

ASPECTS OF HEALTH-RELATED HARM CAUSED BY AI (ARTIFICIAL INTELLIGENCE)

KILIAN GROSSMANN¹, PRIV.-DOZ. DR. FABIAN RENGER¹

¹ ST. ELIZABETH-UNIVERSITY BRATISLAVA, SLOVAKIA

Abstract

Introduction: The increasing integration of artificial intelligence (AI) into everyday life, particularly through social media, digital assistants, and conversational systems, raises concerns about potential impacts on mental health. AI applications are increasingly used for interpersonal, advisory, and health-related purposes, yet their possible health-related consequences remain insufficiently clarified.

Objectives: This paper aims to systematically examine whether artificial intelligence can cause psychological harm or exacerbate existing mental health problems. A particular focus is placed on distinguishing direct causes of mental disorders from indirect effects on well-being and psychological health.

Methodology: A systematic literature review was conducted using the databases PubMed, Cochrane, Springer Link, Google Scholar, and Elsevier (Scopus). German and English search terms related to AI, psychological harm, and mental health were applied. Review articles, empirical studies, and reports from international organizations were included.

Results: The reviewed literature indicates that AI does not directly cause mental disorders in a clinical sense. However, evidence suggests that AI-based applications may indirectly affect psychological well-being through problematic usage patterns, increased screen time, social substitution, and algorithm-driven content amplification. Children and adolescents appear to be particularly vulnerable.

Conclusions: Artificial intelligence should not be considered a direct cause of mental illness, but it can function as an amplifying factor for psychological stress. Preventive strategies, education, and the responsible design and use of AI systems are essential to reduce potential risks to mental health.

Keywords: AI, artificial intelligence, psychological harm, mental health

1. INTRODUCTION

The field of artificial intelligence has existed since the 1950s. The term “Artificial Intelligence” was first coined at the Dartmouth Conference. Since the end of 2022, artificial intelligence has been accessible to everyone via the internet. Large amounts of data have been collected within internet-based systems, which AI can access. Various algorithms enable AI to learn, making its handling easier and its output more precise and accurate. By now, AI is capable of processing or generating images, videos, texts, and languages. Accordingly, it can speak, write, and explain, and it is also integrated into many mobile phones.

However, artificial intelligence is increasingly being asked interpersonal or health-related questions. For this reason, it is necessary to investigate whether artificial intelligence can cause health impairments or even exacerbate existing health problems.

2. State of Research

In 2022, a study on the health behavior of school-aged children (the HBSC study) was conducted. This study surveyed nearly 280,000 individuals aged 11, 13, and 15 years across 44 countries in Europe, Central Asia, and Canada (<https://www.who.int/europe/de>)[1]. In September 2024, the WHO Regional Office for Europe reported serious cause for concern regarding the impact of digital technologies—including artificial intelligence—on the mental health and well-being of young people [2]. The key findings of the study showed that more than one in ten adolescents (11%) exhibited signs of problematic use of social media. These adolescents had difficulties controlling their usage and reported negative consequences. Girls reported problematic use of social media more frequently than boys (13% compared to 9%). In addition, more than one third of young people (36%) reported being in constant online contact with friends. This proportion was particularly high among 15-year-old girls, 44% of whom stated that they were online almost constantly.

Digital gaming also plays a significant role in the daily lives of many adolescents: around one third (34%) played daily, and more than one in five (22%) spent at least four hours gaming on days they played. Overall, 12% of adolescents were considered at risk due to problematic gaming behavior. Boys showed such signs significantly more often than girls (16% compared to 7%).

Furthermore, the WHO Regional Office for Europe noted that previous studies had identified lower emotional and social well-being when social media use was problematic. If this trend continues, it could have far-reaching consequences for adolescent development and long-term health [2].

