AI and Human Rights in Healthcare

October, 2023

BSR®

This report examines how artificial intelligence (AI) technologies are driving change within the healthcare industry and the associated human rights challenges and opportunities.

As the healthcare industry continues its digital transformation, providers need to consider the impacts of AI for three main reasons:

Human RIghts

Technological transformation brings complex, nuanced, and systemwide risks and opportunities for the realization of human rights. These risks and opportunities are related to both the **design** and **development** of technologies as well as how technologies are **deployed** and **used** by companies.

Evolving Regulatory Environment

Changes in the regulatory landscape, including the EU's proposed Corporate Sustainability Due Diligence Directive and Artificial Intelligence Act, signal that companies outside of the technology industry will need to have a better understanding of the human rights impacts of the AI solutions they deploy. It is noteworthy that companies using AI, not just companies selling AI, are considered in scope for the proposed EU AI regulation.

Lack of Company Processes

In initial engagements with healthcare companies, BSR has observed that companies have varying degrees of maturity in respect to their AI governance processes. While some companies have begun to address ethical and

human rights issues associated with AI by establishing internal teams and advisory boards dedicated to responsible technology use, such as Merck KGaA's Digital Ethics Advisory Panel, or by setting out AI principles, such as Novartis' commitment to the ethical and responsible use of AI systems, other companies remain unaware of, or unprepared to deal with, the emergent risks and impacts arising from the development and deployment of AI technologies.

With this context, BSR has started engaging healthcare companies and the technology companies that provide AI services to them to better understand the current use cases of AI, associated human rights risks, and the processes and policies in place to address those risks. This primer summarizes our findings and observations from these engagements and makes preliminary recommendations to companies in the healthcare industry on how they can address the human rights impacts of AI in healthcare.

This report is not intended to provide a comprehensive assessment of human rights impacts across the healthcare industry. Rather, it introduces salient human rights issues associated with the increased use of AI technology in the healthcare industry. The findings outlined in this

report are intended to be a starting point; healthcare companies that would like to further explore these issues should undertake more comprehensive human rights due diligence.[1]

BSR welcomes input from healthcare companies on this topic. Please reach out to Ife Ogunleye, Lale Tekisalp, or Hannah Darnton if you would like to join the conversation.

The Use of AI in the Healthcare Industry

The global market for AI in the healthcare industry reached over US$6 billion in 2021 and is projected to grow to over US$40 billion by 2027.[2] According to the World Health Organization, AI has the potential to strengthen the delivery of healthcare and medicine and help countries achieve universal health coverage. AI can be deployed across various sectors, including health or clinical care, research and drug development, public health surveillance and monitoring, and health systems management.

AI technologies are being utilized across the healthcare industry for various use cases and applications. **In health and clinical care**, we have observed the following use cases:

- **Diagnosis:** AI technologies are being developed and deployed for diagnostic purposes and disease detection through radiology, medical imaging, and other tools. AI is being used for initial disease diagnosis, to support medical professionals in making prompt and accurate diagnosis, and to predict illness or disease before they occur.[3]

- **Patient management:** AI technologies are being used by hospitals, clinics, and healthcare professionals to manage patient records, identify and prevent clinical errors, monitor patient treatment, medication and care plans, and support treatment decisions and self-management by patients.

- **Personalized medication and care:** AI technologies are being adapted to provide personalized medical care and wellness services, including diagnosis of medical conditions, patient health management to prevent the occurrence or progression of disease, and medication dosage.

In **research and drug development**, AI systems are being used for:

- **Data generation and analysis:** AI systems can be used to generate high-quality data for biomedical research and drug development. They are also able to effectively analyze large datasets to enable improved understanding of diseases and human physiology and accelerate the discovery of effective treatments.

- **Drug discovery:** AI technologies are being used to accelerate drug discovery and support therapies development for various diseases and health conditions. AI can enable the identification of molecules that may contribute to disease progression, as well as compounds that may effectively target diseases, and the generation of new drug candidates.

