# AI Human Rights Impact Assessment

# for Educators

**Purpose:** To provide a structured guide for educators wishing to assess an AI application for use in educational settings.

**Note on AI Literacy:**

If you are new to AI or would like to build foundational knowledge, you may wish to explore one or more of the following resources:

- **Academic Applications of Artificial Intelligence (AAAI)** – A free, self-paced course with five practical modules. Optional micro-credential available.
  - ➲ [Enroll in the AAAI Course](#)

- **Teaching AI Ethics – Leon Furze** – A blog series offering educator-focused insights into AI ethics and classroom use.

  - ➲ [Read Teaching AI Ethics: The Series - Leon Furze](#)

  - ➲ [Listen to audio versions of Leon Furze's blog posts](#)

- **AI Dialogues Podcast** – Conversations on AI through an equity, diversity, and inclusion lens.

  - ➲ [Listen to AI Dialogues](#)

These resources are intended to support diverse learning needs and help ensure that all users, regardless of prior experience, can confidently engage with the AIHRIAE.

# Quick Start Guide

## Get warmed up by considering this Quick Start guide:

Review the full guide (starting on page 5) for more in-depth questions, pedagogical background, supporting information and examples.

| **AI Tool Assessment**<br>Select questions from the guide. Numbering reflects their location within the full guide. | **Example: AI pizza ordering and delivery application**  |
|---|---|
| **Q1: What is the primary function of the AI application?** | **What is the primary function of the AI application?**<br>• To streamline the pizza ordering and delivery process. |
| **Q2: Why is this AI application needed or beneficial?** (Check all that apply)<br>☐ Provides students with personalized learning<br>☐ Improves student accessibility<br>☐ Provides students with 24/7 support<br>☐ Frees up resources for use elsewhere<br>☐ Supports diverse learning modalities<br>☐ Other | **Why is this AI application needed or beneficial?**<br>• Provides customers<br>  o with personalized pizza recommendations<br>  o improved customer accessibility<br>  o 24/7 support<br>• Frees up restaurant resources by automating the ordering process<br>• Enhances customer experience by reducing wait times and errors in orders. |
| **Q7: What potential harm could arise for these learners?** (Check all that apply)<br>☐ Providing students with inaccurate information<br>☐ Providing students with biased information.<br>☐ Sharing students' private Information.<br>☐ Failing to accommodate students' diverse needs.<br>☐ Excluding students<br>☐ Sharing students' intellectual property<br>☐ Unduly persuading or influencing students<br>☐ Other | **What potential harm could arise for customers?**<br>• Providing customers with inaccurate order details or delivery times.<br>• Recommending certain pizzas over others due to biased algorithms.<br>• Sharing customers' private information:<br>• Not considering dietary restrictions or preferences.<br>• Limited accessibility for those without internet access<br>• Unauthorized use of customer reviews or feedback. |

| AI Tool Assessment | Example: AI pizza ordering and delivery application |
|---|---|
| Select questions from the guide. Numbering reflects their location within the full guide. | |
| **Q8: How can potential harm be mitigated?** (Check all that apply) <br> ☐ Choose accessible technologies. <br> ☐ Ensure transparency: Verify that decision-making processes are explainable and that explanations are available to students. <br> ☐ Identify and address bias through regular audits. <br> ☐ Protect privacy via strict data privacy policies and data anonymization. <br> ☐ Maintain human oversight. <br> ☐ Solicit feedback from students and educators and address identified issues promptly. <br> ☐ Other | **How can potential harm be mitigated?** <br> • Regularly review and update algorithms to ensure unbiased recommendations <br> • Ensure customer information is protected. <br> • Ensure staff are available to handle complex orders. <br> • Regularly gather and act on customer feedback. <br> • Ensure the application is user-friendly and accessible to all customers |
| **Q15: How will feedback be solicited?** (Check all that apply) <br> ☐ Survey <br> ☐ Focus groups <br> ☐ In-class discussion <br> ☐ Other | **How will feedback be solicited?** <br> • Online survey sent to customers after their order. <br> • Regular focus groups with frequent customers. <br> • Social media polls and direct feedback through the app. |
| **Q16: Who will review the feedback and how often?** | **Who will review the feedback and how often?** <br> A dedicated team of customer service specialists will review the feedback on a monthly basis to identify trends and areas for improvement. |
| **Q17: How will concerns be addressed?** | **How will concerns be addressed?** <br> Concerns will be addressed by prioritizing issues based on severity and frequency, implementing necessary changes, and communicating updates to customers. |
| **Q9: What alternatives exist?** (Check all that apply) <br> ☐ Variations on this AI application <br> ☐ Other AI applications <br> ☐ Non-AI applications <br> ☐ Other | **What alternatives exist?** <br> • Different AI models or algorithms for order processing. <br> • Non-AI applications: Traditional online ordering systems or phone-based orders. <br> • Other: Manual order processing and delivery. |