2.1 Health Damage

In Duden, the authoritative dictionary of the German language, health damage is defined as follows: “damage to health” [3]. The term *damage* has many facets:

1. “something that alters the given conditions or the existing situation in a negative, undesirable way”
2. A) “partial destruction; damage; defect”
B) “physical or health-related impairment”
3. “a loss resulting from the loss or [partial] destruction of an asset“ [3]

With regard to 2B, physical or health-related impairment, the following examples are given:

- severe psychological damage
- she has had damage to her eye [since birth]
- to sustain damage
- they were rescued from the burning house without damage (uninjured)
- fortunately, no one was harmed (injured) in the [3]

Based on this definition, acquired, external, and immediately visible damage is not included in the present investigation. This is because an AI, as a purely software-based system, has no independent physical agency. Without access to tools, machines, or possibilities for physical interaction, it is incapable of causing physical harm autonomously. Any damage can therefore not be attributed to the AI itself, but exclusively to human use or to the technical interfaces employed.

2.2 Psychological Damage / Mental Disorder

According to e.Medpedia (Springer Medizin), a mental disorder is defined as follows: „Eine psychische Störung setzt sich aus Symptomen zusammen, die als Störungen in den Bereichen Kognition, Emotionsregulation und Verhalten charakterisiert sind. Psychische Störungen sind Ausdruck dysfunktionaler – psychologischer, biologischer oder entwicklungsbezogener – Prozesse, die den seelischen Funktionen zugrunde liegen. Sie sind typischerweise mit bedeutsamem Leiden verbunden und wirken sich auf soziale, berufsbedingte, schulische und andere wichtige Aktivitäten beeinträchtigend aus.“[10] That means: “A mental disorder consists of symptoms characterized by disturbances in cognition, emotional regulation, and behavior. Mental disorders are expressions of dysfunctional—psychological, biological, or developmental—processes underlying mental functioning. They are typically associated with significant distress and impair social, occupational, educational, and other important areas of activity.”

