

Mastering the Digital Frontier: The Intersection of Generative AI and Human Rights in the Digital Age

*Syed Shaharyar Ahmed**

Abstract

The world has entered the digital age with the invention of Artificial Intelligence, and it is the harbinger of great implications for human rights. This paper explores these implications at the nexus of generative AI and digital human rights. This paper looks closely at how AI technologies, particularly generative AI, hold great promise but can sometimes threaten rights, such as privacy, freedom of expression, and non-discrimination. It uses Chat-GPT as a typical case example. It foregrounds the need for strict legal and ethical structures to ensure that the deployment of AI technologies reflects and advances human rights. This paper evaluates digital rights through the prism of human rights, with specific reference to international human rights law. This paper supports the opinion that a human-centred approach to AI, emphasising transparency, accountability, and inclusivity, is suitable. It also suggests making AI accountable and discusses possible approaches to receiving remedies should there be any breach of digital rights. The research will, therefore, help contribute to critical discussions on opportunities and challenges introduced by AI in the context of digital human rights. The paper finally attempts to propose ways to minimise the possible repercussions that AI might have on the enjoyment of digital and human rights.

Keywords: ChatGPT, Generative AI, Human Rights, Digital Rights.

Introduction

The global cyberspace has seen rapid progress in AI over the past three years. The amalgamation with Large Language Models referred to as GPT has changed the ways through which AI can be used. Perhaps the most noticeable feature of GPTs is the generation of new data and information from already obtained data that are fed

73

*Lawyer, Junior Partner, Ibrar & Associates, Lahore, Pakistan email: shaharyar.yargcu@gmail.com

Article History: Received; 18 January 2024; Received in revised form 12 June 2024; Accepted: 02 July 2024

Available online: 15 August 2024

to their servers (Nazir & Wang, 2023). This led to the development of the ChatGPT tool, an AI program designed for conversational interactions. Previously, it was not possible for users to directly communicate with AI software in a *two-way communication method*. However, since this conversational AI has been put into use, legal experts have raised questions concerning the ethical implications and possible violations of digital human rights. Digital human rights are those rights available to human beings when accessing digital cyberspace. Some prominent rights are the right to privacy, digital security, freedom of expression on social media platforms, the right to encryption, the right to anonymity, the right to data portability, the right to digital literacy and the right to online participation. There are numerous ways through which generative AI tools such as ChatGPT can trespass into the domain of any of these digital human rights. It is the responsibility of domestic and international regulatory authorities to ensure that data accessible to ChatGPT in no manner violates any of these digital rights (Marr, 2023).

AI programmes are not ethically trained, so they cannot assess whether utilising a certain individual's private data is against the right to privacy (Marr, 2023). The same is the case with the ChatGPT tool. ChatGPT has evolved over a shorter period. It started from GPT-1; now the latest global usage version is GPT-4.0. The launch of GPT 3.5 was a great achievement, however, the AI was confined to only responding to the form of text but with the recent launch of GPT 4.0, it has obtained the capabilities to generate text, audio, video and so on. It suffices to mention that such capabilities must be governed and put to work very carefully to ensure that they are not a source of violation of digital human rights. An example is: the GPT 4.0 tool with advanced audio and video generation might use any copyrighted intellectual property, such as a song, and generate anything out of it without the owner's will (Liu et al., 2023).

The principal objective of this research would be to evaluate the development of generative AI with respect to a ChatGPT tool and make a critical analysis of its usage and implications on digital human rights. Furthermore, this research sheds light on the requirement of sustaining and protecting digital human rights. Drawing on pertinent data related to generative AI, the author has tried to draw a fine line between what constitutes a fair use of the ChatGPT tool and what it means to constitute a digital human rights

violation. Lastly, the impact on third-world countries, compared to first-world countries, will be discussed with a special emphasis on digital human rights as protected under the legal framework of Pakistan.

Despite the growing importance of ChatGPT and its integration with many digital platforms, there has been no specific research evaluating the impact of ChatGPT on digital human rights. Although there is some literature exploring the general ethical dimensions of AI technology, there is still a lack of proper research addressing the specific challenges and opportunities which the ChatGPT presents.

There is currently no proper legal framework which is particularly designed to comprehensively address the direct impact of ChatGPT on digital human rights worldwide. Existing laws and regulations may not appropriately address issues such as the right to privacy, freedom of expression, freedom of speech, and algorithmic and logical bias prevalent in the use of AI technology such as the ChatGPT tool. The use of the ChatGPT tool raises major concerns regarding privacy issues such as the collection, storage, transfer and active usage of personal / user data. Therefore, extensive research is required to examine whether the ChatGPT interactions might violate the individual right to data privacy and data protection, particularly the use of sensitive information within the conversation.

