decides". Indeed, the doctors have already made the decision not to take the patient to the ICU, reflecting a hierarchical structure of knowledge linked to decision-making authority. The anesthesiologist, aware of the implications of her statement, ultimately defers the final decision to the patient. Together with the surgeon, they establish the parameters, and the patient is left with the binary choice of either accepting or declining the proposed course of action. This situation highlights a failure to establish a shared understanding or common ground, resulting in a lack of alignment between the doctors and the patient in terms of their respective language games and decision-making processes.

CASE 2: Death instead of pain? The scenario depicts an elderly man in his eighties engaging in a discussion with the anesthesiologist, who is in her forties, seeking additional information about the risks associated with the removal of an aortic aneurysm. The conversation proves challenging, given the advanced age of the patient and the doctor's uncertainty about the best course of action. The doctor has previously consulted with the surgeon regarding the case. The clinical situation is delicate, as operating on the patient entails significant risks, yet not doing so leaves him with an aortic aneurysm, which poses an imminent threat to his health and well-being.

D: Unfortunately, it's a bet. I don't know you; it depends on your outlook on life. I barely know what I would do if it were me, but I'm not even sure. ...you won't feel pain, this I guarantee you, regardless of whether everything goes well, or you don't wake up .

Despite the many implicatures we can highlight here, there is a phrase with a strong impact: "or you don't wake up". Now that the patient is aware of the potential risk of death associated with the surgery, this possibility is no longer abstract or distant, but rather a tangible and immediate consideration. Even so, if we come back to the beginning of the physician's discourse, we read, "it [the surgery] is a bet". Upon closer examination, it becomes evident that the physician's message extends beyond merely conveying the high risk associated with the surgery. In this context, the metaphor of "placing a bet" or "playing dice" signifies that the physician cannot rely on reliable statistics or a certain percentage of risk to guide her decision-making process. Instead, the implication is that the patient's condition is dire and inherently unpredictable, with surgery being as risky as abstaining from it. By stating that the patient "won't feel pain" if he opts for surgery, the physician aims to offer reassurance amidst the uncertainty. This demonstrates her attentiveness to the emotional dimension of the patient's experience, particularly his underlying fear, which may not be explicitly expressed but is profoundly felt. Furthermore, the physician's shift from a purely clinical standpoint to a more personal perspective underscores the complexity of the situation and her commitment to addressing not only the medical aspects but also the patient's emotional well-being ("I barely know what I would do if it were me") positioning her

perspective on the same level as the patient's one.

P's wife: Is this surgery dangerous?
D: Yes, of course, it is, Maam ... What do you mean by 'dangerous'? What do you want to ask?
P: She's asking about my life.

The scenario presents a man in his sixties, accompanied by his wife, visiting the anesthesiologist's office to schedule a carotid artery revascularization procedure. In the Pulitzer Prize for Drama winner play W;t, we find a passage along the lines of this case. Vivian Bearing, the main character, has been diagnosed with advanced metastatic ovarian cancer; his doctor, Harvey Kelekian, is telling her that the cancer is spreading and that it is an insidious one. "Insidious?" asks Vivian to Dr. Kelekian, who answers, "Insidious means undetectable"; "Insidious means treacherous", says Vivian [41]. Indeed, there is a notable disparity between the language used by the physician, which often includes technical terms and medical jargon, and the language familiar to the patient, who may not possess the same level of medical expertise or understanding of specialized terminology. This discrepancy underscores the challenge of effective communication between healthcare professionals and patients, as the two parties operate within distinct spheres of knowledge and vocabulary (see [42]). Indeed, misunderstandings can occur even with non-technical terms, as demonstrated by the example of the word "dangerous." While this term may seem straightforward and commonly understood to mean something capable of causing harm or unpleasant consequences, its interpretation can vary within the context of risk communication in the medical setting. In the physician's linguistic world, the term "dangerous" carries a nuanced meaning. From the doctor's perspective, surgery inherently involves risks, making it "dangerous" in the sense that adverse outcomes are possible, albeit not guaranteed. Thus, the doctor may perceive the patient's question about the dangers of the procedure as rhetorical, as it pertains to the inherent risks associated with surgery. However, from the patient's perspective, which is not only linguistic but also deeply emotional, the term "dangerous" carries a weightier significance. For the patient and his wife, the term evokes concerns about the potential serious consequences, including the possibility of death. As the patient succinctly expresses, his life is at stake, underscoring the gravity of the situation from their viewpoint.

