# **Briefing sheet and response sheet**

Please study the scenario below.

Please put your responses in the answer boxes below the relevant question and when you have finished, submit this document to the assignment.

It is important to remember that this is an individual piece of work:

* It will be submitted to Turnitin to support the academic integrity of the assignment
* You must work on your own and must not involve anyone else in this assignment work.

## **Scenario**

Ade is a 13-year-old student who has been struggling with her mental health. She has found it difficult to focus on her studies due to low mood and excess worry. While searching the internet, Ade found a free AI chatbot service that helps users with mental health difficulties. Based on her conversations with the chatbot, it suggested a diagnosis of bipolar disorder due to trauma from her childhood.

Ade discussed this diagnosis with her parents who went on to research the website. Their first discovery was that the service stores all conversations with the chatbot for further training of its underlying models. Secondly, they noticed that the website did not clearly state if the bot had been appropriately trained in psychotherapy techniques, or if the service has received the regulatory approval required for offering such advice.

## **Questions and answer boxes**

|  |  |
| --- | --- |
| **(a)**  **40 marks**  **Word limit: 350 words** | Identify **ONE** ethical issue that is relevant to the scenario above and analyse it from a **Consequentialist** perspective. |
| **Answer:** The consequentialist perspective comes under a framework of ethical theory known as Consequentialism, the doctrine that the moral worth of an action is determined only by its consequences.  For example: releasing a software update that will enhance the functionality of 80% of users’ phones, but prevent the remaining 20% from receiving updates due to hardware limitations. Utilitarianism is a prominent example of consequentialism (Pyper, 2024).  In this scenario, Ade using the AI chatbot may offer short-term benefits in emotional support and reduce stigma by using the service for help. Still, the risk of low data privacy and protection becomes a vivid issue from a consequential point of view. We learn from the case that the AI chatbot service stores all conversations with clients for further training of its models. This could have adverse consequences if the sensitive data of users is mishandled or exploited by probably sharing with third parties or not properly safeguarding the data.  There is also the case where the AI chatbot service did not clearly state that the model has been trained in psychotherapy techniques and has not received approval from licensed regulatory bodies. There could be a chance of inaccurate diagnosis or advice being given resulting in negative effects, in case Ade and her patients believed the chatbot and started to seek medication which could delay getting professional care and cause long-term harm.  From a consequentialist perspective, the harm caused by the risk of misuse of sensitive data and the risk of inaccurate or misleading advice outweighs the benefits the chatbot may offer in terms of immediate emotional relief to the client. Though Ade may experience some level of support from the AI chatbot, the long-term consequences such as compromised privacy, potential data breaches, and the possible delay of professional care pose significant risks to her well-being. |
| **(b)**  **30 marks**  **Word limit: 250 words** | Using the **Ethical OS toolkit**, identify **ONE** risk zone that is most applicable to the case study above with its related unintended consequences. Justify the reasons for your choice. |
| **Answer:** The Ethical OS Toolkit is a framework designed to facilitate better product development, faster deployment, and more impactful innovation while striving to minimise technical and reputational risks. It helps the makers of tech, engineers, product managers, and others get out in front of problems before they happen (“A guide to anticipating the future impact of today’s technology ”, 2018).  In the case of Ade using the AI chatbot, the most relevant risk zone to consider from the Ethical OS Toolkit is Risk Zone 7: Implicit Trust and User Understanding. This zone focuses on how personal data is collected, stored, used, and shared. It assesses whether a technology or service does anything the users do not know about, or would probably be surprised to find out about (“Risk Mitigation Checklist”, 2018).  Considering Ade being a minor using this AI chatbot service to discuss sensitive matters like her mental health raises concerns on privacy and data security issues. The website may have cookie prompts for users the first time they visit the platform but given the age group of this user, she is less likely to read and understand privacy policies which makes her vulnerable to a breach in data privacy and security because she may not know that information about her are kept online and could potentially hurt her later.  And ultimately, the unintended risk of exposing or misusing sensitive data could cause emotional harm to Ade and her parents, damaging her trust in digital mental health tools, and even have long-term consequences on her well-being.  This makes the Implicit Trust and User Understanding a relevant risk zone from the Ethical OS Toolkit about this scenario. |
| **(c)**  **30 marks**  **Word limit: 250 words** | Identify the subsection of Principle 1 (**Public Interest**) of the **BCS Code of Conduct** that is the most relevant to this scenario. Justify the reasons for your choice. |
| **Answer:** The British Computer Society (BCS) Code of Conduct sets rules and standards that direct the behaviour of its members in professional matters. It is expected that these rules and standards will be higher than those established by the general law and that they will be enforced through disciplinary action and any breach of the Code of Conduct will be considered under the Institute’s disciplinary procedures.  Looking at Principle 1 (Public Interest) of the BCS Code of Conduct, subsection 1.1 which states “You shall have due regard of public health, privacy, security, and wellbeing of others and the environment” makes a relevant section to evaluate this scenario (“Code of Conduct for BCS Members”, 2022).  In this case, the AI chatbot provides a mental health service to Ade, a vulnerable 13-year-old, who is struggling with mental health challenges such as low mood and excessive worry. The chatbot suggests the diagnosis of bipolar disorder and its engagement in sensitive mental health matters raise significant concerns about the public interest, particularly the health, safety, and well-being of individuals like Ade.  The AI chatbot provides mental health diagnoses to help Ade but could have serious implications if the diagnosis is inaccurate, premature, or unsupported by appropriate psychotherapy evaluations. Providing such a diagnosis without proper authorisation or safeguards can jeopardise the well-being of the individual, especially young people who are in a vulnerable state emotionally and mentally.  The service storing conversations (Ade’s personal and sensitive mental health data) for further model training without Ade or her parents’ informed consent raises significant concerns about privacy and security which is essential for protecting the public interest. In particular, the website does not make clear whether it meets regulatory standards for data storage and mental health counselling which could expose users to risks like data breaches or misuse of sensitive information. |

**References**

Institute for the Future and Omidyar Network (2018). *A guide to anticipating the future impacts of today’s technology* [PDF]. Ethical OS. Retrieved from Lectorial 03 Online Web Site: https:// [herts.instructure.com/courses/116886/pages/lectorial-03-ethical-risks-with-autonomous-](http://herts.instructure.com/courses/116886/pages/lectorial-03-ethical-risks-with-autonomous-) vehicles?module\_item\_id=3911966

Institute for the Future and Omidyar Network (2018). *Risk and Mitigation Checklist* [PDF]. Ethical OS. Retrieved from Lectorial 03 Online Web Site: <https://herts.instructure.com/courses/116886/> pages/lectorial-03-ethical-risks-with-autonomous-vehicles?module\_item\_id=3911966

Pyper, A. *Ethical issues and organisational culture* [PDF document]. Retrieved from Lecture 03 Online Web Site: <https://herts.instructure.com/courses/116886/pages/lecture-03-ethical-issues-> and-organisational-culture?module\_item\_id=4379737

BCS, The Chartered Institute for IT (2022). *Code of Conduct for BCS Members* [PDF]. BCS Code of Conduct. https://www.bcs.org/media/2211/bcs-code-of-conduct.pdf