ChatGPT raises difficult questions regarding freedom of expression and censorship in supporting communication and the dissemination of information. There are no concrete scientific findings that could establish how the ChatGPT tool algorithms limit or affect users in case of full self-expression online; it could affect freedom of speech and public debate.

Research has shown that, even in the case of artificial intelligence tools like ChatGPT, algorithmic bias has, over time, served to promote discrimination against some minorities or disadvantaged groups. Further studies are required to assess bias in the ChatGPT replies and their effects on the marginalised community along with measures that can be taken to avoid them to guarantee equal opportunities in using the digital environments (Nazir & Wang, 2023).

Another problem is that it affects people's health and social lives, which is the consequence of the extensive use of ChatGPT (Shany, 2022). There is, therefore, a need for research to find out how the

long-term use of the ChatGPT tool impacts the user's perceptions, attitudes, and social interactions, including the social impacts of sending the studies of intelligence agencies' communications (Wu et al., 2023). Therefore, this research is important for developing laws, guidelines and interventions regarding the use of technology to protect the digital rights of the people in society since artificial intelligence is getting stronger each day.

Research Methodology and Operational Framework

Qualitative, deductive and descriptive research methodologies have been utilised for this research. The primary sources include scholarly articles, relevant statutory instruments (where applicable), and international agreements. The research is categorised under the following heads: in the first part, a brief overview of the concepts of generative AI and digital human rights has been provided. In the second part, the history of generative AI, evolution of ChatGPT and its implications for the digital cyberspace have been discussed. In the third part, the concept of digital human rights along with regulatory mechanisms and agencies responsible for ensuring safety in digital cyberspace has been explained. In the last part, the impact of ChatGPT as a generative AI tool on the enjoyment of digital human rights has been discussed, with possible recommendations and suggestions for minimising the consequences.

Literature Review

ChatGPT is the formal manifestation of a generative AI language model designed by Open AI. This tool is a conversational AI which uses larger data sets to respond to queries put forth. It is a smart chatting AI tool which is capable of comprehending and generating responses in various languages (Nazir & Wang, 2023). It has been improved over time and the latest operating version of ChatGPT is 4.0 but only the ChatGPT 3.5 is available for free use. ChatGPT 3.5 was able to only converse in textual form while the later version is capable of processing and responding to data in the form of audio, video, text and much more. ChatGPT has some shortcomings which are needed to be further improved. Some of these drawbacks include reliability, consistency, limited data set availability, explicit knowledge monitoring, and extraordinary costs

of research and development (Wu et al., 2023). The first-ever demo of ChatGPT was released as early as the 30th of November, 2022, which quickly went viral on several social media platforms. It grabbed the attention of more than one million users in a short span of mere five days. A whopping 175 billion parameters were used in the data set instruction training for the development of the GPT-3 version. ChatGPT has redefined the concept of artificial intelligence, thereby proving the fact that even language models and machines can “learn” human-oriented language and interact in a two-way manner (Marr, 2023).

Nazir and Wang (2023) have discussed the ethical considerations regarding ChatGPT. They highlighted that generative AI tools such as ChatGPT created adverse ethical consequences which were related to the digital human rights such as the right to privacy, digital security, transparency, anonymity as well as user awareness. They have provided an example of biased response generation by ChatGPT depending on the data set which is fed into such Generative AI. For instance, if a data set is prejudicial towards a particular community, then it is more likely that responses generated by ChatGPT will be biased and prejudiced too.

Shany (2022) mentioned the hurdles in the adaptation of international human rights to digital cyberspace. The principles governing the right of freedom of expression are somewhat similar in the online / digital application as well. Freedom of speech and expression should, therefore, be strictly guarded and respected on social media platforms like Facebook, Instagram and others. Furthermore, it is the responsibility of domestic and international authorities to ensure that state citizens and the populace of a global village are able to enjoy these human rights in the digital world.

Javaid, Haleem and Singh (2023) defined a Generative Pretrained Transformer (GPT) as a sort of Artificial Intelligence language model which is prepared to produce human language text. It is a language model which has a pre-trained neural network. ChatGPT is a tool which uses this GPT architecture and it has been created by OpenAI. ChatGPT has been trained with a much larger data set than other language models, hence making it the most popular GPT tool. They also argued that ChatGPT has caused negative consequences in its use in the healthcare industry. It has compromised the confidentiality of patient information, which is the basic element of a trust-based relationship between doctor and

patient. They are of the view that consistent unchecked and unregulated usage of ChatGPT will lead to breaches concerning data security and data privacy.