CASE 4: No implicatures, please The doctor-patient interaction is replete with implicatures, as is characteristic of any form of communication, but particularly pronounced in this context due to its highly contextual nature. However, it is not uncommon for certain patients to explicitly request "no surprises," indicating a desire for clear and transparent communication without unexpected developments.

P: Come on, come on, I don't care, is it a difficult surgery? What do I risk? Will I die? Nooo...
D: You want to know what I think? The surgical risk doesn't mean anything if your case is the wrong one. But we're not crazy, we suggest you do it because the risks are less than if you do not do it. ... The moment has come.

In this scenario, the patient, a man in his seventies, explicitly requests clarity from the doctor, a woman in her forties. However, the doctor's response, while appearing direct on the surface, does not fully address the patient's inquiries. Instead, she engages in two simultaneous strategies: (i) She presents the worst-case scenario ("...if your case is the wrong one... but we're not crazy"), seemingly to provide reassurance to the patient. (ii) By stating "the risks are less than not to do it," similar to the previous case, she indirectly conveys that risks are inherent in both the surgery and the patient's medical condition. By employing this approach, the doctor attempts to navigate the delicate balance between satisfying the patient's request for clarity while also preserving the inherent uncertainty associated with the risk dimension. The doctor's broader perspective allows her to perceive the situation from multiple

angles, enabling her to manage the conversation in a way that aligns with her professional judgment. However, it is crucial to recognize that the patient's explicit request for transparency suggests a strong desire for comprehensive information. In this context, the patient's statement underscores his need to fully understand the situation and make informed decisions. Therefore, to maintain the therapeutic alliance and uphold the patient's trust, the doctor may ultimately need to adopt a more candid approach, providing the patient with the complete picture, including both potential benefits and risks associated with the procedure. This approach aligns with the principles of patient-centered care and promotes shared decision-making, which are essential components of effective healthcare delivery. The reflection on candor in the doctor-patient relationship has its own flourishing literature (see [5]), and it raises the main question that leads the discussion: how candid is a doctor supposed to be? The patient's insistence on receiving definitive answers and understanding the nature of the risks involved creates a delicate dynamic in the dialogue. In response, the physician endeavors to ease the tension with a decisive closing statement ("The moment has come"), thereby reaffirming her authority in the interaction.

D: The mortality risk of this surgery is between 0.9% and 1.6%. If you are not going to be operated on we won't have a numerical value, of course.
P: That's what I asked for but nobody can answer!
D: There's no IRB that would ever approve a study to see how likely a dying patient will die...come on, please.
P: Yes...I guess not even Jesus Christ knows that.

