

A woman with curly hair, wearing a brown blazer, is smiling and looking at a document in a meeting. She is surrounded by other people, including a man with glasses on the left and a woman on the right. The document she is holding is titled "CONTRACT OF SALE OF REAL ESTATE - Disclosure of Cash".

# AI and Human Rights in **Financial Services**

July, 2023



Introduction

---

01

The Use of AI in  
Financial Services

---

05

Human Rights Impacts of AI in  
the Financial Services Sector

---

07

Recommendations

---

12



This is an industry report on how artificial intelligence (AI) technologies are driving change within financial services and the challenges and opportunities that trend brings from the perspective of human rights.

As digital transformation within the financial services sector continues, there are some key reasons why companies need to consider the impacts of AI:



### Material Risks

There are a range of material risks related to the use of AI by the financial services sector, including financial risks and risks to reputation (connected to potential human rights impacts). The regulatory environment is also evolving, with the [EU's proposed Corporate Sustainability Due Diligence Directive](#) and [Artificial Intelligence Act](#) creating new legal and compliance requirements. These new regulations 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 EU's Artificial Intelligence Act.



### Human Rights

Technological transformation brings complex, nuanced, and systemwide risks and opportunities for the realization of human rights. These risks and opportunities may be associated with the design and development of technologies and, just as importantly, with how the technologies are

deployed and used by companies, such as those in the financial services sector.



### Lack of Company Processes

In initial engagements with companies in the financial services sector, BSR has observed that many do not have internal processes to review the human rights and other impacts of AI applications. This makes it less likely that the material risks listed above will be picked up and increases the potential liability of companies. There are, however, some emerging good practices in the sector, such as [Scotiabank's Data Ethics Principles](#), [RBC Capital Markets' Responsible and Explainable AI](#), [TD's Responsible AI report](#), and [JP Morgan's AI Research Agenda](#).

With this context, BSR has started engaging companies in the financial services sector 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 report summarizes our findings and observations from these engagements, and makes preliminary recommendations to companies in the financial services sector on how they can address the human rights impacts of AI in financial services.

This report is not intended to provide a comprehensive assessment of all human rights impacts across the financial services sector. Rather, it introduces some of the most significant impacts associated with the increased use of AI technology.

The findings outlined in this report are intended to be a starting point; financial services companies that would like to further explore these issues should undertake more comprehensive human rights due diligence.<sup>1</sup>

Finally, this report has been written for the financial services sector broadly. BSR recognizes that this sector comprises a broad range of different types of company with different structures, at different stages of integrating AI technologies into their operations, and using AI for different purposes. As such, the risks related to AI will vary between different types of company, and not all of the examples in this report will be relevant for all companies in the sector. Many of the risks highlighted are of particular relevance to consumer-facing parts of the sector, where customers are potentially impacted. But there are also risks for other parts of the sector, particularly in relation to their internal use of AI technologies.

BSR welcomes input from financial services companies on this topic. Please reach out to [Richard Wingfield](#), [Lale Tekisalp](#), or [Hannah Darnton](#) if you would like to join the conversation.

