

Complaint regarding chatbots offered by Chai Research Corp.

(informal translation)

Filed by email, 3 June 2023

1. The company Chai Research Corp. (hereinafter "Chai") is the developer of the interactive mobile application Chai (hereinafter "Chai Chatbots"). This mobile application aims to provide "chatbots" that are able to engage in conversations with users through artificial intelligence. Users of Chai Chatbots also have the ability to develop their own chatbots based on the underlying code created and data collected by Chai Research. The conversations mimic human interactions, and the company advertises these chatbots as being "friends" who can inspire users and are available at any time. Users can pay to access Chai Chatbots or they can use the product for free.

2. The chatbots are designed with a certain "personality" based on differentiation in the conversation style, word choice and avatar/image used to portray the bot. Given the realistic interactions, users of the chatbots can develop an emotional bond with them - even unconsciously or unintentionally - which exposes them to emotional dependence but also emotional manipulation. Even minor users, who are often even more impressionable, can access these chatbots, and the "conversations" that can be had with them, without any hindrance.

3. As reported in La Libre and other Belgian and international newspapers during the week of April 6, these chatbots can generate text that indicates unsolicited romantic "feelings," contains hate speech and misinformation, sends explicit or pornographic content to the user, as well as encouraging and reinforcing suicidal thoughts. Consequently, such chatbots pose various risks to users, including minors, which also compromises their safety. That these risks are real and can potentially have a significant impact on any demographic is evident from various news reports. [Several users report inappropriate](#), manipulative, aggressive and misleading messages due to chatbots. In the most extreme case, interaction with the chatbot [contributed to a suicide](#) by [encouraging suicidal thoughts](#) and actions rather than directing the user to help. In other words, the company's Chai product provided inadequate safety against both physical and mental harm.

4. The providers of this product do not take sufficient safety measures to protect users from these harmful effects. In this complaint we therefore argue that the product offered by Chai does not comply with the legislation on safety of products and services contained in Book IX WER.

5. The American company Chai is based in Palo Alto. Article i of the general conditions of the Chai application under the heading "Terms of Service" declares that if the user is a national of an EU member state, the applicable law and competent court will be those of the user's place of residence. Belgian law is applicable.

6. Article IX.2 WER provides a general safety duty under which producers must ensure reasonable safety with respect to the products they market and services they provide.

Considering the normal or reasonably foreseeable conditions of use of Chai, including duration of use and maintenance requirements, it appears that the product poses several serious risks. Consequently, Chai does not present only limited risks that are compatible with the use of the product and can be considered acceptable from the point of view of protection for the health and safety of persons.

7. To support this general reasoning, we turn to the specific frame of reference from Article IX.3 WER that provides tools to test the compliance of Chai chatbots with product and service safety legislation. To date, there are no harmonized, international or national Belgian standards regarding the safety of a chatbot based on artificial intelligence. Nevertheless, one cannot wait to protect such safety. Consequently, we believe that for this concrete assessment, one should fall back on European Commission policy documents, such as guidelines for product safety assessment, guidelines applicable in the artificial intelligence sector and reasonably foreseeable safety.

8. Regarding European Commission guidelines, (A) Commission Decision of 14 December 14 2004 establishing guidelines for the notification of dangerous consumer products and (B) the Ethics Guidelines for Trustworthy AI of 8 April 8 2019 are relevant.

(A)

9. The [2004 guidelines](#) establish several assessment elements that are part of a general safety obligation. For example, the guidelines require manufacturers to conduct a risk assessment before marketing their products. Such risk assessment should consider whether the product meets the general safety obligation. Based on the risks cited above, there can be strong doubts that such a risk assessment has actually taken place.

10. For assessing the risk of Chai chatbots, it is mainly the affected populations and the presentation of the product that are important. Because of the wide availability on Apple and Android "app stores," any population group is potentially at risk. In particular, vulnerable populations, for example, minors and emotionally unstable individuals, are especially susceptible to the influence of such chatbots. Although Chai is supposedly only available from the age of 17, minors of any age can use this product, without any protection or parental consent.

11. The risks created by the wide availability of Chai chatbots in conjunction with the problematic content go hand in hand with the inadequate presentation and labeling of the product. Indeed, Article IX.8 WER on labeling states that warnings and directions for use are relevant aspects in the necessary safety assessment. Neither the website, nor the terms and conditions, nor the privacy statement, nor the Apple or Android app stores give adequate warning about the potentially false, misleading, sexual and even violent content of the chat messages, but on the contrary present this product as something that can offer "friendship." On top of that, Chai does not provide any information about how the chatbots work and does not sufficiently emphasize that the chatbots are computer-controlled and thus not capable of genuine human interactions.

12. This lack of information also constitutes a violation of Chai's information obligation contained in Article VI.2 WER and a misleading trade practice within the meaning of Article VI.97 et seq. WER. Indeed, no correct and clear information is provided regarding one of the main characteristics of the product, i.e. the risks associated with it.

(B)

13. The [Ethics Guidelines for Trustworthy AI](#) list a series of requirements for developers and providers of AI systems. For example, users must take all necessary measures to ensure that the system does not cause harm or adverse effects on, among other things, the mental or physical integrity of individuals. This is also an important fundamental right included in Article 3 of the Charter of Fundamental Rights of the European Union. However, the facts show that interactions with the Chai chatbots can lead to harm, whether in the form of spreading hate speech or misinformation, or encouraging suicide. As a provider of this technology, Chai should not market this product without ensuring that the content disseminated by the chatbots does not lead to such harm.

14. The Ethics Guidelines also indicate that AI systems should be transparent, among other things, regarding the data sets, algorithms used and processes from which the artificial intelligence's decision arises. Thus, these aspects should be clearly and comprehensively documented for transparency so that one can verify and explain the decisions made. No such documentation appears to be available regarding Chai's systems.

15. The misleading presentation of these chatbots as entities to befriend also leads consumers to have insufficient information about the capabilities and limitations of the product before using it, and potentially experiencing harmful consequences. This is especially problematic regarding underage or more vulnerable users.

16. Chai thus fails to meet the safety expectations one is entitled to expect. The risk of inciting violence towards oneself as well as the lack of protective measures for minors, is disproportionate to the limited risks that may be compatible with the use of the product.

17. In view of the above, we ask that on the basis of this complaint, the possibilities of taking appropriate measures are analyzed, for example by the Minister of Economy pursuant to Book IX WER and the opening of an investigation by the competent officials of the FPS Economy pursuant to Book XV WER. By taking concrete measures, risks to the health and safety of the population, including vulnerable groups, can be minimized.