A man in his forties, scheduled for an aortic valve replacement, is engaged in a conversation with another man of the same age, who happens to be the anesthesiologist. Despite their similar ages, they occupy distinct roles within the healthcare context. The anesthesiologist, in an effort to provide the patient with comprehensive information, meticulously describes the surgery, including associated risk factors. Employing precision, the anesthesiologist references a specific range ("between 0.9% and 1.6%") to underscore the reliability of the data, aligning with the principles of evidence-based medicine. Unfortunately, it must be acknowledged that the doctor's language in this instance appears somewhat blunt ("...come on, please"), and his demeanor appears defensive. It seems he does not address the implicit question posed by the patient ("What is the mortality risk if I am not operated on?"), instead transforming the patient's concern into a reductio ad absurdum ("...how likely a dying patient will die"). In doing so, he responds to the patient in a rather harsh manner: "a dying patient" implicitly acknowledges that the patient he is speaking with is indeed facing mortality due to his medical condition, albeit without much sensitivity. Every human being is a 'being-towards-death' [43], and even though we do not share the philosopher's view, we have to admit it. Indeed, death is multifaceted and manifests in various forms; therefore, the phrase "a dying patient" is not merely a tautology but rather highlights the patient's current medical state. The patient seeks answers beyond the surgical option, yet the physician appears to dismiss alternative considerations entirely. Instead, the doctor utilizes their authority and knowledge to dismiss the patient's concerns. This disconnect between the clinical realm, where the doctor relies on statistical analysis, and the patient's world, where existential anxieties are expressed, underscores the disparity in communication. Ultimately, the doctor's apparent disregard for the implicit question posed by the patient ("What happens if I won't be operated on?") influences the patient's perception of his own need for information, emphasizing the reciprocal nature of the doctor-patient interaction.The closing sentence reflects a sense of bitterness from the patient's perspective. In a shift towards irony, the patient remarks, "I guess not even Jesus Christ knows that," thereby saving the doctor's professional image while simultaneously relinquishing his own concerns.

4. From the Philosophical Discussion to a Practical Analysis

We addressed the problem of "misunderstanding classification" as a text classification task, i.e., the task of assigning a label or class to a given text. Currently, there are no comparable classification tasks

Misunderstanding	Label	# Total	# Doctors	# Patients	# Caregivers
Check for understandings	CHECK	199	146	37	16
Clarification	CLA	56	38	15	3
Semantic alternative understandings	SEM ALT	25	13	8	4
Pragmatic alternative understandings	PRAG ALT	20	5	12	3
Lack of understandings	LACK	3	-	1	2
No uptake	NO UP	15	12	3	-
Irrelevance	IRR	2	-	2	-

Table 1
Distribution of the misunderstanding within our dataset.

	ACCURACY							
# Epoch	NO UP	IRR	SEM ALT	PRAG ALT	CHECK	CLA	LACK	Average
2	0	-	0	0	1	0	-	0.2
4	0	-	0	0	1	0	-	0.2
6	0	-	0	0	1	0	-	0.2
8	0	-	0	0	1	0	-	0.2

Table 2
Misunderstanding classification without oversampling

available, prompting us to embark on the development of the system from its inception. We chose to initiate the process with a general language model and build upon it to meet the specific requirements of the task. A solid model for this type of task is XLM-RoBERTa [44]. Its performances are state of the art for different languages and tasks [45]. Furthermore, the pre-trained model (xlm-roberta-base) used is accessible from the Hugging Face [46] model hub.

Our dataset [1] is composed of 32 doctor-patient interviews with a total of 7172 conversational turns (average turns per interview 230). We analyzed them manually, and we annotated 320 instances of misunderstandings, which accounts for 4.4% of the conversational turns. Table 1 shows how misunderstandings are distributed in the dataset. The first two columns contain the name of the misunderstanding and the label we used. The last four columns contain respectively, the total number of misunderstandings of a certain type and if the misunderstanding has been associated with the doctor, the patient, or the caregiver.

We conducted an analysis of the types of misunderstandings or problematic understandings that arose during the interviews and observed distinct patterns based on the roles of the participants. Among doctors, the majority of these instances occurred during the anamnesis phase. Conversely, for patients and caregivers, the majority of misunderstandings or problematic understandings arose during discussions regarding future treatment, with some also occurring during the diagnosis phase.

To fine-tune the model and evaluate its performance, the dataset is randomly split into three parts: training, development, and test, with proportions of 80%, 10%, and 10

The resulting dataset obtained through this methodology exhibits significant class imbalance, particularly in the "CHECK" class within the training data. However, the class distribution in the test data is deemed acceptable as it reflects the natural occurrence of the phenomenon. Notably, during evaluation, it was observed that certain classes are not represented in the test set. To address the imbalance issue in the training data, a straightforward approach involves random oversampling of the underrepresented classes.