Wu, Duan and Ni (2023) explained that the integration of ChatGPT into daily lives has given birth to several privacy, security and ethical issues. ChatGPT is now being misused for negative actions such as for the generation of hate speech, fake news and propaganda. They think that it is imperative to address these concerns as soon as possible. They have pointed out the use of ChatGPT in e-commerce and the healthcare sector. It has rendered positive benefits for them but this tool needs regulation and that too within due time.

ChatGPT and its Impact on the Right to Education

As far as the academic community is concerned, the development of the ChatGPT tool has received two major viewpoints. People who support ChatGPT are of the opinion that the inclusion and usage of ChatGPT will lead to improved work quality and work delivery by helping out humans with certain tasks. These tasks include but are not limited to; writing emails, editing videos, script writing, copywriting, language translation, answering queries as a chatbot etc. The use of ChatGPT in the education sector can help to democratise education and support diversity in students' participation in learning through providing an easily available and personalised learning environment. It can ensure quality education which is in light with the fundamental human right to education (Malik, Khan & Hussain, 2023).

However, on the other hand, people who oppose the use of ChatGPT strictly argue that the content generated by ChatGPT is prone to scrutiny (Malik, Khan & Hussain, 2023). It will be difficult for an average human being to differentiate between content written by a human and the one generated by AI tools like ChatGPT. The elements of genuineness, novelty, creativity and integrity are at risk if we deploy the ChatGPT tool for common practice. In academia, this translates into ChatGPT-generated assignments and answers used by students to gain maximum marks. This would seriously hurt the integrity of results and the quality of assessment methods (Rasul et al, 2023).

The equal access of the ChatGPT tool to students is another concern. The students who don't have access may be left behind compared with the students who have access to all the modern technological equipment. The equitable access to technology becomes a challenge and it becomes the duty of higher education institutions to ensure that equal access is available to every student (Rasul et al, 2023). This is a requirement for the fulfilment of the right to education, equality and equal treatment. The positive use of ChatGPT can fasten the self-learning process through which students can learn on their own in a manner which suits them the best (Rasul et al, 2023). Educators and teachers can integrate ChatGPT into their teaching, instruction and evaluation methods to make the classroom environment more inclusive and practical. The integration of ChatGPT can lead to the automation of tasks, freeing up the teacher's time in generating effective content. Nonetheless, ethical concerns are still there regarding intellectual property ownership, academic integrity, privacy issues and various potential biases which are required to be addressed when utilising such AI-based technologies in academic institutions (Malik et al, 2023). Therefore, it is the duty of policymakers, legal experts, lawmakers, and educators to ensure the responsible, regulated, and ethical use of AI technologies such as the ChatGPT tool in an academic environment. All the stakeholders should come together and devise universal guidelines for proper use, providing effective training to the teachers, and integrating AI technology in the overall curriculum with an appropriate balance and consistent pace, so it can be ensured that AI tools like ChatGPT are being used in accordance with such guidelines and they align with an institution's academic objectives so that the right to education can be guaranteed (Malik et al, 2023).

ChatGPT and its Impact on the Right to Life

ChatGPT has several uses in the healthcare sector. It has led to improving adherence to treatment plans and providing access to healthcare management professionals so that they can monitor patients who are residing in remote or underrepresented areas. Furthermore, ChatGPT also provides a rapid production of medical reports which are often delayed in the usual routine. ChatGPT can effectively take over some human tasks which will spare more time for doctors and the hospital staff to attend to patients. The tasks of

record-keeping, data management and transcription of medical records can be taken over by ChatGPT. Patients can also benefit from the ChatGPT tool by obtaining guidance. This will reduce their tension, anxiety and confusion as they get along with the diagnosis of their health condition. ChatGPT can also be used to share data between the insurance sector and health care. Insurance providers and healthcare professionals can get timely access to integrated information regarding a patient in one single database by deploying the ChatGPT tool over such a dataset (Javaid, Haleem & Singh, 2023). Thus, there are ways and methods through which a beneficial use of ChatGPT can be made. ChatGPT can offer intelligent support for decision-making, the capability of translating different medical texts from one particular language to another, and the ability to communicate between caregivers and patients (Javaid, Haleem & Singh, 2023).

ChatGPT can turn out to be very transformational in the healthcare sector since it has the potential to understand natural language-based questions and respond in appropriate contexts. It can be controlled with administration for functions such as the triage of the patient, medical diagnosis, and prescription recommendation. It can also help in health education by answering questions about the symptoms, medical conditions and treatment of various diseases. The ChatGPT tool can also generate health tips, lifestyle recommendations, and proper mental health care support (Marr, 2023). The ChatGPT can also work as a virtual doctor, with a patient's input of the symptoms given out to the system and responded to. It can also provide medication reminders which can alert patients to schedule follow-up appointments. Patients can also utilise the ChatGPT to access their medical records, keep track of their health status as well as report side effects of any medication (Moore, 2023).