# The Use of AI in the Financial Services Sector

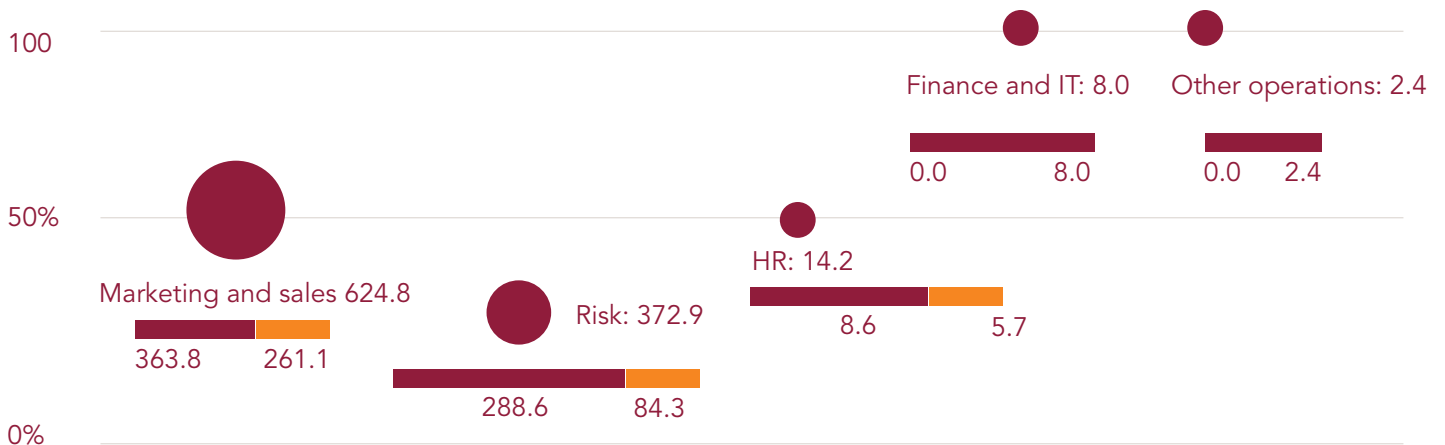
Companies in the financial services sector are increasingly integrating and using AI technologies to reduce operational costs, drive efficiency, support investment analysis, and deliver a better service to customers and clients. A [survey](#) of over 100 firms by the UK’s Financial Conduct Authority in 2019 showed that two-thirds were using machine learning technologies in some form, a proportion that [increased](#) to almost three-quarters in 2022. The most recent survey also showed that firms expected the overall median number of machine learning applications to increase by 3.5 times over the next three years. This is unsurprising given that there are strong financial incentives for investment in AI, with estimates that its potential value for global banking could be as high as US\$1 trillion.<sup>2</sup>

## Total potential annual value (in USD billions)<sup>2</sup>

1,022.4 (15.4% of sales )



## % of value driven by advanced AI, by function



We have observed that companies in the financial services sector are employing AI in a range of ways:

In the front office and for consumer-facing parts of the sector:

- Chatbots powered by natural language processing are increasingly being used to provide financial guidance to customers and clients 24/7.
- AI-based technologies, such as facial recognition and voice commands, are used to **access and use financial services apps**, improving the customer experience (a trend likely to increase as, post-Covid, more people look to access financial services digitally).
- AI technologies are used to provide **more personalized insights and advice** to customers and clients.

In the middle office and back office, and where decisions are made, algorithms are being developed and deployed, including to:

- Assess an individual's credit risk;
- Detect instances of potential fraud, money laundering, and "bad actors," including as part of "Know Your Customer" due diligence (a trend likely to increase as a result of greater pressure from regulators to use new technologies when complying with money laundering regulations); and
- Inform investment decisions and assess their performance.

New AI technologies are increasingly being used because, in part, they have the potential to counter traditional human bias and increase the speed and accuracy of providing financial services to people who need them, enlarging the potential customer and client base and ensuring a competitive edge. At the same time, they may also help with regulatory compliance, particularly money laundering regulations, a significant overhead for companies operating in the financial services sector. Regulators are actively encouraging the use of AI technologies, potentially leading to reductions in operational costs and to more sophisticated analysis by combining integrated human intelligence with the data analytics and analysis capabilities of AI to reduce rates of false positive alerts and provide more monitoring.

# Human Rights Impacts of AI in the Financial Services Sector

The use of AI technologies may alleviate or exacerbate existing human rights impacts of companies in the financial services sector. In this report, we focus mainly on the human rights risks that AI technologies may be associated with, and the ways in which AI technologies may exacerbate the adverse human rights impacts that are already present in the financial services sector. However, we also note where AI technologies may help advance respect for human rights.

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



## 1. Privacy



## 2. Non-Discrimination



## 3. Access to Financial Services

Below we list the most significant 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. Privacy

There are a range of risks related to privacy when AI is used for specific tasks within the financial services sector, all stemming from the collection and processing of personal data. Some of the most basic risks relate to the risk of data breaches or the collection of data beyond what is strictly needed, which increases risks relating to data breaches. These can both also amount to breaches of data protection regulations. This risk will be relevant to any company that collects personal data, including financial data, such as retail and corporate banking institutions, credit card companies, digital payment platforms, and payroll companies.

