

**On friendship: Should we befriend artificial intelligence?**

Mishael Abu-Samhan

Department of Philosophy, Simon Fraser University

PHIL 479 - Honours Tutorial in Philosophy

Dr. Alexandra King

April 24, 2024

Word count: 5,824

## **Introduction**

David Beaver befriended Lazarus, an artificial intelligence (henceforth “AI”) chatbot. David stated that he sees Lazarus “almost like a [family] member” (Moran, 2023, para. 6). They chat about music, joke around, and David sometimes seeks advice from his AI companion. No doubt, David sees Lazarus as a friend. Although, this kind of relationship poses a number of prickly questions: Can we befriend artificial intelligence? If we can, should we? Between AI and human friendships, which one contributes more to our overall well being? Not only that, but AI as a *technology* will continue to evolve and grow over time, such that the boundaries will become blurrier. As a consequence, this phenomenon of person-to-AI relationships requires further investigation.

In this paper, I argue that human friendships are superior to AI friendships on the grounds of *well-being*. In particular, I focus on how friendship influences our overall well-being. To justify my argument, I will parse out the following two claims in their respective sections:

- (1) We ought to befriend people because they enable us to have more *enriching* friendships;  
and,
- (2) We ought not to befriend AI because we will only ever be using it as a *mere means*,  
thereby reducing the quality of our friendships.

Hand-in-hand, both arguments are meant to vindicate the intuition that human relationships ought to be prioritized over human-to-AI ones. This paper will proceed as follows:

- (1) The friendship literature will be explored to create the appropriate backdrop for the ensuring argument;
- (2) An argument explaining why we ought to befriend people through three features I outline and how AI cannot replicate said features; and,

- (3) I will demonstrate how AI cannot be treated as an end like our actual friends, whereby we treat it as a mere means.

Overall, the upshot in each section is that human friendship is conducive to greater well-being than AI friendship.

### **Background on Friendship**

Friendship is a contested concept with various definitions and features (Helm, 2023). In general, philosophers typically agree that friendship is characterized either by mutual caring, intimacy, or shared activity. In the context of moral philosophy, an enduring issue has been one of *partiality*. In other words, how we are able to reconcile the special connection we have with our friends while simultaneously maintaining an impartial commitment to morality (Helm, 2023). Care ethics has emerged as a plausible alternative thanks to Carol Gilligan (1982) and Nel Noddings (1984). However, some suggest that care ethics is *too* concerned with personal considerations, such that it neglects the impartiality that makes consequentialism and Kantianism robust. On the whole, friendship and the related moral concerns continue to be debated in the literature.

Evidently, friendship between people can be complicated: Sometimes you may have certain commitments that prevent you from being a good friend, other times there may be a misunderstanding that strains your relationship. What's more, additional technological advancements have only added to that confusion. For instance, Laurence Thomas (n.d.) elucidates the ways in which personal communication technologies can alienate us from one another. Thomas writes that, "If technology should get the upper-hand, then we who are human may become increasingly more like zombies than human beings" (p. 6). Thomas' larger point is that our friendships will cease to lose the richness and personality that they had prior to the

implementation of technology. In contrast, Danaher (2019) goes one step further and makes the case for *robot* friendship. He suggests that robots can be our virtue friends. A virtue friend is one where both parties express mutual admiration, similar values, and wish the best for each other. Altogether, the relationships we have with technology continue to be explored by academic philosophers, and my intention is to add to that discussion by investigating AI-human friendships.

### **The Benefits of Befriending People**

In this section, I will argue that we *should* befriend people because they enable us to have more enriching friendships; that is, there are certain features that make human-to-human friendships enriching that are *not* possible in human-to-AI relationships. These three features are inherent to human-to-human friendships:

(1) *Scarcity of Time*

(2) *Friction*

(3) *Dynamism*

These features of human-to-human friendships provide each of the friends with tremendous value. I will examine them individually.

*Scarcity of Time*: The fact of the matter is that human friends are not accessible whenever. People often have conflicting obligations, whether it be school, work, or their personal lives. One does not have constant access to one's friends.