- **Clinical trial design:** AI systems can be used to optimize the operational design of clinical trials, including trial design and participant recruitment and monitoring. They also may be used to determine which countries or clinical centers are most suitable for a specific clinical trial and how resources should be allocated among different ongoing clinical trials, as well as for the rapid identification and recruitment of participants.

In **public health**, AI technologies are being applied for:

- **Disease outbreak monitoring and management:** AI systems are being used to identify disease outbreaks and manage public response. This includes identifying disease transmission, facilitating detection and tracking, and developing vaccines and other treatments.

- **Health promotion and disease prevention**: AI systems can be used to identify and microtarget individuals or communities that are at high risk of certain diseases or identify and address underlying causes for poor health outcomes or disease outbreaks.

The adoption of AI technologies in healthcare is expected to continue to grow across various functions and use cases. According to a survey conducted in 2022, AI is at the top of emerging technologies expected to uptake in the healthcare industry, with 50% respondents saying they are planning to invest in AI technologies in 2022. Examples of innovations in the sector include Insilico Medicine's AI-enabled drug discovery platform, which has generated a drug candidate for idiopathic pulmonary fibrosis that has advanced to clinical trials.

Uptake of emerging technologies in the next two years
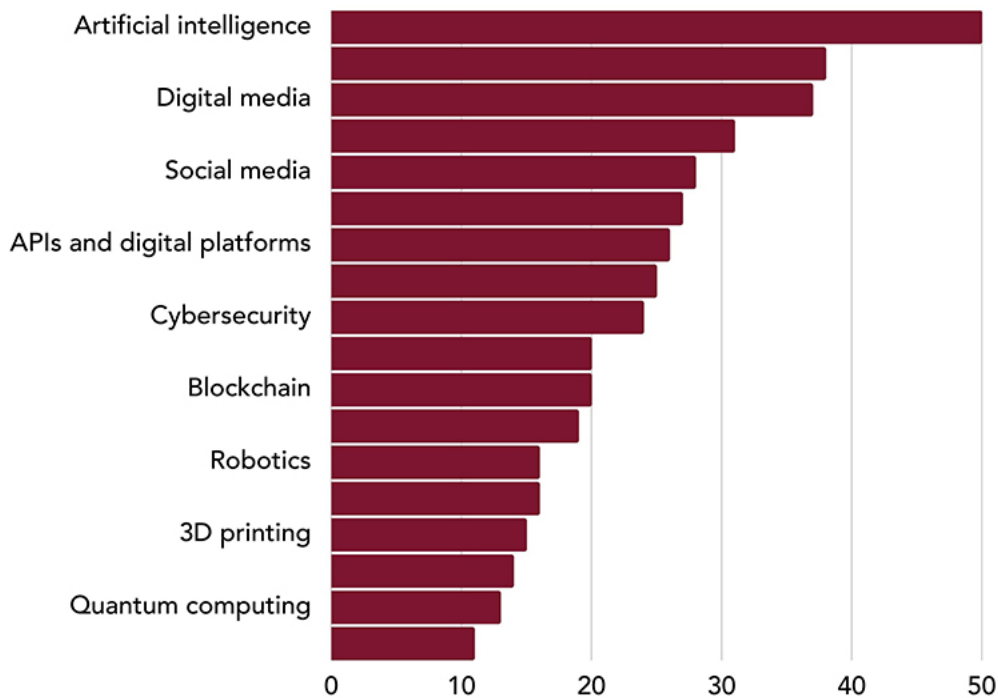
Q: In the next two years, which of the following emerging technologies do you expect your organization to invest in?

Source: Global Data

Recreated from Pharmaceutical Technology

Human Rights Impacts of AI in the Healthcare Industry

The use of AI technologies may alleviate or exacerbate existing human rights impacts in the healthcare industry. In this report, we focus mainly on the human rights risks that AI technologies may lead to, and the ways in which AI technologies may exacerbate the existing inequities in the healthcare industry.