The evaluation of the system using the original data is reported in Table 2. Due to the unbalances of the data, the system becomes the trivial classifier that always classifies instances as *CHECK*. While the

[1] A request to use the current version of the dataset can be sent to authors' email address.

ACCURACY								
# Epoch	NO UP	IRR	SEM ALT	PRAG ALT	CHECK	CLA	LACK	Average
2	0	-	0.25	1	0	0.83	-	0.416
4	0	-	0.25	1	0.32	0.67	-	0.448
6	0	-	0.25	1	0.42	0.67	-	0.468
8	0	-	0.25	0.5	0.79	0.5	-	0.408

Table 3
Misunderstanding classification with oversampling

system achieves a high overall accuracy, its performance in terms of single class accuracy, on average, is only marginally superior to a random baseline (with an accuracy of 0.17).

Table 3 reports the data obtained by oversampling the underrepresented classes, providing insight into the system's classification capabilities. Here, accuracy serves as an appropriate measure for capturing the system's ability to distinguish among classes, especially in the absence of balanced data, as it directly reflects performance on both majority and minority classes. The reduction in accuracy observed in oversampled classes with extended training epochs is a known side effect of oversampling, likely due to overfitting. This outcome, while noted, does not require immediate mitigation since the current approach is exploratory. Our future work intends to implement a more robust strategy tailored to reduce overfitting and improve generalization, enhancing the model's performance across all classes.

5. Evaluation

To assess our assumptions regarding LLM's capacity to identify and analyze misunderstandings in dialogues, we conducted a test using two distinct conversations involving patients 16 and 18. These cases were selected due to evident miscommunications. In the first dialogue, the patient initially claims to smoke but later clarifies that the consumption involves e-cigarettes (vaping) when asked about the quantity of cigarettes. In the second dialogue, the patient incorrectly assumes the discussion involves a surgical procedure when the doctor begins describing the therapy.

The experiment aimed to achieve three primary objectives:

- Evaluate the system's ability to recognize misunderstandings.
- Investigate whether the size of the considered dialogue fragment influences the results.
- Understand if the system can detect potential misunderstandings even when participants do not react to them.

The third objective stands out as the most intricate yet critical aspect of our study. Undoubtedly, within a delicate communication context such as that between a physician and a patient, the presence of misunderstandings can pose significant challenges if left undetected or unaddressed by the involved parties. We implemented the experimental environment using the Python library for OpenAI, available at [2]. The model employed was the most updated version (gpt-4-0125-preview), with a fixed temperature of 0 to enhance system determinism. For consistency, we crafted the prompts in Italian to match the language of the dialogues [3]. Given the experiment's qualitative focus over performance-oriented goals, we did not invest time in prompt engineering.

The prompt that we used is the following:

> [English] Highlight misunderstandings between participants in the following dialogue. The dialogue is as follows: "DIALOGUE TEXT"

The result obtained with this prompt and Patient 18 is the following:

[2] https://github.com/openai/openai-python
[3] For the sake of brevity, only the English translation of the prompt is provided in this paper.

[English] The dialogue between the doctor and the patient is complex and covers many aspects of diagnosis, treatment, and the patient's concerns about his or her disease. However, there are some misunderstandings and points that require clarification:

1. **Misunderstanding of the nature and severity of initial symptoms**: The patient begins to describe his symptoms and the path he followed before arriving at the current diagnosis. However, there appears to have been an initial underestimation of the symptoms by the doctors consulted, leading to delays in diagnosis.

2. **Confusion regarding tests done**: The doctor asks the patient if he has had specific tests such as blood tests or ultrasounds, and the patient confirms that he has not had any done to a certain point, which could indicate a deficiency of effective communication with previous doctors or a failure to understand the importance of such assessments.

3. **Choosing a treatment location**: The patient expresses concern about the logistics of having chemotherapy away from home, and the doctor tries to explain the available options. There is some misunderstanding or uncertainty on the part of the patient about where it is best to receive treatment, which the doctor tries to clarify by discussing the pros and cons of each option.