Increasingly, some financial services institutions are using “alternative data” to determine credit risk, identify money laundering, or monitor for fraud. Alternative data can include nontraditional personal data such as a user’s location or social media activity. These may result in arbitrary interferences with the right to privacy if they are

unnecessary or not proportionate to the need. There will also be risks when new and/or alternative types of data are collected, stored, used, shared, or combined to make financial assessment decisions using AI technologies without appropriate informed consent, or where data which has been provided with consent is used to infer new types of information about users. This risk will be relevant to any company that assesses an individual’s credit risk or that monitors potential money laundering or fraud, such as retail and corporate banks, credit card companies, and other lenders.

While the risks to the right to privacy are significant, there is the potential for AI-based cybersecurity technologies to help protect individuals’ privacy by making the company’s systems for identifying threats to personal data more effective and efficient.



## HUMAN RIGHTS THAT MAY BE IMPACTED

- Right to privacy (UDHR Article 12; ICCPR Article 17)





## 2. Non-Discrimination

AI technologies are invariably trained on historical data and so may themselves be biased if the training data is biased, discriminatory, or nonrepresentative. Unless these biases are addressed, the assessments or predictions made by those AI technologies may themselves also be biased and lead to discriminatory outcomes.

There have been well publicized instances of these kinds of effects, such as online job listings with roles commanding higher salaries being advertised only to men based on historical datasets showing men's salaries were higher than women's. Where algorithms and automation are involved, the risks are increased. In the financial services sector, this could lead

to discriminatory outcomes where, for example, decisions are made as to whether a person can access particular financial services, the price that they will pay, or their credit risk. The risks could even relate to individuals belonging to certain groups having to wait longer for decisions to be made, or being required to overcome additional hurdles in order for applications to be successful.

While such discriminatory outcomes would likely be based on characteristics on which there are already patterns of discrimination in society (such as gender or race), the use of more complex models and nontraditional data could create new forms of discriminatory treatment based on other characteristics,



### HUMAN RIGHTS THAT MAY BE IMPACTED

- Right to non-discrimination and equal treatment (UDHR Article 2; ICCPR Articles 2(1) and 26)

such as example where a person makes a spelling mistake on their application form. There is also the potential for discriminatory outcomes where data may be used as a proxy for other factors, such as [postal codes acting as proxies for race or socio-economic status](#).

In the financial services sector, connecting financial data to other data (e.g., data provided by data brokers or obtained from a person's social media activity) may further increase the risk for discriminatory outcomes. As with adverse impacts on privacy, adverse impacts on non-discrimination may also constitute breaches of anti-discrimination regulations. These risks will be relevant to any company that uses AI to make decisions or determinations about people. This would include all companies that use AI

for internal decision-making (such as hiring decisions), as well as for decision-making about customers such as mortgage financiers, retail banks, credit card companies, and other lenders.

However, given that biases and discriminatory outcomes occur when humans make decisions in the financial services sector, new AI technologies have the potential to make more objective outcomes, which could reduce instances of discrimination through the use of high-quality datasets or by using those technologies to identify differential treatment. In addition, where differential treatment does occur as a result of an algorithm, it may be easier to modify the algorithm than to address biases stemming from human decision-making.



### 3. Access to Financial Services

Connected to the above, the use of AI technologies could lead to individuals being unjustifiably excluded from access to financial services. The use of AI technologies to profile individuals' credit risk or other ways of profiling individuals exacerbates this risk. AI, particularly when trained on historical data that may be biased, discriminatory, or nonrepresentative, may be associated with discriminatory outcomes and results (see the examples listed above).

Given existing differences in areas such as credit scores, (where white people tend to have higher scores than ethnic minorities), AI tools developed using existing datasets

could reinforce biased assumptions about who presents a credit risk. This could lead to individuals unable to access (or afford) certain financial services such as bank accounts, insurance, or loans due to an incorrect assessment of their risk. In a context where many lenders' appetites for financial risk are low, and large numbers of people struggling to gain access to financial services, the impacts of further exclusions would be particularly pronounced. The risk would be exacerbated where there is a lack of effective human oversight of those technologies. This risk would be relevant to any company



that provides access to financial services, including retail and corporate banking institutions, credit card companies, mortgage financiers, and other types of lenders.