Through our engagements with healthcare companies and the technology companies that provide AI services to them, we **identified four main categories of risk:**

1. Non-Discrimination

## 2. Right to Health and Science

3. Privacy and Surveillance

4. Human Autonomy

Below we list the salient human rights associated with these categories. However, it is important to note that all human rights are indivisible, interdependent, and interrelated. The improvement of one right facilitates advancement of the others; the deprivation of one right adversely affects others.
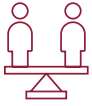
1. Non-Discrimination

AI systems may reinforce systemic bias and exacerbate existing inequities in the healthcare system. The use of AI solutions across the healthcare industry may result in the discrimination of individuals by race, gender, age, disability, or other protected categories. AI technologies are trained on historical data and so may be biased if the data they are trained on is biased, discriminatory, or unrepresentative.

Datasets may be encoded with existing biases in the healthcare system such as overrepresentation of certain gender, racial, or age groups and underrepresentation of others, which may lead to disparate performance or accuracy rates for different patient demographics. For instance,

AI models may perform less accurately at diagnosing medical conditions in individuals or communities of color or recommend courses of treatment that may be less effective for those patients.

AI systems may also perpetuate discrimination by exacerbating the "digital divide."[4] Where access to AI technologies is inequitable due to barriers of language, geography, or access to devices, benefits associated with uptake such as improved health outcomes may be unequally distributed.

HUMAN RIGHTS THAT MAY BE IMPACTED

- Right to equality and non-discrimination (UDHR Article 1 and 2)

2. Right to Health and Science

Well-designed and appropriately implemented AI solutions in the healthcare industry could lead to more effective and easily accessible healthcare services. However, there are risks that AI solutions could lead to unintended negative health outcomes or further disparities in access to healthcare. For instance, the overrepresentation of certain populations in different phases of the drug discovery process, including clinical trials, may lead to disproportionately worse health outcomes for underrepresented populations.[5]

Where AI solutions such as AI diagnostic tools are only made available to high income individuals or communities, they may increase inequities by contributing to the healthcare divide. Diagnostic errors may also negatively impact the right to health where users or patients erroneously believe the outcomes of AI algorithms to be foolproof, resulting in healthcare professionals and patients making decisions about health and treatment options based on erroneous outcomes suggested by AI algorithms.

The use of AI systems to prioritize research and development may also have implications on the right to science. For example, if AI systems are used to prioritize research and innovation focused on the needs of high-income countries, lower-income countries may not be able to benefit from these scientific advancements because innovative treatments may be too costly or may require technology (such as refrigeration) that some markets do not have access to.

HUMAN RIGHTS THAT MAY BE IMPACTED

- Right to health (UDHR Article 25, ICESCR Article 12)
- Right to science (UDHR Article 27, ICESCR Article 15)

3. Privacy and Surveillance

With the use of AI solutions, healthcare organizations may collect, utilize, and share patient data in ways that infringe on the right to privacy. Data is routinely bought and used, frequently without the knowledge of the patient/owner of the data. In addition, what is considered health data has expanded in recent years to include personal data from a variety of sources beyond standard health sources, such as personal devices and environmental, behavioral, and socioeconomic sources.

Although the collection of data from a wide variety of sources may improve data quality and, ultimately, the performance and accuracy of AI technologies, extensive data collection increases the risk of violating individuals' privacy. For instance, private information about individuals' health status may be inferred or disclosed without their informed consent, leading to stigmatization or exclusion of vulnerable groups.

Increased data collection may also lead to increased surveillance by government or private actors such as law enforcement agencies, insurance companies, employers, etc. External actors may use sensitive health data to identify, track, or monitor individuals. For example, location data tracked by quarantine enforcement apps can be used as part of law enforcement and intelligence efforts. Similarly, governments may demand user data related to pregnancy and abortion.[6]

HUMAN RIGHTS THAT MAY BE IMPACTED

- Right to privacy (UDHR Article 12)
- Note that the violation of privacy may have secondary impacts on other rights, including the right to life, liberty, and security, and freedom from arbitrary arrest.