However, it is also pertinent to mention here that the ability of ChatGPT to access and generate accurate data is dependent upon the data quality and accuracy fed into its data set (Milmo, 2023). Therefore, ChatGPT needs to be improved and updated by providing the latest information and data available regarding healthcare so that it can provide accurate results which could cause any misinformation leading to potential safety risks for patients. ChatGPT often uses unreliable sources of information and data such as Wikipedia, Medium, and LinkedIn for the generation of

responses which underscores the importance of having a data integrity regulation mechanism (Milmo, 2023). Using a virtual doctor service in the form of a chatbot based on ChatGPT infrastructure is also risky due to the possibility of erroneous results. Patients can be wrongly diagnosed. The privacy of patients' records could also be at risk. Since the data is being accessed and used by ChatGPT there are chances of a data breach or ChatGPT using one patient's data to diagnose and treat another patient. ChatGPT is an emerging technology and it offers many significant benefits. However, it cannot in any way serve as a substitute for human expertise. Therefore, in this manner, the fundamental human right to life as well as the digital human right to privacy and data security are under threat of being compromised (Nerozzi, 2023).

Digital Human Rights

Digital human rights concern the protection of human rights in the digital space. These rights include the right to privacy, freedom of speech, and equal access to technology and information. Digital citizenship means the responsible use of digital technology, which includes engagement in ethical online behaviour, protection of a person's privacy, and being aware of one's digital rights and responsibilities (Nerozzi, 2023). Digital literacy is the capacity to efficiently utilise and critically examine digital media. This also includes various technical skills along with the ability to reason, to think critically, and to communicate thoughts through digital tools. The concepts of digital literacy, digital citizenship and digital rights are deeply interconnected (Fui-Hoon Nah, 2023). The development of AI and the creation of the ChatGPT tool has influenced the function and overall nature of these concepts. There are various case laws of South African and European Courts which highlight the evolution and importance of digital human rights such as: -

- In the case of *Amnesty International Togo v. The Togolese Republic (2020)*, the Community Court of Justice of the Economic Community of West African States (ECOWAS) declared that the shutting down of the internet during the protests in September 2016 amounted to a violation of the fundamental right to freedom of expression. The Court was of the view that access to the internet is a “*derivative right*” as it is

responsible for “*enhancing the exercise of freedom of expression.*” Therefore, “*Internet access is a right that requires the protection of the law and any interference with it must be provided for by the law specifying the grounds for such interference.*”

- Digital technologies have impacted the arena of human rights. The development of ChatGPT has led to the recognition of certain digital rights which are being considered to be threatened by the unregulated usage of this tool. In the case of ***Jamii Media v. The Attorney General of Tanzania and Other (2017)***, the petitioner was a website operator who used to provide users a platform for posting content anonymously and engaging in debates concerning social, political or economic aspects without being disclosed. However, in pursuance of the Cybercrimes Act 2015, orders were issued by the police which demanded the complete disclosure of information concerning the users of that platform, and then the police threatened the petitioner that he would be prosecuted if he did not comply. The matter went up to the High Court of Tanzania. The Court declared that the right to data privacy and user anonymity is recognised, however, there are certain exceptional circumstances in which a government’s request for a “*disclosure of user information by an internet intermediary may be justified.*”
- The right to privacy is another fundamental human right which has been protected under various international instruments such as the International Covenant on Civil and Political Rights (ICCPR) and the Universal Declaration on Human Rights (UDHR). The concept of conducting mass surveillance contrasts with the right to privacy. This debate was finally settled in the landmark case of ***10 Human Rights Organisations v. The United Kingdom (2021)***. The case began in the UK around 2013 when a person named Edward Snowden revealed that the UK’s GCHQ had been secretly intercepting, processing and storing sensitive data which included millions of people’s private communications. This data collection was extensive enough not to exclude people