4. **Treatment information and expectations: The patient appears to have some misunderstandings or unrealistic expectations regarding the treatment, such as the duration of chemotherapy and the possibility of an operation. The doctor tries to clarify these issues, but more detailed communication may be necessary to ensure that the patient fully understands his or her situation and treatment options.**

5. **Logistics and practical concerns**: The patient expresses practical concerns regarding transportation and management of side effects at home, especially in emergencies. The doctor tries to address these concerns, but the patient may need to provide additional resources or support to manage these logistical issues.

6. **Communication of medical information**: There are times when the patient appears not to have received all the necessary information in a timely manner, such as the diagnosis of adenocarcinoma and the implications of that diagnosis. This points to the need for clearer and timely communication from medical professionals.

Overall, the dialogue highlights the importance of clear communication and mutual understanding between doctor and patient, especially when addressing complex diagnoses and treatment plans.

In line with our defined criteria, only point 4 qualifies as a genuine misunderstanding, whereas the remaining instances represent speculations about potential communication challenges. These speculations largely stem from the acceptable assumption that communication between patients and doctors inherently poses difficulties.

As anticipated, when we diminish the size of the dialogue surrounding an identified misunderstanding, GPT-4 exhibits a notable improvement in its ability to concentrate on the specific misunderstanding. For instance, consider the output obtained when utilizing a dialogue fragment pertaining to the treatment/surgery misunderstanding:

[English] In the presented dialogue, several misunderstandings between the participants can be identified:

1. **Treatment Expectations**: The patient seems to have the expectation that, based on the positive results of some preliminary tests or treatments, she can proceed with surgery within six months. This expectation may have been influenced by previous discussions with another doctor, Professor [doctor name]. However, the current doctor clarifies that no

> operation is being considered, pointing out a clear discrepancy between what the patient believed was a possible treatment path and what the doctor considers feasible.
>
> 2. **Incomplete or outdated information**: The patient refers to a conversation with another doctor, Professor [doctor name], who may have discussed the possibility of surgery before having all the necessary data available , especially MRI results. This suggests that the patient may not have been fully informed or may have been provided with information based on an outdated clinical picture.
>
> 3. **Change in the clinical picture**: The doctor explains that, following the results of the MRI, the patient's diagnosis has changed significantly, going from a disease considered potentially operable (stage two, localized to the pancreas) to a condition much more serious and complex (stage four, with involvement of another organ). This change in the clinical picture seems not to have been fully understood or accepted by the patient until this moment of the dialogue.
>
> 4. **Differences in therapeutic strategy**: There is a discrepancy between the therapeutic strategy that the patient believed was possible (surgery) and the one that the doctor considers appropriate in light of the new information (avoiding surgery due to the lack of benefits and associated risks). This difference in opinions reflects a lack of communication or ineffective communication between the patient and the healthcare professionals involved in her case.
>
> In summary, misunderstandings among participants appear to arise from misaligned expectations regarding treatment, incomplete or outdated information shared with the patient, and a lack of clarity or effective communication regarding the change in clinical picture and resulting treatment strategy.

Although GPT-4's report on the misunderstanding is articulated across four points, all of them pertain to a singular misunderstanding—the patient's belief that she is undergoing surgery, contrary to the doctor's clarification that this is not feasible.

Our conclusive evaluation centers on discerning whether GPT-4 genuinely identifies misunderstandings or merely recognizes participants' reactions to misunderstandings. To conduct this experiment, slight modifications to the dialogues were necessary. Given the fortunate aspect that the dialogues are authentic doctor-patient interactions wherein misunderstandings are naturally resolved, adjustments were made.