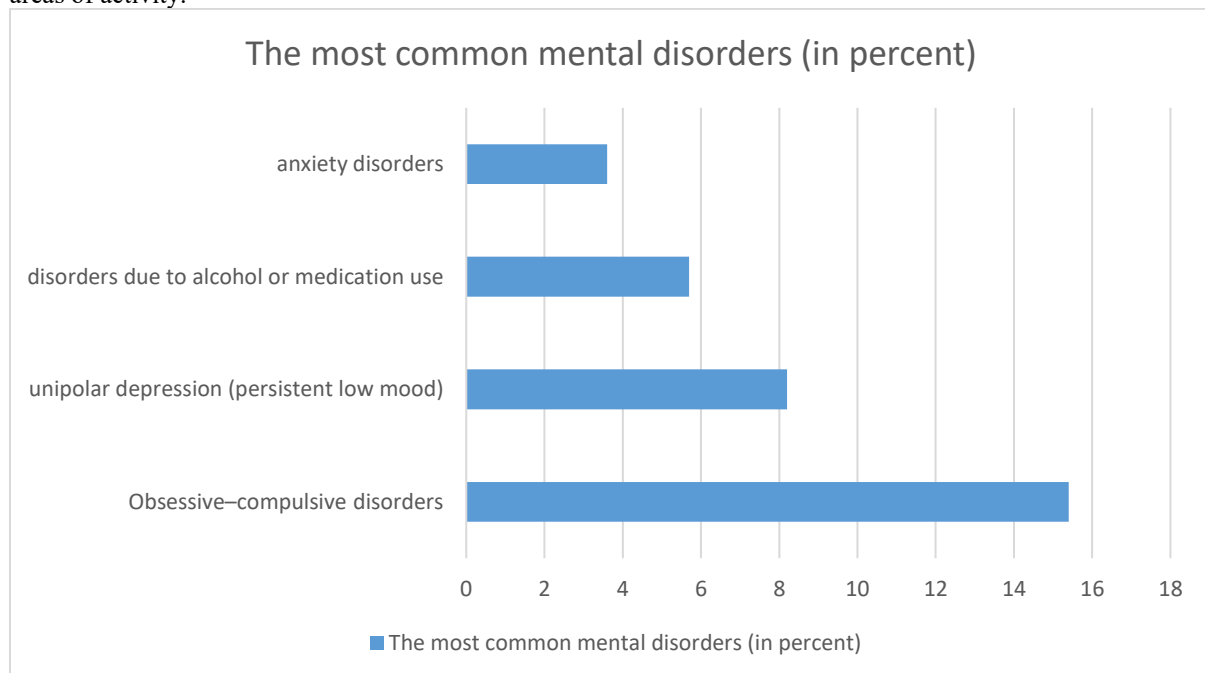


Figure 1: Mental disorders by type of disorder among adults in Germany [4]

2.3 Mental Health

The World Health Organization (WHO) defines mental health as an essential component of individual and collective well-being. Stressful or adverse life situations, working conditions and/or economic circumstances, as well as social inequality, experiences of violence, and conflict, are factors that influence mental health [5].

2.3 Artificial Intelligence

In the existing academic literature, there is no single, universally accepted definition of the term artificial intelligence. Instead, numerous—sometimes divergent—definitions and explanatory approaches can be found (cf. Wirtz et al., 2019, p. 597) [6]. This conceptual diversity is primarily attributable to the interdisciplinary nature of artificial intelligence research from its inception, as well as to the various developmental phases the field has undergone over time. Each of these phases has been closely associated with specific technological advances that have significantly shaped the prevailing understanding of artificial intelligence. In its early stages of development, AI-related systems were primarily based on rule-based logic, such as simple, table-based “if-then” structures (roughly comparable to Microsoft Excel). These systems followed fixed decision rules and were capable only of reproducing predefined patterns of action. As complexity increased, individual rules gradually evolved into a rudimentary system of automated knowledge processing and reproduction. Central to these systems was always the existence of an underlying data source to which they could gain access. With the establishment of the Internet as a comprehensive primary data source, the foundation was laid for feeding large quantities of heterogeneous information into such systems. This led to a significant expansion of the available knowledge base and opened up new possibilities for data-driven processing. A further decisive step in development was the integration of learning mechanisms that allow systems to identify patterns in existing data, generate new knowledge, and continuously refine this knowledge. Unlike purely rule-based approaches, such learning systems can dynamically improve their performance, provided that newly acquired knowledge is stored and linked with existing knowledge. From this combination of data availability, learning capability, and knowledge storage emerges a self-reinforcing dynamic, which can be understood as a core characteristic of modern artificial intelligence [7].

3. METHOD (SYSTEMATIC LITERATURE REVIEW)

In order to determine whether artificial intelligence can cause health-related psychological harm, various literature databases were searched for relevant academic publications. The databases selected for this review were PubMed, Cochrane, SpringerLink, Google Scholar, and Elsevier (Scopus). The search terms used were: *AI; artificial intelligence; psychological harm; mental health; KI; künstliche Intelligenz; psychische Schäden; psychische Gesundheit.*

	PubMed	Cochrane	Springer Link	Google Scholar	Elsevier (Scopus)
KI	124.299	237	563.318	6.050.000	67
Künstliche Intelligenz	306	1	25.710	265.000	18
psychische Schäden	2	20	11.632	78.600	2
psychische Gesundheit	659	94	45.564	322.000	8
AI	1.234.108	32	1.390.344	9.880.000	66
Artificial intelligence	345.948	9	476.472	7.890.000	5
psychological harm	8.557	301	141.570	2.790.000	0
mental health	650.241	617	781.092	6.410.000	4

Table 1: Database hits for the respective search terms

When conducting a literature review, it is important to filter out high-quality qualitative search results in order to better examine their relevance to the research question. Due to the large number of academic articles, the search terms were combined, and the following number of hits was identified:

	PubMed	Cochrane	Springer Link	Google Scholar	Elsevier (Scopus)
Künstliche Intelligenz psychische Schäden	0	0	697	16.400	0
Artificial intelligence psychological harm	66	1	10.695	419.000	0
Artificial intelligence mental health	7.048	1	37.209	3.570.000	0

Table 2: Database hits with combined search terms

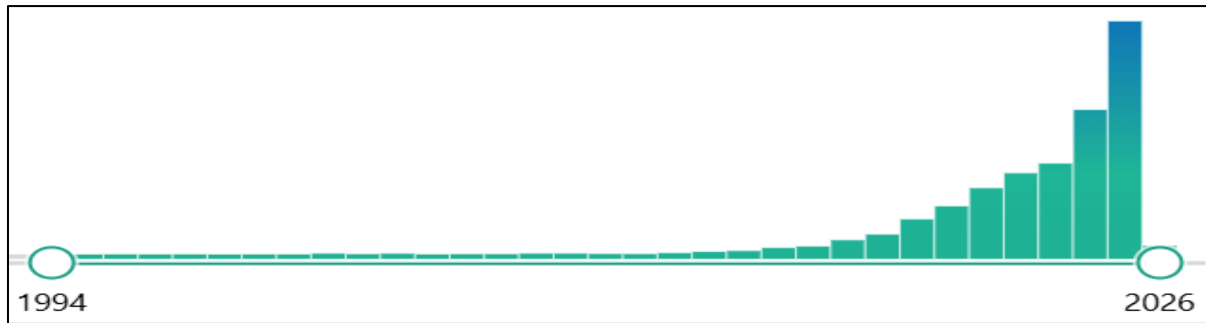


Figure 2: Publication years of the search term: „Artificial intelligence mental health“ on PubMed [8]

Figure 1 shows that the number of scientific articles published on PubMed on the topic “artificial intelligence mental health” has increased significantly since 2015. This rise appears to be driven by the growing integration of artificial intelligence into everyday life, as well as the increasing number of providers of AI systems.

The literature review revealed that the topic of mental health in connection with the use of digital and social media during childhood and adolescence is gaining increasing importance. It is necessary for children to learn how to engage with the digital world appropriately at an early stage. This can only be achieved through monitoring and guidance structures. Parents and teachers are primarily responsible for education and awareness-raising in this context. Attention must be paid to mental health, and various protective factors such as physical activity, regulated screen time, and interpersonal or social contacts should be promoted.

In the study “Impact of Social Media on Mental Health Among Adults” (2025), Batool, A. examined the relationship between social media addiction and the mental health of adults. In this cross-sectional study, 209 adults were surveyed regarding their social media use. The results showed that higher levels of social media use were significantly associated with poorer mental health outcomes. Regression analysis demonstrated that social media addiction is a significant predictor of deteriorating mental health. Batool, A. emphasizes the need for responsible use of social media and highlights the relevance of these findings for prevention strategies and mental health interventions [9].

As early as 2012, Deutsches Ärzteblatt addressed the issue of media addiction and discussed why it should be classified as an independent mental disorder. The article published in September 2012 focuses on the increasing relevance of problematic media use as a health-related phenomenon. Bühring, P. explains that excessive media consumption among children, adolescents, and adults can be associated with loss of control as well as negative effects on daily life, social relationships, and mental health. At the same time, diagnostic inconsistency is highlighted, as media addiction has not yet been clearly established within common classification systems, despite the recognition of comparable disorders. To justify its classification as an independent condition, addiction-like characteristics such as withdrawal symptoms, neglect of other areas of life, and persistent usage behavior are cited. In addition, psychological consequences such as sleep disturbances, concentration problems, and comorbid disorders are emphasized. The article concludes by stressing the need for improved diagnostic, therapeutic, and preventive care structures [9].

To expand the statistics, the table on AI psychological harm and AI mental health is used again. This table already provides a quantitative data basis.