who were not directly linked to any intelligence interest such as “The Tempora Programme”. The UK government was conducting mass surveillance of people belonging to the UK and beyond without any authorisation or public consent having no basis in law and without adequate safeguards. The applicants took the matter to the Court. They argued that the practices of mass surveillance conducted by the UK are in grave violation of Articles 8 (right to privacy), 6 (right to a fair trial), 10 (freedom of expression) as well as 14 (prohibition on discrimination) of the European Convention on Human Rights (ECHR). In 2018, the *First Section of the European Court of Human Rights* declared that the UK laws which enable mass surveillance regime were violative of the rights to privacy and freedom of expression provided under ECHR. The Court referred the matter to the Grand Chamber for a decisive ruling. On 25th May 2021, the *Grand Chamber of the European Court of Human Rights* confirmed that UK mass surveillance laws breached the right to privacy and freedom of expression. The ruling condemned mass surveillance conducted by the UK government. This judgment will have a long-lasting impact as it will not only affect the UK but the entire European region and beyond. The deployment of the ChatGPT tool in conducting mass surveillance would have proved to be dangerous and since mass surveillance itself is against the law, therefore, ChatGPT should not be used to conduct mass surveillance and gather private data. This is one of the major future implications of ChatGPT on digital human rights that it can be used as a medium to conduct effective mass surveillance.

- The ChatGPT tool is being deployed in conducting digital surveillance by some states while digital surveillance itself is an uncharted territory. For instance, in the case of *AmaBhungane Centre for Investigative Journalism NPC and Another v. Minister of Justice and Correction Services and Others (2021)*, it was declared by the Constitutional Court for South Africa

that the state's bulk interception legislation is unconstitutional owing to several aspects, which include the lack of notification of a person who is under surveillance. A positive notice is required and insufficient measures concerning the management and security of such information are not in accordance with the law.

The international human rights law framework is rapidly changing owing to its interaction with digital human rights, which has generated new possibilities for the promotion of human rights in the digital era. Furthermore, the territorial nature of the human rights enforcement framework is being changed since people are now becoming global citizens. Existing human rights are being reinterpreted to meet the needs of the digital age, which is leading to the recognition of new rights and duties. For instance, the evolution of digital cyberspace has led to the recognition of the right to privacy on the internet, the right to anonymity, the right to data security, the right to data integrity and protection from hacking etc. (Shanty, 2023).

Positive Impacts of ChatGPT on Digital Human Rights

Generative AI such as the ChatGPT tool has impacted digital human rights both positively and negatively. As far as the positive implications are concerned, the ChatGPT tool has led to open access to information concerning human rights. This access to information has also enhanced the reporting and investigation mechanism of human rights violations globally, which is a positive step. The AI algorithms are capable of analysing huge amounts of data from a plethora of sources such as social media, webpages etc. They can also detect certain patterns that remain unnoticed under human scrutiny (Shanty, 2023). Therefore, AI can provide new insights regarding the human rights issues which might have gone unchecked. ChatGPT tool can also enhance global communication networks by translating text into a real-time conversation. Laws and information governing human rights can also be translated by the ChatGPT tool which will remove the language barrier previously faced by people who wanted to obtain information regarding human rights. It will make the dissemination of digital human rights

information much easier for those facing a language barrier (Rodríguez-Garavito, 2023).

Access to information is quite valuable for the marginalised segments of society as it can help them realise their human rights. ChatGPT tool also provides support to human rights activists who need data analysis concerning human rights abuses (Rodríguez-Garavito, 2023). ChatGPT tool can be used as a chatbot by human rights organisations to provide an immediate response for the victims whose rights have been infringed. That chatbot can provide practical solutions by guiding the victims regarding appropriate forums to contest their cases and obtain relief. Policymakers and lawyers working for NGOs can utilise the ChatGPT tool to scrutinise data dealing with social problems, discrimination and violation of human rights (Haristiani, 2019).

Negative Impacts of ChatGPT on Digital Human Rights

One of the major potential implications of the ChatGPT tool on digital human rights is the risk of inherent prejudice and discrimination. If the data set governing or regulating AI models contains biases or discriminatory content, it can lead that AI model to behave in that particular way since AI models work based on the machine learning process (Tripathi, 2023). Therefore, the AI model can learn from the content that is fed into its neural networks. This would lead to a situation in which persons belonging to a certain class will face severe discrimination based on their gender, race, ethnicity, social status etc. Furthermore, the job security of human beings is in grave danger. ChatGPT tools are being incorporated as Chatbots in many major multi-national companies which means that the human beings hired for customer service or chatting agent jobs would be replaced with Chatbot tools. This paradigm shift will lead to job displacement, unemployment and poverty, which can cause great instability in the economic landscape of a country (Shanty, 2023). Moreover, large-scale data collection and data analysis also amount to a breach of data privacy rights. AI models such as the ChatGPT tool can also be used in decision-making processes with significant implications for human rights, such as determining bail amounts or guilt or innocence of the accused persons; if these tools are inherently prejudiced, they can lead to an unjust outcome against human rights (Tripathi, 2023).