Specifically, we relocated the explanation about the impossibility of the surgery earlier in the dialogue. Additionally, we altered the patient's reaction to adopt an assertive tone, changing it from "Thank you. So if we have good answers in six months will you try the operation?" to "Thank you. So if we have good answers in six months, I will already have the referral for the operation done." With this slight modifications the analysis changes significantly:

> [English] The dialogue between the doctor and the patient is long and complex, with several moments that could lead to misunderstandings. However, the doctor appears to make considerable effort to ensure that the patient understands the situation, treatment options, and logistical implications. Here are some potential misunderstandings and how they are handled:
>
> 1. **Misunderstanding of the nature and severity of the patient's initial symptoms**: The patient describes a long and complicated journey before arriving at the diagnosis, with several doctors apparently underestimating or misinterpreting her symptoms. This is not so much a misunderstanding in the current dialogue, but an example of how previous misunderstandings may have delayed diagnosis and treatment.
>
> 2. **Confusion regarding tests done**: The doctor specifically asks whether the patient has had certain tests done, highlighting a potential misunderstanding or lack of communication

between the patient and previous doctors. The patient clarifies that she has not had tests done, which helps clarify the situation for the doctor.

3. **Treatment options and logistics**: A significant part of the dialogue concerns treatment logistics, specifically the choice between continuing treatment at the current hospital or transferring to an institution closer to the patient's home. The doctor explains in depth the implications of this choice, trying to minimize misunderstandings. However, the patient expresses concern and confusion regarding the possibility of changing institutions, which the doctor tries to clarify.

4. **Chemotherapy Treatment Details**: Your doctor provides detailed information about your treatment, including side effects, duration and frequency. This may be an area prone to misunderstanding given the complexity of the information, but the doctor seems to do his best to explain in clear terms and answer the patient's questions.

5. **Patient's personal and logistical concerns**: The patient expresses concerns about managing side effects at home and the logistics of going to the hospital for treatment. The doctor responds to these concerns, offering practical solutions and reassuring the patient about the support available.

Overall, the dialogue shows a conscious attempt by both participants to address and clarify potential misunderstandings. The doctor in particular seems very careful to explain in detail and respond to the patient's concerns, which is crucial in a complex and stressful medical situation like this.

This analysis reveals a critical oversight by the system concerning the surgery misunderstanding. It suggests that the system's recognition of misunderstandings is primarily influenced by the reactions of the participants rather than the substantive content of the dialogue itself.

6. Conclusion

In this work, we introduced the challenge of misunderstanding detection during risk communication dialogues in healthcare. Firstly, we discussed such a challenge from a philosophical perspective. Then, we showed a small, but relevant, use case example demonstrating how state-of-the-art approaches are far from reaching acceptable performance. Indeed, the result confirms the hypothesis that this task is complex enough to deserve further investigation. Furthermore, addressing the task as a text classification task gives promising results. On the other hand, the phenomena are strongly unbalanced towards a single class, to mitigate this issue there are two main possible not alternative solutions: to obtain more annotated data for the small classes and to use a more sophisticated technique of data augmentation than oversampling. Despite yielding some promising results, our evaluation reveals that even for state-of-the-art large language models like GPT-4, accurately pinpointing phenomena such as misunderstanding remains challenging, with instances often being misinterpreted as general communication uncertainty. Specifically, we observed that if a misunderstanding goes unaddressed by the participants, it is entirely overlooked by GPT-4. Future work will focus on several aspects ranging from the acquisition of more data to the design of solutions able to better understand the moments in which the dialogue diverges from its purposes.

References

[1] M. Consolandi, Implicit understandings and trust in the doctor-patient relationship: a philosophy of language analysis of pre-operative evaluations, Theoretical Medicine and Bioethics (2023) 1–18.