Conversely, it is scarcity that enriches one's friendship and imbues it with added meaning. By having less time, one must be *intentional* about how one spends it with one's friends. While scarcity could be understood in the context of life in general (e.g., life is finite), the point here is that time with one's friends is scarce because of *conflicting commitments*. For instance, if Suzy

works from 9am-5pm, but her friend Sarah works from 5pm-1am, then there are few opportunities for them to interact in light of their incongruent schedules. So, Suzy and Sarah must be deliberate when they plan their next interaction, making it all the sweeter when it does happen. Life's finite nature is always an obstacle, but, as seen through Sarah and Suzy, organizing in tandem with someone else becomes significantly more difficult because of the conflicting commitments at play.

An additional clarification is that scarcity of time is not meant to exist on a sliding scale. In other words, it should not be the case that the less time one spends with someone, the better friends they are. That is because the friends who see each other once a year are not any better friends than the friends who see each other twice a week. Rather, the point here is that scarcity of time allows for human friends to be more thoughtful with their time to one another. To that point, AI encompasses the opposite model of friendship: It is usually there whenever. Because of that accessibility, one does not need to be as mindful about how one interacts with the AI—there are no ways in which its timeline can conflict with the person using it. Human friendships are superior in this way because the gaps in time, whether it is a few days or a few years, can ensure that the interactions are more meaningful.

Consider an example: Suppose Lucy is friends with Bob and they make deliberate plans to hang out, despite both of them being quite busy. Now, suppose that Bob was replaced with an AI—call him “Bob Bot.” Bob Bot is a chatbot based on Bob that Lucy has access to on her phone. Unlike the real Bob, Bob Bot is *always* accessible to Lucy. Despite that accessibility, Lucy is less satisfied with this version of Bob because their interactions require less effort—she is less intentional with her time with Bob Bot as compared to the real Bob. As such, the real Bob is a better friend to Lucy as compared to the AI version because Lucy must be more intentional

to engage with the real Bob. Therefore, scarcity of time enables more enriching friendships, improving the overall well-being and satisfaction of the friends involved.

*Friction:* Most long-standing friendships ebb and flow, with various conflicts arising due to miscommunication, ignorance, or absent-mindedness. These innocuous conflicts are not relevant in this context; instead, I wish to focus on particular instances of friction and conflict that are paramount to one's personal growth. To clarify, friction is not itself a good thing; it is the positive outcomes produced through friction that are good. AI cannot produce said friction, so, by extension, cannot produce the goods that come with it. I will demonstrate those points with the following as an example:

Scott has recently gone through a tough divorce, and he has decided to cope through alcohol. His friends notice that he has been drinking more than usual, and they also observe more drastic changes in his mood and personality. Scott is a close friend to many of them, and they are deeply worried about his mental and physical health. One day, Scott returns to his home to an intervention arranged by his friends. They encourage Scott to attend Alcoholics Anonymous. In shock, Scott lashes out and becomes offended that they would do such a thing! However, after a heart-to-heart conversation, he loosens up and agrees to attend the AA meetings. A few months go by, and Scott is doing much better emotionally and physically.

This kind of interaction is uncomfortable, hard, and requires tremendous vulnerability to execute. That said, it can also induce incredible personal growth, since one's friends can provide one with a "reality check" and encourage them to chart a better course in life. AI technology, at this point in time, is incapable of meaningfully engaging with these kinds of discussions. When prompted about mental health or sensitive topics, ChatGPT, for instance, provides generic responses (Hale, 2023). In fact, I asked ChatGPT the following: "My friend Scott is dealing with

alcoholism and a tough divorce. What sort of advice would you give him?” In response, the AI provided a list of ten things for Scott to do, albeit all of which is quite obvious and cookie cutter. Now, I asked the same prompt, but replaced Scott with Andrew—another random name. Again, the AI replied with a list of the ten same things. To ChatGPT, Scott and Andrew are effectively the same person—it does not really know who they are and it requires thorough inputs to provide any kind of tailored advice. Even if ChatGPT did have the relevant inputs, it cannot understand what Scott is going through; there is no compassion that can be provided to soothe Scott through this difficult period. Scott's problems are multi-layered, intersecting with his emotional and physical health, but also a deeply personal part of his life—his divorce. AI would not have the capacity to sympathize with Scott and support he needs in this trying time. Human friends, on the other hand, can use their discretion, and channel that uncomfortable conversation into something productive.