Another implication of the ChatGPT tool on digital human rights is the potential dissemination of fake news and tampered information. ChatGPT tool solely relies on the information that is fed to it or made available in its data set and may not be able to counter-check or verify whether the information is trustworthy (Lausevic, 2023). This can lead to the potential dissemination of propaganda, various conspiracy theories, or even fake news that can cause violations of civil and political rights, such as the right to obtain accurate information or the right to participate in free and fair elections. Moreover, the strict reliance on AI models such as the ChatGPT tool can lead to unwanted consequences, like dependence on decisions obtained through the algorithmic process which can violate the right to due process or fair trial when a person is tried by a competent court relying on ChatGPT for its guidance. Therefore, strict accountability and transparency are required in the deployment and consistent usage of AI models, so that its negative consequences on digital human rights can be minimised (Lausevic, 2023).

Lausevic (2023) mentioned that when the ChatGPT tool was asked about its contribution towards human rights, it provided that the impact of its technology has led to increased access to information and enhanced data analysis. It also acknowledged that AI has posed three major threats to human rights. These are the reproduction of biases present in the previous data fed to it, violation of privacy through utilising user's data without permission and enabling the abuse of data without any accountability. As per the responses obtained, it became evident that the potential of generative AI such as ChatGPT tool to increase misinformation and fake news is getting stronger (Tripathi, 2023). The definition and interpretation of fundamental human rights can also be challenged by the ChatGPT tool through its ability to manipulate public opinion, generate false narratives, and sway social and political outcomes. This can cause severe trouble for the democratic systems to protect human rights. Therefore, lawmakers and policy analysts need to evaluate how tools like ChatGPT are capable of challenging human rights while devising possible methods to reduce these risks and promote human rights (Lausevic, 2023).

Deep fakes are AI-generated audio and video-based information. ChatGPT tool can generate content which is accurate apparently but can directly threaten the integrity of evidence.

ChatGPT 4.0 is being used by hackers and malicious users to generate deep fakes to blackmail people (Dash & Sharma, 2023). It can also lead to a new wave of next-generation propaganda and dissemination of fake news through artificially generated videos, audio and pictures. This possibility has increased the responsibilities of NGOs such as the Human Rights Watch, to counter-check and verify the information circulating on social media before it causes an irreparable loss or unwanted consequence (Bacciarelli & Aufiero, 2023).

Suggestions and Recommendations

The following suggestions and recommendations can be implemented in order to effectively utilise generative AI:

- Governments and international organisations need to collaborate to establish a strong legal system which specifically deals with the ethical regulations governing AI technologies such as the impact of the ChatGPT tool on digital human rights. The legislative framework should involve provisions regarding personal protection, freedom of expression, principles of non-discrimination and algorithmic accountability. Both national and international stakeholders need to establish a robust legal framework for governing the usage and function of generative AI. This initiative will aid in the protection of digital human rights. The future legislation needs to deal with matters concerning privacy, freedom of expression as well as the issue of non-discrimination which is linked with AI tools like ChatGPT.
- International collaboration is required among all the states to enter into a convention or treaty with the objective of developing a mutual understanding and drafting rules governing the use of AI tools like ChatGPT which are universal in nature. Furthermore, a global perspective is required to ensure effectiveness in addressing the cross-border implications of AI on digital human rights.

- There should be transparency as well as accountability in the working and monitoring of AI systems. The developers of AI technologies should be obliged to disclose the algorithms and mechanisms behind the decision-making process of tools like ChatGPT so that it can be ensured that there is no inherent discrimination or bias fed into the tool itself. This software-level accountability will help prevent any adverse impact on digital human rights. Therefore, it is the responsibility of the developers and service providers to ensure adherence to transparency and accountability. This includes the disclosure of information regarding the training of the ChatGPT algorithm and the identification of potential biases. The creation of a process for users to report the nature of the breach or violation is also required.
- Initiatives should be promoted to empower the users so that they can understand and utilise the ChatGPT tool in an effective manner. The users should be made aware of the possible biases of the tool, and the possibility that the information generated by such a tool might not be accurate. Business participants must comply with the ethical codes of conduct for the development and use of intellectual property when using generative AI. This process must primarily protect the rights and freedoms of users, ensure the fairness and justice of algorithmic decisions, and protect the principles of human dignity and freedom of speech.
- All of the AI applications should be strictly regulated. For this purpose, the implementation of a sector-specific regulatory mechanism would be an optimal approach. Therefore, tools like ChatGPT should be separately regulated. Manufacturers must actively work to identify and eliminate prejudice, reducing bias in the ChatGPT algorithm to prevent discrimination against marginalised communities. This will include differentiating training dates, using bias detection tools,

and performing regular audits to assess the integrity of the algorithm.