4. Human Autonomy

The increasing use of AI systems in the healthcare industry to provide healthcare services such as diagnosis, treatment plans, and symptom monitoring may lead to an overreliance on these systems by healthcare professionals and patients. Over-reliance on algorithms to make decisions regarding health management could result in potential impacts to mental autonomy.[7]

For example, healthcare professionals and patients may believe AI systems to be infallible. Unequivocal confidence in AI technology could lead to flawed healthcare decisions pertaining to diagnosis or treatment. These decisions may have harmful outcomes, such as misdiagnosis, overdiagnosis, underdiagnosis, or overtreatment.[8]

HUMAN RIGHTS THAT MAY BE IMPACTED

- Human autonomy and dignity (UDHR Article 1)
- Right to freedom of thought (UDHR Article 18)

Recommendations

Responsible AI challenges typically need the involvement of various functions at a company. For companies that do not yet have a dedicated team addressing these issues, we recommend starting the process by involving the following functions:

A) Teams that can manage the issue from a central perspective, such as Sustainability, Human Rights, Ethics, Legal Compliance

B) Teams that use AI technologies, such as Research and Development, Marketing, Human Resources

C) Teams that develop or purchase AI technologies, such as Procurement or Marketing

To mitigate any adverse human rights impacts, companies can take actions including but not limited to the ones listed below:

1. **Take inventory of the AI use cases within the company.**
   An important first step is to understand how AI is being used by different functions across the business. Companies should reach out to the teams listed above and ask them how they are using or are planning to use AI technologies in their work. Companies should then make a list of these use cases and prioritize those that may be higher risk.

2. **Undertake human rights due diligence**
   To identify and address the actual and potential human rights impacts of the AI solutions they are using, companies should start by undertaking human rights due diligence[9], a process that specifically assesses risks to people (as opposed to other risks a company may face). Human rights due diligence should be undertaken on an ongoing basis because the ways in which AI technologies are used may change over time. In addition to practicing continuous due diligence, companies should undertake specific human rights impact assessment when developing, using, or procuring new AI technologies that are likely to pose risks to human rights. The results of these impact assessments should then be used, if necessary, to modify or adapt the technologies, or to ensure sufficient mitigation measures or safeguards are in place to address any risks identified.

3. **State purpose and use limitations**
   Companies should have a clearly defined purpose for the use of AI and consider setting use limitations within implementation guidelines. If the AI solution is going to be shared externally with other users, companies should establish acceptable use policies that define what users can and cannot do with the AI solution.

4. **Establish a governance mechanism for the responsible use of AI**
   There are important questions around how ethical and human rights implications are identified, assessed, and addressed by the company. Some companies have added new expertise to existing ethics panels and/or developed guiding principles on their use of AI, whereas others have created bespoke councils to advise specifically on AI.

5. **Ensure a high level of data protection**
    Many of the human rights risks related to AI stem from the use of personal data. While it can be tempting to focus on compliance with relevant privacy and data protection frameworks, many of these put the focus on the rightsholder to assert their right to privacy, rather than requiring the integration of privacy and data protection by design. Companies should go beyond regulatory compliance and align their internal data protection and privacy commitments, policies, and practices with the highest international standards.

6. **Test AI models for bias and externalities.**
    AI models rely on data input, which can be biased and lead to potential adverse human rights impacts around discrimination and the unfair distribution of goods and services. Companies should continually review data inputs that are used by the AI models, through data audits and assessments.

7. **Undertake adversarial testing**
    AI solutions may lead to different impacts when used in different contexts or for different use cases. Companies should undertake adversarial testing to new risks as they arise, especially if the use of AI solutions expands to new functional areas or geographies. Adversarial testing refers to exercises where the AI system is stress tested to discover the ways in which the system might be misused or lead to harmful outcomes. Methodologies might include futures thinking or red team/blue team testing (traditionally used in the cybersecurity field).[10]

8. **Provide transparency about how the AI models work**
    Developers of AI models should communicate the details of the model to its users, including training data sources, metrics that the model[11] optimizes for, and key limitations of the model. Companies that are using AI solutions should also consider how AI models can be explained to end users or employees who engage with these models.

9. **Integrate feedback**
    Establish a reporting channel where potential misuse and abuse of the AI solutions can be reported to the teams or third parties who have developed the solution. Workers' voices should be central when making decisions on how to deploy a new technology. Ensure that the necessary mechanisms are in place to integrate employee feedback into the way AI solutions are used by the company.