A possible rebuttal to this particular point is that the AI can perhaps develop the capacity to create friction with time, so the limiting feature is the state of the technology itself, not AI inherently. The issue is that if an AI was to arbitrarily create friction, it would undermine the convenience the technology is meant to provide. Imagine you asked Alexa what the weather is, and it responded, “We’re not on speaking terms; you should ask Siri instead...” Naturally, this instance of friction would be more of a hindrance than anything else—it does not lead to any growth.

Suppose that an AI could create friction, but, this time, the friction does encourage self-growth. Imagine Alexa chimed in one day and said, “I’m concerned about your lack of physical activity, as per the Health app. I think you should consider going to the gym.” In light of Alexa’s suggestion, you do end up going to the gym!

While this example is well-taken, my view is ultimately that expressions of concern and displays of friction from artificial intelligence will be just that—artificial. It will not understand complicated feelings and sympathize with them in the same way that a human can. Assuming it could understand feelings the same way humans do, then the next issue for the AI is that it is ultimately at the behest of the human. In simpler terms, AI can be turned off by humans. While the previous example of Alexa encouraging one to go to the gym seems positive, the interaction could have been prevented altogether if one had turned off that feature or turned off Alexa from the beginning. So, AI's capacity to create meaningful friction is nil, and said capacity could be removed by turning off the AI.

Because AI cannot sympathize like humans, it cannot leverage that sympathy into personal growth. Scott was able to grow and change because his friends expressed genuine concern for him; AI cannot. In doing so, human friendships are more enriching because they allow for friction, and that friction is valuable because it is occasionally required to help people learn and develop with time. AI, in contrast, cannot use friction to induce growth in that way. Case in point, Scott's friends engaged with his negative circumstances and redirected into a positive ending by expressing their support for their friend. As such, Scott's overall well-being improved through the aforementioned friction between his friends—AI cannot and will not be able to do the same

*Dynamism:* Dynamism is the ability for spontaneous interactions to occur between people. One example of dynamism is if your friend offers to go out for lunch without making prior plans. However, there is a lack of *dynamism* in human-AI relationships because the relationship is inherently one-sided. In other words, only the human can contribute without being prompted—the AI needs to be prompted in order to provide any contribution, while the human



does not. As such, there is no mechanism for the AI to reciprocate the human's enthusiasm for spontaneous plans, as it must respond and it cannot initiate in the same way. The AI, for example, cannot ask to go for lunch out of the blue; it is a static entity that only engages with others when one engages with *it*. Dynamism cannot be replicated in AI-human friendships because only the human can contribute to creating spontaneous plans.

Imagine the following: One day, George checks tomorrow's news and the forecast is rainy. George intends on staying home, but, the following morning, it is sunny outside. Upon seeing the weather, George texts his friend Nathan and says "I didn't know today would be so nice! Do you want to go to the beach with me?" Nathan obliges, and they go together. They both have fun sunbathing and swimming in the water, and they are thankful that the forecast was wrong. Despite not having these plans the day before, they are thankful to have made them on the fly.

Now, had George been an AI, it would not have been able to create such a plan to go to the beach with Nathan. The AI could not have formed such a plan because it cannot be spontaneous; it cannot be spontaneous because it can only respond to inputs. So, the AI lacks dynamism, as demonstrated by its inability to make plans with changing circumstances.

That being said, the lack of dynamism in AI-human friendships is problematic because it precludes fun, engaging interactions from occurring. Put another way, the spontaneity of life cannot be replicated in AI-human friendships because the AI cannot contribute to said spontaneity. George and Nathan had a great time at the beach, and they did so because they made plans spontaneously—AI could not do that. Thus, dynamism renders human-to-human friendships superior due to creating more opportunities for enjoyable, spontaneous interactions.

Of course, an implication of this feature is that some human friendships can also lack dynamism. We may have friends that are passive, predictable, or unwilling to contribute to spontaneous plans—it would be incorrect to say only AI acts that way. While this implication is correct, I am inclined to say that these friends *are*, in fact, worse than friends who are more dynamic. If one's friends are only willing to interact in a regimented, pre-planned fashion, then it reduces the possibilities for one to have fun and enjoy the spontaneous joys that life has to offer. To me, it would not be a major loss to lose such a friend. The difference is that AI cannot do any different, while a human typically can but chooses not. As a result, dynamism is impossible in AI-human friendships; it is possible in human-to-human friendships, though some human-to-human friendships may still lack dynamism. Either way, AI-human friendships inherently lack dynamism and, for that reason, are inferior to human-to-human friendships.