- Regular audits and assessments should be conducted of the popular AI applications so that any non-compliance or breach of any legal or ethical standard might be addressed at the earliest. This should involve hiring independent firms, which specialise in the identification and rectification of potential biases, to ensure compliance with legal and ethical standards.
- Users must be provided with clear information about how the ChatGPT tool works, including its limitations and potential risks to individuals' digital rights. Users should be allowed to exercise control over every aspect of their interaction with the ChatGPT tool including the ability to turn off certain functions and request features to provide explanations for algorithmic decisions.
- Organisations implementing ChatGPT should conduct human rights impact assessments regarding policies and procedural practices regardless of user practice. This would provide a roadmap in terms of mitigation and would also ensure that ChatGPT transmission is aligned with human rights.
- There is a need to establish training programs to instil an understanding of the consequences of interacting with ChatGPT, particularly in terms of privacy, biases, and the importance of analysing AI's content for the users.
- It is important to conduct regular research to monitor the changing impact of ChatGPT on digital human rights and identify challenges and opportunities. Interdisciplinary research efforts should explore the psychological, social, legal, and ethical aspects of the ChatGPT tool, to inform evidence-based policy and practice.
- There should be a mechanism to ensure the quality and accuracy of the data used by AI tools to control the spread of discriminatory content. Furthermore, user consent regarding data usage by ChatGPT should be strengthened and strictly regulated. Heavy fines should

be imposed on companies violating the user consent standard. ChatGPT should not be allowed to process private users' data without explicit and authorised consent.

- The implementation of digital literacy programmes will help educate the users about the limitations and capabilities of AI technologies. This will assist the users to know their digital rights and to safely navigate in AI-dominated cyberspace. Professionals should be trained in using ChatGPT and other similar tools. Conducting special training programmes is an effective method for achieving this purpose.

Conclusion

Considering the positive and negative impacts which the ChatGPT tool might have on digital human rights, it can be said that the implications of ChatGPT on digital human rights underscore the dire need for a comprehensive strategy to navigate and address numerous opportunities and challenges provided by the AI technologies. Although the ChatGPT tool provides access to unparalleled information and content, it demands urgent scrutiny over how these can harm digital human rights, such as the right to privacy, freedom of expression, and the right to non-discrimination.

Given its multiple influences on digital human rights, it is important to consider it carefully from the perspective of various stakeholders such as governments, non-governmental organisations, academicians, and businesses. With further penetration of ChatGPT into the digital communication platform and its takeover of online communication, there will be open avenues, along with associated challenges, for discussions on the right to privacy, freedom of speech, non-discrimination, and freedom of use.

In a digital environment that is constantly on the move, it helps to think in terms of the human environment while regulating and keeping a check on the use and development of generative AI. A balancing framework between technological innovation and the protection of human rights has to be orchestrated through strong legal frameworks codified with safeguards and ethical and legal guidelines under a transparent accountable system.

The ChatGPT application should be used for the greater good, employing its data processing and analysis abilities towards the service and advocacy of human rights. On the other hand, the reduction of the negative implications is imperative to achieve the maximum output. This can be done by balancing issues concerned with biases, privacy breaches, user anonymity, and data integrity that require global, agreed and collaborative efforts from all the states of the world. While we have to manage an information technology revolution underpinned by artificial intelligence, we also have to be alert to rising challenges, continuously evaluate the impact of the protection afforded by ChatGPT to people's digital rights, and fine-tune policies and practices.

Embedding human rights in the development and deployment of ChatGPT would draw out the transformational potential of the technology while ensuring that its risks are oriented towards underpinning a digital ecosystem consistent with the rights and dignity of all individuals. There needs to be a global agreement toward the standardisation of ethical norms. States must begin to invest in standard ways of researching generative AI. It could be one way of being proactive toward this paradigmatic transformation and keeping states well-equipped with the necessary tools.

Therefore, the international community needs to support the responsible deployment of AI, as this approach is valid and requires affirmative measures to ensure a secure future where generative AI coexists with human digital rights.

References

- AmaBhungane Centre for Investigative Journalism NPC and Another v Minister of Justice and Correctional Services and Others; Minister of Police v AmaBhungane Centre for Investigative Journalism NPC and Others (CCT 278/19; CCT 279/19) [2021] ZACC 3; 2021 (4) BCLR 349 (CC); 2021 (3) SA 246 (CC) (4 February 2021)
- Amnesty International Togo v The Togolese Republic (2020, ECW/CCJ/JUD/09/20)
- Bacciarelli, A., & Aufiero, P. (2023, May 3). Pandora's Box: Generative AI Companies, ChatGPT, and Human Rights.