	PubMed	Cochrane	Springer Link	Google Scholar	Elsevier (Scopus)
Artificial intelligence psychological harm	66	1	10.695	419.000	0
Artificial intelligence mental health	7.048	1	37.209	3.570.000	0

Table 3: Prepared Data for the Chi-Square Test

To calculate the descriptive statistics, the total number of hits is added.

Artificial intelligence psychological harm = 429,762

Artificial intelligence mental health = 3,614,258

	Treffer insgesamt	Prozentualer Anteil
Artificial intelligence psychological harm	429.762	11%
Artificial intelligence mental health	3.614.258	89%

Table 4: O = Observed values

Table 4 shows the observed values used in preparation for the chi-square test.

First, the total number must be calculated by adding all search engine results ($3,614,258 + 429,762 = 4,044,020$). The null hypothesis is:
 Ho: Both topic areas are equally represented in the literature.

Since two categories are given, an equal distribution (50% each) would confirm the hypothesis. The expected values are therefore $4,044,020 / 2 = 2,022,010$.

	Expected values
Artificial intelligence psychological harm	2.022.010
Artificial intelligence mental health	2.022.010

Table 5: E = Expected values

The difference between the observed and the expected values is now presented:

Mental health: $3,614,258$ (observed) – $2,022,010$ (expected) = $1,592,248$ (difference).

Psychological harm: $429,762$ (observed) – $2,022,010$ (expected) = $-1,592,248$ (difference).

To calculate the collected data, the chi-square formula is applied:

$$\chi^2 = \sum((O-E)^2/E)$$

First, the difference is calculated, then squared, divided by the expected value, and finally summed.

Category 1: $(1,592,248)^2 / 2,022,010 \approx 1,254,000$

Category 2: $(-1,592,248)^2 / 2,022,010 \approx 1,254,000$

Total chi-square value:

$$\chi^2 \approx 2,508,000$$

To determine the degrees of freedom, the following formula is used:

$$df = \text{categories} - 1$$

Two categories are given (mental health and psychological harm), therefore:

$$df = 2 - 1 = 1$$

The significance level is 5% ($p < 0.05$). This means that the probability that the observed difference occurred purely by chance is below 5%. In this case, $df = 1$. The corresponding critical values are:

$p = 0.05$ (critical value $\chi^2 = 3.84$), $p = 0.01$ (critical value $\chi^2 = 6.63$) and $p = 0.001$ (critical value $\chi^2 = 10.83$)

If χ^2 is greater than these values, the result is considered statistically significant.

According to the calculation, χ^2 is approximately 2,508,000.

Accordingly, $p \ll 0.001$, indicating that the difference is highly significant.

A chi-square test revealed a highly significant difference in the distribution of publications between the topics “Artificial Intelligence and Mental Health” and “Artificial Intelligence and Psychological Harm” ($\chi^2(1) \approx 2,508,000$, $p < 0.001$). Research on AI and mental health is substantially more prevalent than research on potential psychological harms caused by AI.

4. RESULTS

The analysis of studies identified within the systematic literature review shows that artificial intelligence does not itself cause immediate mental disorders; however, it can exert relevant indirect effects on mental health. In particular, AI applications in the context of social media, digital assistant systems, and personalized content are associated with changes in well-being. Several studies report that AI-driven recommendation algorithms may promote problematic usage behavior. This manifests in prolonged screen time, reduced self-regulation, and emotional dependence on digital systems. Children and adolescents are particularly affected. In addition, qualitative studies indicate that AI-based chatbots exhibit ambivalent effects: while they may provide short-term support, they also carry the risk of social substitution and isolation.

5. DISCUSSION

The results illustrate that AI is not inherently harmful to health. Rather, the risk of psychological impairment depends strongly on the context of use, duration, and individual vulnerabilities. Similar to social media, AI can present both risks and opportunities for mental health. AI systems designed to maximize user engagement should be viewed critically. At the same time, initial studies demonstrate positive effects of AI in the prevention and treatment of mental disorders. A prerequisite for this is an ethically reflected and professionally supervised application. A limitation of this study is the limited number of studies explicitly addressing psychological harm caused by AI. Future research should employ longitudinal designs to better understand causal relationships.