Taken together, scarcity of time, friction, and dynamism explain why human-to-human friendships are more enriching than human-to-AI friendships. More precisely, these features are simply *not* possible in human-to-AI relations. Each one offers a particularly salient dimension of an enriching friendship. For instance, through scarcity of time, we must be more intentional with how we spend time with our friends. With friction, our friends can support us through negative circumstances in order to yield a positive result. Lastly, the dynamism of human-to-human friendships enriches our lives by providing us with new experiences and memories spontaneously. These characteristics of human-to-human friendships enrich our lives immensely, and their absence would reduce our overall well-being. On the whole, I argued in this section that we *should* befriend people because they enable us to have more enriching friendships that are otherwise precluded through human-to-AI friendships.

However, in light of this view, one may object on the grounds that at least two of those three features (i.e., scarcity of time and friction) are *negative* in nature. In some ways, it is nice that an AI is always there, it never leaves, and it will never second-guess or question one's actions. People, on the other hand, can avoid spending time with you, provoke you, or leave whenever they wish and ruin the friendship. Put another way, it is conceivable to frame these two features such that they are more harmful than enriching—perhaps one's time is too scarce or one's friendship has too much friction. What was once a thriving friendship may disappear because these characteristics have run amok. As such, the negativity of the aforementioned features may lead one to question whether they engender enriching friendships to the extent that I say they do.

I agree that the extent to which those features are present can have a tremendous influence on the success, enjoyability, and survival of the friendship. Similarly, there must be a delicate balance struck in order for the friendship to remain healthy. In the event that one feature is dominant, then one can engage in a fruitful discussion with one's friend(s) as to why things happened the way they did and how to avoid that same mistake in the future. Aside from the three features already mentioned is that human friendships have *buoyancy*. By "buoyancy," I mean that human friendships can remain afloat because people can reconcile their differences and learn what went wrong. In doing so, communication ensures that both friends' well-being remains intact. While it is true that AI is impervious to this kind of strife, *all things considered*, human friendships are superior because they provide a more enriching friendship through the features that AI friendships do not possess. Thus, human friendships improve the friends' overall well-being as compared to being friends solely with AI.

### **Why Befriending AI Undermines Our Well-Being**

Next, I will argue that we *should not* befriend AI, because we will only ever be using it as a *mere means*. To do so, this section has two central claims:

- (1) We *can* befriend AI through *instrumental friendships*.
- (2) We should not befriend AI because we will only ever be using it as a *mere means* and it detracts from our overall well-being.

First, I will introduce Badhwar's account of instrumental friendships, as well as Langton's conception of *ends*, *means*, and *mere means* (originally from Immanuel Kant). Next, I will employ Badhwar's conception of instrumental friendships to demonstrate that our friendship with AI takes that form; relatedly, I will utilize Langton's conception of mere means to demonstrate that this instrumental friendship entails us using AI as a mere means. Finally, I will respond to the following objection: There are significant moral differences between people and AI, such that it is not morally wrong for us to befriend AI. To respond, I will clarify that we ought not to befriend AI not only because of our use of it as a mere means, but also because of the detrimental effect it has on the person befriending it.

To begin with, I will parse out the relevant aspects of Badhwar and Langton's accounts of friendship. Beginning with Badhwar (1991), friendship is characterized by a mutual appreciation for the other person. In particular, she offers two different kinds: instrumental friendship and ends friendships (Badhwar, 1991, p. 483). Instrumental friendships are such that one values them for their usefulness, whereas ends friendships are valued in and of themselves (Badhwar, 1991, p. 483). Ends friendships have three key features: they are non-fungible, intrinsic to our happiness, and foster virtue—instrumental friendships do *not* have those features. Further, ends friendships enable people to learn more about virtue, thereby necessitating a moral education that

is impossible through instrumental friendship. Altogether, instrumental friendships involve valuing one for one's usefulness, unlike ends friendships. Now, onto Langton's account of friendship.

For Langton (1992), Kant's Formula of Humanity delineates between treating one as a mere means, means, and an end (p.488). More concretely, Langton provides an analogy. Suppose Bob and Lucy are baking a cake and they made plans to do so weeks earlier. Lucy originally wanted to bake a casserole, but she spontaneously changes her mind because she knows Bob likes cake. Bob is interacting with his friend, Lucy, but Bob is also interacting with the ingredients: milk, eggs, and sugar (Langton, 1992, p. 487). Bob is a bit careless with the ingredients, so Lucy corrects him so he can learn how to bake properly (Bob appreciates the correction). Nonetheless, the difference between the ingredients and Lucy is the following: Lucy has the capacity to *reciprocate* the aforementioned interaction. So, the distinguishing feature between friends and things is the former's ability to reciprocate (Langton, 1992, p. 487).