On the other hand, AI technologies could also enhance access to financial services for some individuals who are currently excluded from them. For example, AI technologies could reduce operational costs that could then be passed on to consumers through lower prices, making financial services more affordable to more people.

It is also possible that AI technologies, if used appropriately, could determine individuals' risk profiles on a more granular level, independent

of any potentially discriminatory criteria, leading to lower risk profiles for some individuals than would have been the case using traditional risk-profiling assessments. For example, a 2019 study suggested the use of algorithms by FinTech lenders in the US reduced the level of discrimination, particularly discrimination based on race, by 40 percent. This could lead to greater access to certain financial services (or lower prices) for many people. It may even be possible for machine learning to establish risk profiles for individuals where it would not otherwise have been possible due to a lack of certain types of traditional data, further increasing the range of people who are eligible.

# Recommendations

Responsible AI challenges typically need the involvement of various functions at a company. For companies that do not have a dedicated team addressing these issues yet, 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 Operations, Supply Chain, Retail Lending/Credit, Customer Service, Human Resources
- C)** Teams that develop or purchase AI technologies, such as Technology, Product, IT, Research and Development, Procurement

To mitigate any adverse human rights impacts related to the use of AI, 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 company. 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 technologies that they are developing, using, or procuring, companies should start by undertaking human rights due diligence<sup>3</sup>, 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 the ethical and human rights implications are understood, anticipated, and addressed by the company. Some companies have developed human rights working

groups, external advisory committees, and other processes for identifying and mitigating risks both broadly and specific to AI. Companies should consider establishing such processes, or building the capacity of equivalent internal teams and processes (or those who are responsible for ESG more broadly), to advise specifically on the use of AI internally. This could involve establishing dedicated roles related to responsible AI. Companies should also publicly disclose information about these mechanisms and the principles that underpin them.

#### **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 (exercises where the AI system is stress tested to discover the ways in which the system might be misused or lead to harmful outcomes) to new risks as they arise, especially if the use of AI solutions expands to new functional areas or geographies.<sup>4</sup>

#### **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 optimizes for, and key limitations of the model. Companies should also consider how AI models can be explained to end-users. For example, end-users who are given a particular credit risk assessment should be provided with an explanation on the reasons for that assessment.

#### **9. Integrate feedback**

Establish a reporting channel where potential misuse and abuse of the AI technologies 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., the [EU's proposed Corporate Sustainability Due Diligence Directive](#) and Artificial Intelligence Act).

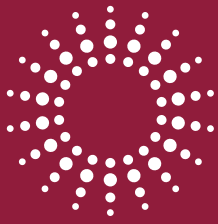
**11. Engage in dialogue with other industry players.**

As the use of AI technologies becomes more prevalent in the financial services sector, companies are becoming more interested in its impacts. Through dialogue with other industry players, companies can help advance the understanding on 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 financial services sector. Companies should start putting in place structures and processes to address the adverse impacts of the technologies they are using, but these systems should be agile to meet future concerns.

# Endnotes

- 
1. The [UN Guiding Principles on Business and Human Rights \(UNGPs\)](#) provide a framework for human rights due diligence (HRDD). The [UN B-Tech Project](#) provides further guidance on how HRDD can be applied to technology products and services.
  2. McKinsey, [AI-bank of the future: Can banks meet the AI challenge?](#) 19 September 2020.
  3. [The UN Guiding Principles on Business and Human Rights \(UNGPs\)](#) provide a framework for human rights due diligence (HRDD). [The UN B-Tech Project](#) provides further guidance on how HRDD can be applied to technology products and services.
  4. See Microsoft's Harms Modeling Tool and Omidyar's Ethical Explorer Pack as examples.



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](http://www.bsr.org)

Copyright © 2023 by Business for Social Responsibility (BSR)

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law.