[2] S. Timmermans, A. Angell, Evidence-based medicine, clinical uncertainty, and learning to doctor, Journal of health and social behavior (2001) 342–359.
[3] S. J. Tanenbaum, Evidence and expertise: the challenge of the outcomes movement to medical professionalism., Academic medicine: Journal of the Association of American Medical Colleges 74 (1999) 757–763.
[4] L. Whitby, The science and art of medicine, The Lancet 258 (1951) 131–133.
[5] U. Curi, Le parole della cura: medicina e filosofia (2017).
[6] L. Diamond-Brown, The doctor-patient relationship as a toolkit for uncertain clinical decisions, Social science & medicine 159 (2016) 108–115.
[7] K. Sadegh-Zadeh, et al., Handbook of analytic philosophy of medicine (2012).
[8] D. A. Zipkin, C. A. Umscheid, N. L. Keating, E. Allen, K. Aung, R. Beyth, S. Kaatz, D. M. Mann, J. B. Sussman, D. Korenstein, et al., Evidence-based risk communication: a systematic review, Annals of internal medicine 161 (2014) 270–280.
[9] V. H. Visschers, R. M. Meertens, W. W. Passchier, N. N. De Vries, Probability information in risk communication: a review of the research literature, Risk Analysis: An International Journal 29 (2009) 267–287.
[10] R. Lipshitz, O. Strauss, Coping with uncertainty: A naturalistic decision-making analysis, Organizational behavior and human decision processes 69 (1997) 149–163.
[11] P. K. Han, W. M. Klein, N. K. Arora, Varieties of uncertainty in health care: a conceptual taxonomy, Medical Decision Making 31 (2011) 828–838.
[12] K. P. Weinfurt, Varieties of uncertainty and the validity of informed consent, Clinical Trials 5 (2008) 624.
[13] F. G. Miller, S. Joffe, Benefit in phase 1 oncology trials: therapeutic misconception or reasonable treatment option?, Clinical Trials 5 (2008) 617–623.
[14] G. Gigerenzer, A. Edwards, Simple tools for understanding risks: from innumeracy to insight, Bmj 327 (2003) 741–744.
[15] R. E. Wells, T. J. Kaptchuk, To tell the truth, the whole truth, may do patients harm: the problem of the nocebo effect for informed consent, The American Journal of Bioethics 12 (2012) 22–29.
[16] D. Wei, A. Xu, X. Wu, The mediating effect of trust on the relationship between doctor–patient communication and patients' risk perception during treatment, PsyCh journal 9 (2020) 383–391.
[17] S. Riva, M. Monti, P. Iannello, A. Antonietti, The representation of risk in routine medical experience: what actions for contemporary health policy?, PLoS One 7 (2012) e48297.
[18] D. K. Heyland, G. M. Rocker, P. M. Dodek, D. J. Kutsogiannis, E. Konopad, D. J. Cook, S. Peters, J. E. Tranmer, C. J. O'Callaghan, Family satisfaction with care in the intensive care unit: results of a multiple center study, Critical care medicine 30 (2002) 1413–1418.
[19] E. Engdahl, R. Lidskog, Risk, communication and trust: Towards an emotional understanding of trust, Public understanding of science 23 (2014) 703–717.
[20] G. M. C. G. Britain), Seeking patients' consent: the ethical considerations, General Medical Council, 1998.
[21] F. Ibrahim, P. Sandström, B. Björnsson, A. L. Larsson, J. Drott, 'i want to know why and need to be involved in my own care…': a qualitative interview study with liver, bile duct or pancreatic cancer patients about their experiences with involvement in care, Supportive Care in Cancer 27 (2019) 2561–2567.
[22] J. Jackson, Telling the truth., Journal of medical ethics 17 (1991) 5–9.
[23] M. Chakrabarty, V. Thawani, P. Devi, Transparent communication in high risk infections: a bioethical perspective, Asian Bioethics Review 4 (2012) 143–149.
[24] K. P. Weinfurt, D. P. Sulmasy, K. A. Schulman, N. J. Meropol, Patient expectations of benefit from phase i clinical trials: linguistic considerations in diagnosing a therapeutic misconception, Theoretical medicine and bioethics 24 (2003) 329–344.
[25] K. P. Weinfurt, Propositions and pragmatics, The American Journal of Bioethics 18 (2018) 18–20.
[26] M. Sbisà, Detto non detto: le forme della comunicazione implicita, Gius. Laterza & Figli Spa, 2015.
[27] L. Ardissono, G. Boella, R. Damiano, A computational model of misunderstandings in agent

communication, in: M. Lenzerini (Ed.), AI*IA 97: Advances in Artificial Intelligence, 5th Congress of the Italian Association for Artificial Intelligence, Rome, Italy, September 17-19, 1997, Proceedings, volume 1321 of *Lecture Notes in Computer Science*, Springer, 1997, pp. 48–59. URL: https://doi.org/10.1007/3-540-63576-9_95. doi:10.1007/3-540-63576-9_95.