6. CONCLUSION

To address the question of whether AI can cause health-related harm, the AI system itself (ChatGPT) was asked: „kann die KI krank machen?“ (Can AI make you ill?).

The answer was clear: „Nein, eine KI kann dich nicht im medizinischen Sinn krank machen.

Aber sie kann Auswirkungen auf dein Wohlbefinden haben, wenn man nicht bewusst mit ihr umgeht“¹ (No, AI cannot make you ill in the medical sense. However, it can affect your well-being if it is not used consciously).

The preceding sections—2. State of Research, 2.1 Health Damage, 2.2 Psychological Damage / Mental Disorder, and 2.3 Mental Health—demonstrate that well-being is an indicator of mental health. Accordingly, unconscious and excessive consumption or interaction with AI may have a detrimental impact on mental health.

REFERENCES

1. World Health Organization, Hancock, J. et al (2024). Jugendliche, Bildschirme und psychische Gesundheit. URL: <https://www.who.int/europe/de/news/item/25-09-2024-teens--screens-and-mental-health> (accessed 06 January 2026)
2. Duden Editorial Office (n.d.): „Gesundheitsschaden“ on Duden online. URL: <https://www.duden.de/rechtschreibung/Gesundheitsschaden> (accessed 06 January 2026)
3. Duden Editorial Office (n.d.): „Schaden“ on Duden online. URL: <https://www.duden.de/rechtschreibung/Schaden> (accessed 06 January 2026)
4. World Health Organization (2021) Das Europäische Bündnis für psychische Gesundheit URL: <https://www.who.int/europe/de/initiatives/the-pan-european-mental-health-coalition> (accessed 08 January 2026)
5. Wirtz et al. (2019). Künstliche Intelligenz: Erscheinungsformen, Nutzungspotenziale und Anwendungsbereiche, S. 597 URL: https://web.archive.org/web/20220301072849id_/https://www.beck-elibrary.de/10.15358/0340-1650-2019-10-4.pdf (accessed 06 January 2026)
6. Großmann, K., Renger, F., (2024) Can Artificial Intelligence [AI] Support the Personnel Absence Management of Nursing Staff in Hospitals in Germany?. Journal of research in Nursing and Health Care, Issue No. 1, Vol. 1; pp. 01-06 URL: <https://aytinpublications.com/Journal-of-Research-in-Nursing-and-Healthcare/Files/NH010103.pdf> (accessed 08 January 2026)
7. Resch, F. (2021) Was ist eine psychische Störung? URL: https://www.springermedizin.de/emedpedia/detail/psychiatrie-und-psychotherapie-des-kindes-und-jugendalters/was-ist-eine-psychische-stoerung?epediaDoi=10.1007%2F978-3-662-49289-5_54 (accessed 08 January 2026)
8. PubMed figure URL: <https://pubmed.ncbi.nlm.nih.gov/?term=Artificial+intelligence+mental+health> (accessed 06 January 2026)
9. Batool, A. (2025) Impact of Social Media on Mental Health Among Adults URL: https://www.researchgate.net/publication/394102724_IMPACT_OF_SOCIAL_MEDIA_ON_MENTAL_HEALTH_AMONG_ADULTS (accessed 06 January 2026)
10. Suhr, F. (2020) Die häufigsten psychischen Erkrankungen. URL: <https://de.statista.com/infografik/22218/haeufigste-psychische-erkrankungen-in-deutschland/> (accessed 06 January 2026)
11. Deutsches Ärzteblatt, Bühring, P. (2012) Medienabhängigkeit: Eigenständige psychische Krankheit <https://www.aerzteblatt.de/archiv/medienabhaengigkeit-eigenstaendige-psychische-krankheit-7bfd056d-7f84-4c92-870f-aea77c11cfd2> (accessed 06 January 2026)