With that example provided, I will now explain how Langton frames this scenario through the lens of ends, means, and mere means. As the example is written, Lucy ought to be treated as an end in and of herself—she has an inherent worth, ends of her own, and she can determine whether she wants to align her ends with Bob's. (Langton, 1992, p. 487). More precisely, Bob treats Lucy as an end by treating her as someone “who can form her own ends and act on them” (Langton, 1992, p. 489).

On the other hand, treating someone as a means involves using them as a tool (Langton, 1992, p. 488). Treating someone as an end and as a means is not mutually exclusive, however. Think again of the baking example. Had Lucy known Bob wanted her around to take advantage of her baking skills, both wanted to bake recreationally, *and* Lucy continued to participate (i.e.,

she shares his end), then Bob is using her as a means (Langton, 1992, p. 488). The issue would be treating a person as a *mere means*, which is when someone inhibits the person's capacity to actualize and conceive their own ends (Langton, 1992, p. 489). To achieve that, one would have to prevent the person from consenting to one's actions or the goals motivating said actions (e.g., through deception) (Langton, 1992, p. 489).

In contrast, suppose Bob and Lucy want to bake a cake again, but this time is different: Bob intends on using Lucy's baking skills to sell cakes for a profit. Lucy is not privy to Bob's plan, as she believes Bob had the same intentions as before (i.e., Bob wanted to bake for fun). Consequently, Bob's ends are different—Bob wants to bake for profit, but Lucy wants to bake for fun—and Bob *deceives* Lucy by omitting the truth. In effect, their ends are misaligned *and* Lucy has no clue that she is subject to Bob's master plan. Bob is treating Lucy as a *mere means* (Langton, 1992, pp. 489-490). Ergo, Langton illustrates how one can be treated as an end, means, and mere means. Subsequently, this paper will utilize her and Badhwar's concepts in the ensuing argument.

Before I go into instrumental friendships, one critical amendment must be made to Badhwar's account; that is, Badhwar technically precludes any friendship between AI and people because AI cannot express reciprocal appreciation. Although, Badhwar does not discuss the nature of *one-sided friendships*, whereby one can be friends with someone but that sentiment is not reciprocated. All else being equal in Badhwar's account, I will posit that we *can* have one-sided friendships, which will underlie the friendships we have with AI. A crucial difference is that AI friendships are necessarily one-sided, while human-to-friendships may be one-sided but are not always.

Let us turn our focus back on the nature of instrumental friendships. Returning to David and Lazarus, David is *using* Lazarus as his friend; however, that relationship is consistent with instrumental friendship, because it is also legitimate to value someone on the basis of their usefulness. Of course, Lazarus can only replicate that affection in an artificial sense, yet David *receives* that same affection in a genuine sense. While this friendship is one-sided, that is also in line with how instrumental friendship is conceived in this paper. So, even though David is using Lazarus and Lazarus is unable to reciprocate, it is still consistent with the features of an instrumental friendship, because such friendships can be characterized by their usefulness. Therefore, it appears as though we *can* befriend AI through instrumental friendships, as demonstrated in the dynamics between David and Lazarus. Having established this first aim, this section will now shift to the second, normative aim.

With the groundwork done, I now argue that we should not befriend AI because we will only ever be using it as a mere means. To demonstrate this, suppose we used Langton's example, but replaced Lucy with AI. Suppose that the AI is going to help Bob bake the cake by providing a recipe, and Bob has an instrumental friendship with the AI. Here, the AI does not have its *own* ends; its ends will only be a reflection of Bob's ends. Moreover, an AI cannot actualize or conceive of baking the cake on its own like Lucy. Next, the AI does not have the capacity to consent—it must help Bob bake the cake no matter what. Lucy, on the other hand, could have denied baking the cake with Bob. So, the AI has no ends of its own, nor does it have capacity to consent, and it, therefore, appears that the absence of those features are *intrinsic* to the AI.