Retrieved January 18, 2024, from <https://www.hrw.org/news/2023/05/03/pandoras-box-generative-ai-companies-chatgpt-and-human-rights>

Dash, B., & Sharma, P. (2023). Are ChatGPT and Deepfake Algorithms Endangering the Cybersecurity Industry? A Review. *International Journal of Engineering and Applied Sciences (IJEAS)*, 10(1), 2394-3661. <https://dx.doi.org/10.31873/IJEAS.10.1.01>

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277–304. <https://doi.org/10.1080/15228053.2023.2233814>

Haristiani, N. (2019). Artificial Intelligence (AI) Chatbot as Language Learning Medium: An inquiry. *Journal of Physics: Conference Series*, 1387(1), 012020. <https://doi.org/10.1088/1742-6596/1387/1/012020>

Jamii Media Company Ltd v. The Attorney General of Tanzania and other (2017) TLS LR 447

Javaid, M., Haleem, A., & Singh, R. P. (2023). ChatGPT for healthcare services: An emerging stage for an innovative perspective. *Bench Council Transactions on Benchmarks, Standards and Evaluations*, 3(1), 100105. <https://doi.org/10.1016/j.tbench.2023.100105>

Lausevic, V. (2023, May 28). ChatGPT and human rights? Retrieved January 17, 2024, from Vlad's politics website: <https://www.publish0x.com/vlads-politics/chat-gpt-and-human-rights-xgjkyln>

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., . . . Ge, B. (2023, April 4). Summary of ChatGPT-Related Research and Perspective Towards the Future of Large Language Models. <https://doi.org/10.1016/j.metrad.2023.100017>

- Malik, A., Khan, M. L., & Hussain, K. (2023, April 18). How is ChatGPT Transforming Academia? Examining its Impact on Teaching, Research, Assessment, and Learning. <https://doi.org/10.2139/ssrn.4413516>
- Marr, B. (2023, May 19). A short history of ChatGPT: How we got to where we are today. *Forbes*. Retrieved from <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=6c13d2c7674f>
- Milmo, D. (2023, February 2). ChatGPT reaches 100 million users two months after launch. Retrieved from <https://www.theguardian.com/technology/2023/feb/02/chat-gpt-100-million-users-open-ai-fastest-growing-app>
- Moore, S. (2023, March 27). What does ChatGPT mean for Healthcare? Retrieved January 15, 2024, from <https://www.news-medical.net/health/What-does-ChatGPT-mean-for-Healthcare.aspx>
- Nazir, A., & Wang, Z. (2023). A Comprehensive Survey of ChatGPT: Advancements, Applications, Prospects, and Challenges. *Meta-radiology*, *1*(2), 100022. <https://doi.org/10.1016/j.metrad.2023.100022>
- Nerozzi, T. (2023, February 10). Bill Gates says ChatGPT will “change the world,” make jobs more efficient. Retrieved January 15, 2024, from <https://www.foxbusiness.com/technology/bill-gates-says-chatgpt-will-change-world-make-jobs-more-efficient>
- Pangrazio, L., & Sefton-Green, J. (2021). Digital Rights, Digital Citizenship and Digital Literacy: What’s the Difference? *Journal of New Approaches in Educational Research*, *10*(1), 15. <https://doi.org/10.7821/naer.2021.1.616>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., Oliveira Santini, F. D., Ladeira, W. J., . . . Heathcote, L. (2023, May 10). The role of ChatGPT in higher education: Benefits, challenges, and

future research directions | Journal of Applied Learning and Teaching. <https://doi.org/10.37074/jalt.2023.6.1.29>

- Rodríguez-Garavito, C. (2023, May 25). ChatGPT: What's left of the human in human rights? Retrieved from OpenGlobalRights website: <https://www.openglobalrights.org/chatgpt-whats-left-human-rights/>
- Shany, Y. (2022, October 23). Digital Rights and the Outer Limits of International Human Rights Law: *German Law Journal*, *Hebrew University of Jerusalem Research Paper No. 23-2*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4270794
- Tripathi, S. (2023, March 7). We asked ChatGPT about its impact on human rights and business. Here's what it told us |. Retrieved January 17, 2024, from <https://www.ihrb.org/focus-areas/information-communication-technology/we-asked-chatgpt-about-its-impact-on-human-rights-on-business-heres-what-it-told-us>
- Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q., & Tang, Y. (2023). A brief overview of ChatGPT: the history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, *10*(5), 1122–1136. doi: 10.1109/jas.2023.123618
- Wu, X., Duan, R., & Ni, J. (2023, July 26). Unveiling Security, Privacy, and Ethical Concerns of ChatGPT. Retrieved from <https://arxiv.org/abs/2307.14192v1>