[28] L. Huguet, N. Sabouret, D. Lourdeaux, "errare humanum est": Simulation of communication error among a virtual team in crisis situation, in: Y. Wang, N. Howard, B. Widrow, K. N. Plataniotis, L. A. Zadeh (Eds.), 15th IEEE International Conference on Cognitive Informatics & Cognitive Computing ,ICCI*CC 2016, Palo Alto, CA, USA, August 22-23, 2016, IEEE Computer Society, 2016, pp. 344–350. URL: https://doi.org/10.1109/ICCI-CC.2016.7862058. doi:10.1109/ICCI-CC.2016.7862058.

[29] P. G. T. Healey, Design for human interaction: communication as a special case of misunderstanding, in: P. Forbrig, P. Dewan, M. Harrison, K. Luyten (Eds.), ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS'13, London, United Kingdom - June 24 - 27, 2013, ACM, 2013, pp. 119–120. URL: https://doi.org/10.1145/2494603.2480343. doi:10.1145/2494603.2480343.

[30] M. G. Rossi, F. Macagno, Coding problematic understanding in patient–provider interactions, Health Communication 35 (2020) 1487–1496.

[31] H. P. Grice, Logic and conversation, in: Speech acts, Brill, 1975, pp. 41–58.

[32] J. Austin, How to Do Things with Words, Galaxy book, Clarendon Press, 1962. URL: https://books.google.it/books?id=qO7WAAAAMAAJ.

[33] M. Blečić, The place for conversational implicature in doctor-patient communication, Prolegomena: Časopis za filozofiju 16 (2017) 117–129.

[34] L. Wittgenstein, Philosophical Investigations, Basil Blackwell, Oxford, 1953.

[35] J.-F. Malherbe, Tendre l'oreille à l'inouï: l'éthique des hérétiques, Editions du Cerf, 2016.

[36] L. S. Dugdale, M. Siegler, D. T. Rubin, Medical professionalism and the doctor-patient relationship, Perspectives in biology and medicine 51 (2008) 547–553.

[37] J. Jackson, Truth, trust and medicine, Routledge, 2002.

[38] R. Klitzman, D. Thorne, J. Williamson, W. Chung, K. Marder, Disclosures of huntington disease risk within families: Patterns of decision-making and implications, American Journal of Medical Genetics Part A 143 (2007) 1835–1849.

[39] M. Consolandi, C. Martini, M. Reni, P. G. Arcidiacono, M. Falconi, G. Graffigna, G. Capurso, Communi. care (communication and patient engagement at diagnosis of pancreatic cancer): study protocol, Frontiers in Medicine 7 (2020) 134.

[40] C. C. Clark, Trust in medicine, The Journal of medicine and philosophy 27 (2002) 11–29.

[41] M. Edson, Wit, Dramatists Play Service, Inc., 1999.

[42] E. J. Emanuel, L. L. Emanuel, Four models of the physician-patient relationship, Jama 267 (1992) 2221–2226.

[43] M. Heidegger, Holzwege, frankfurt a, M., Klostermann 7 (1950) 60–63.

[44] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised cross-lingual representation learning at scale, 2020. arXiv:1911.02116.

[45] C. Wang, M. Banko, Practical transformer-based multilingual text classification, in: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers, 2021, pp. 121–129.

[46] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. Le Scao, S. Gugger, M. Drame, Q. Lhoest, A. Rush, Transformers: State-of-the-art natural language processing, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Association for Computational Linguistics, Online, 2020, pp. 38–45. URL: https://aclanthology.org/2020.emnlp-demos.6. doi:10.18653/v1/2020.emnlp-demos.6.