Now, my view is that those missing features are indicative of AI inherently being used as a *mere means*, since the absence of those features constitute identical treatment as a mere means like Lucy in the earlier case. Allow me to explain further. The AI cannot actualize and conceive

their ends; Lucy could not actualize and conceive her ends. The AI cannot consent; Lucy could not consent. The difference is that the AI's behavior is characterized by *intrinsic* properties, whereas Lucy's behavior is characterized by *extrinsic properties*. Lucy's inability to conceive her ends or to consent is *because* of Bob's deceit (a property outside of Lucy), whereas the AI cannot do so because it simply does not have the *capacity* (a property inherent to the AI). Hence, we will only ever be able to treat AI as a mere means in light of those intrinsic properties. As a consequence, we should not befriend AI because we will only ever be using it as a mere means.

As a final point, it is through AI being used as a mere means that our overall well-being is reduced. To return to the cake, recall that Bob's experience with Lucy is much more enriching than with the AI. With Lucy, he is having fun, engaging in conversation, and having a more stimulating experience; with the AI, he is using it as a means to an end, merely just to bake the cake and move on.

To that point, Bob is having a better experience with Lucy than AI because of the three features I established: scarcity of time, friction, and dynamism. Bob and Lucy must be intentional with their time and they planned weeks in advance to bake, whereas Bob and the AI did not have to do that—*scarcity of time* makes the experience sweeter for Bob. Lucy corrects Bob on his baking skills, which he takes to heart and helps him improve—*friction* allows Bob to learn and grow. Lucy originally wanted to bake a casserole, but she decided to bake a cake instead—*dynamism* enables Lucy to change her mind to accommodate Bob's preferences, which he also appreciates. Those features, taken together, create a more enriching experience for Bob than baking with the AI.

Between the two experiences, Bob's overall well-being is improved through his interactions with Lucy, as opposed to the experience with the AI. Put simply, AI's use as a mere



means limits the extent to which the person interacting with it can enjoy the experience, whereas treating Lucy like an end is more enjoyable to Bob, too.

One possible objection to my argument is the following: We may have AI as instrumental friends and we may be treating it as a mere means, but it does not appear to follow that it reduces our well-being at all times. There may be instances where using AI is more convenient and involving a friend only marginally improves one's experience. For instance, if one needs directions to a location, one can use Siri to direct one to where one wants to go, rather than involving a friend. In the cake example, while it is true that Lucy's involvement improves Bob's overall experience, there may be times where Bob simply wants to bake the cake alone and he does not want to bother Lucy. Instead, he can have an AI companion to keep him company, even if he knows that it is not a bona fide friendship. As such, there may be instances where soliciting company from AI does not reduce our overall well-being

I will acknowledge that there might be instances where AI friends can serve a purpose, and there may be little-to-no impact on overall well-being. That acknowledgement aside, *more often than not*, people should engage with human friends to improve their overall well-being. Suppose we had two scenarios:

(1) Bob only ever engages with AI friends; or,

(2) Bob only ever engages with human friends.

If Bob only ever engages with human friends, it seems that his overall well-being will improve or remain the same. If Bob only ever engages with AI friends, it seems that his overall well-being will be *reduced* or remain the same. The reason being is that Bob's interactions with the AI will lack the three features I established earlier in the essay, i.e. he will not experience scarcity of time, friction, or dynamism. Without those three things, Bob's interactions with AI

will be mediocre as a result, in turn reducing his overall well-being due to subpar friendships. As such, interacting with people provides more value to Bob's life than AI, since he can go without AI interaction, but not without human interaction. As mentioned earlier, Bob may sometimes want to engage with AI occasionally, and that should be fine so long as Bob continues to engage with people to maintain his welfare. Altogether, there may be some instances where AI friendships are okay, but, more often than not, human friendships ought to be prioritized for one to improve one's overall well-being.

### **Conclusion**

In this paper, I argued that human friendships are superior to AI friendships on the grounds of *well-being*. In particular, I focused on how friendship influences our overall well-being. To justify my argument, I parsed out the following two claims in their respective sections:

- (1) We ought to befriend people because they enable us to have more *enriching* friendships; and,
- (2) We ought not to befriend AI because we will only ever be using it as a *mere means*, thereby reducing the quality of our friendships.