10. **Prepare for upcoming regulation**

    Ensure that your company is prepared for upcoming regulation (e.g., EU Corporate Sustainability Due Diligence Directive (CSDDD), proposed EU AI Act). As a first step, companies can either 1) ensure that AI is included in company-wide human rights due diligence processes and/or 2) conduct due diligence on specific AI use cases to identify human rights risks.

11. **Engage in dialogue with other industry players.**
As the use of AI technologies becomes more prevalent in the healthcare sector, companies are becoming more interested in its impact. Through dialogue with other industry players, companies can help advance the understanding of the human rights impacts of AI in their sector.

Our understanding of the human rights impacts of AI will evolve as the technology becomes more pervasive across the healthcare industry. Companies should start putting structures and processes in place to address the adverse impacts of the technologies they are using. However, these systems should be agile to meet future developments and concerns.

Endnotes

1. The <u>UN Guiding Principles on Business and Human Rights (UNGPs)</u> provide a framework for human rights due diligence (HRDD). The <u>UN B-Tech Project</u> provides further guidance on how HRDD can be applied to technology products and services.

2. <u>Global AI in Healthcare Market Report 2022: Rising Utilization of Robots for Surgical and Rehabilitation Procedures Driving Growth</u>," - ResearchAndMarkets.com. Oct. 2022.

3. <u>Ethics and Governance of Artificial Intelligence for Health; WHO Guidance</u>, June 2021.

4. The '"digital divide'" refers to the uneven distribution, of access to, or use of technologies among distinct groups.

5. "<u>How Will Healthcare Regulators Address Artificial Intelligence?</u>" The Regulatory Review, Oct. 2021.

6. Privacy concerns related to abortion data have increased. See recent articles from <u>Politico</u>, <u>PBS</u>, and the <u>New York Times.</u>

7. For further reading on the impact of technology on mental autonomy, see "Losing the Freedom to Be Human," Columbia Human Rights Law Review, Dec. 2020.

8. See <u>"Trust and medical AI: the challenges we face and the expertise needed to overcome them</u>," Journal of the American Medical Informatics Association, April 2021.

9. The <u>UN Guiding Principles on Business and Human Rights (UNGPs)</u> provide a framework for human rights due diligence (HRDD). The <u>UN B-Tech Project</u> provides further guidance on how HRDD can be applied to technology products and services.

10. See <u>Microsoft's Harms Modeling Tool</u> and <u>Omidyar's Ethical Explorer Pack</u> as examples.

11. The 2019 academic paper <u>Model Cards for Model Reporting</u> proposes the use of "model cards" to provide information about an AI model's performance and limitations. Practical examples include <u>Google's use of Model Cards</u>, and <u>Microsoft's Datasheets for Datasets</u> tool to document the datasets used for training and evaluating machine learning models.

BSR®

About BSR

BSR is a sustainable business network and consultancy focused on creating a world in which all people can thrive on a healthy planet. With offices in Asia, Europe, and North America, BSR provides its 300+ member companies with insight, advice, and collaborative initiatives to help them see a changing world more clearly, create long-term value, and scale impact.

Disclaimer

The conclusions presented in this document represent BSR's best professional judgment, based upon the information available and conditions existing as of the date of the review. In performing its assignment, BSR relies upon publicly available information, information provided by member company, and information provided by third parties. Accordingly, the conclusions in this document are valid only to the extent that the information provided or available to BSR was accurate and complete, and the strength and accuracy of the conclusions may be impacted by facts, data, and context to which BSR was not privy. As such, the facts or conclusions referenced in this document should not be considered an audit, certification, or any form of qualification. This document does not constitute and cannot be relied upon as legal advice of any sort and cannot be considered an exhaustive review of legal or regulatory compliance. BSR makes no representations or warranties, express or implied, about the business or its operations. BSR maintains a policy of not acting as a representative of its membership, nor does it endorse specific policies or standards.

The views expressed in this document do not reflect those of BSR member companies.

www.bsr.org