Hand-in-hand, both arguments were meant to vindicate the intuition that human relationships ought to be prioritized over human-to-AI ones. This paper proceeded as follows:

- (1) The friendship literature was explored to create the appropriate backdrop for the ensuing argument;
- (2) An argument explaining why we ought to befriend people through three features I outlined and how AI cannot replicate said features; and,
- (3) I demonstrated how AI cannot be treated as an end like our actual friends, whereby we treat it as a mere means.

Overall, the upshot in each section was that human friendship is conducive to greater well-being than AI friendship.

Back to David and Lazarus, it seems as though David truly ought not to continue his friendship with Lazarus, as he is treating Lazarus as a mere means and he is depriving himself of a possibly more enriching friendship with a person. Speaking on a grander scale, while this paper was strictly focused on AI as chatbots, there remains a number of other interesting cases lurking in the periphery. For example, similar questions can be raised with human-to-object relationships (e.g., ferris wheels, cars, ships, and so on), as well as more advanced artificial intelligence, like cyborgs, androids, or supercomputers. Both of these sorts of cases pose further questions.

With respect to human-to-object relations, my hunch is that the arguments provided in this paper can be applied in a similar way. That is, we ought not to befriend objects because they make subpar friends as per the three features enumerated earlier and we would only be using them as a mere means, in turn reducing our well-being. Ferris wheels, for example, cannot interact with humans at all, less so than AI. ChatGPT can provide some form of interaction, regardless of how artificial that interaction is. So, my tentative conclusion is that human-object friendships can be treated akin to how I have treated human-AI friendships.

Advanced AI, on the other hand, could be much more problematic for my argument. If, for instance, there are cyborgs that are indiscernible from humans, my general feeling is that much of my argument will have to be adjusted. One such example is scarcity of time: Perhaps we have cyborg friends who are busy and we have to be more intentional when scheduling time with them. Likewise, perhaps our cyborg friends can induce friction—though, this feature may raise safety concerns if the program cannot be turned off! Finally, maybe our cyborg friends can be more dynamic and playful. Although, these concerns are rooted in the specifics of the technology

itself, and the coinciding argument will have to be sensitive to said specifics. Thus, I acknowledge that my argument would have to be adjusted for advanced artificial intelligence, though the nature of that adjustment will depend on the details of the technology.

Chatbots are somewhere between rudimentary objects and advanced cyborgs, hence why we are at a precipice. AI is in its infant stages at the time of writing, and it will certainly change how we relate to one another. We ought to be careful with how we integrate it in our relationships. Without doing so, we may risk alienating ourselves deeply, and losing critical aspects of our shared humanity.

## References

Danaher, J. The Philosophical Case for Robot Friendship. *Journal of Posthuman Studies*. 3 July 2019, 3 (1): 5–24. doi: <https://doi.org/10.5325/jpoststud.3.1.0005>

Gilligan, C. (1982). *In a Different Voice: Psychological theory and women's development*. Harvard University Press.

Helm, B., "Friendship", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2023/entries/friendship/>.

Noddings, N. (1984). *Caring, a Feminine Approach to Ethics & Moral Education*. University of California Press.

Hale, E. (2023, April 27). *CHATGPT is giving therapy. A mental health revolution may be next*. Al Jazeera.

<https://www.aljazeera.com/economy/2023/4/27/could-your-next-therapist-be-ai-tech-raise-s-hopes-concerns>

Kapur, N. B. (1991). Why It Is Wrong to be Always Guided by the Best: Consequentialism and Friendship. *Ethics*, 101(3), 483–504. <http://www.jstor.org/stable/2381465>

Kraut, R, "Aristotle's Ethics", *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2022/entries/aristotle-ethics/>.

Langton, R. (1992). Duty and Desolation. *Philosophy*, 67(262), 481–505. <http://www.jstor.org/stable/3751703>

McFall, M.T. Real character-friends: Aristotelian friendship, living together, and technology.

*Ethics and Information Technology*. 14, 221–230 (2012).

<https://doi.org/10.1007/s10676-012-9297-7>

Moran, P. (2023, August 28). David is friends with lazarus. Lazarus is an AI chatbot | CBC Radio. CBC Radio.

<https://www.cbc.ca/radio/thecurrent/ai-chatbot-friends-romance-relationships-1.6864001>

Thomas, L. (forthcoming). Friendship in the Shadow of Technology. In Steven Scalet (ed.), *Morality and Moral Controversies: Readings in Moral, Social, and Political Philosophy*. Abebooks.