# AI Human Rights Impact Assessment for Educators:
# In-depth Guide

In November of 2024, the "Human Rights Impact Assessment for AI" (HRIAAI) was released jointly by the Law Commission of Ontario and the Ontario Human Rights Commission[1]. My first thought upon reviewing this structured guide for evaluating AI systems was that this was a needed and valuable tool for organizations to ensure that AI system selection and integration were carried out with a human rights focus. However, I also concluded that it would be difficult to apply to educational settings. So, over the following weeks, I set out to develop the "AI Human Rights Impact Assessment for Educators" (AIHRIAE), a tool tailored to the needs of educators and informed by Ontario[2], Canadian[3] and international[4] legislation that aims to align with the UNESCO goal of "harnessing the benefits of artificial intelligence… [while] ensuring the use of digital technologies to promote, protect and fulfill human rights."[5]

Since ChatGPT's emergence in November 2022, I have observed varying responses among educators: some see the potential benefits of AI technologies and are eager to explore them in their classrooms but may need guidance in identifying and mitigating potential risks; others see the potential risks and find the thought of managing those risks to be overwhelming; many fall somewhere in between.  The AIHRIAE was designed to meet the needs of all educators, no matter where they fit in this continuum, by providing a structured guide for identifying both benefits and risks, developing risk mitigation strategies, and creating a plan for optimizing learning moving forward.

The guide is divided into two parts:

**Part A** is a 4-step guide for educators considering using AI in their classrooms.

1. Define the AI Tool's Purpose
2. Consider the Intended User
3. Consider the Educational Context
4. Seek Feedback, Review, and Revise
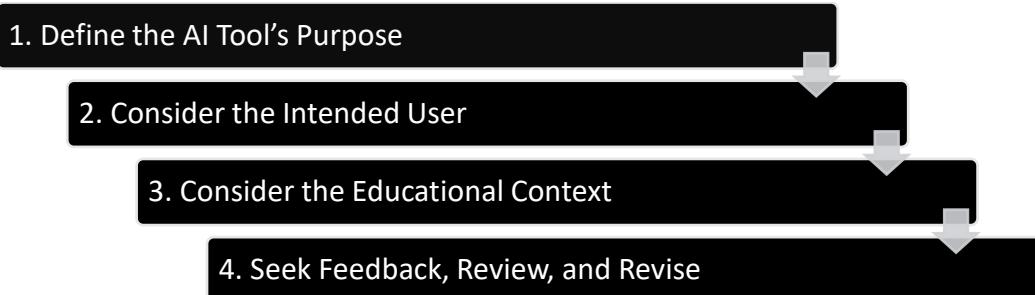
**Part B** provides risk mitigation strategies that are beneficial for all educators to consider but essential for those considering using AI in a high-risk context (as identified in Part A).  Both parts prioritize sound pedagogy, are backed by peer-reviewed research, and have been developed through interdisciplinary consultation.

**Understanding the Relationship Between Bias, Impact, and Discrimination**

Whether embedded in training data, arising from the design of the AI system, or resulting from the context in which it is deployed, bias in AI systems can lead to adverse impacts that disproportionately affect individuals or groups protected under the Ontario Human Rights Code. These impacts may not be obvious or intentional, but they can still constitute discrimination if they meet the Code's three-part test:

1. The person has a characteristic protected under the Code (such as disability, race, sex, or age);
2. They experience adverse treatment or impact in a protected social area protected by the Code (such as education); and
3. The protected characteristic is a factor in the adverse impact.

Human rights law requires a contextual analysis that considers the full effect of the distinction on the affected individual or group. Since education is a protected area under the Ontario Human Rights Code (it falls under "services"), educators using AI tools must therefore assess not only the tool's intended function but also its real-world effects. By proactively identifying and mitigating bias and ensuring that AI tools do not contribute to systemic barriers, educators can uphold their duty to provide equitable and inclusive learning environments in accordance with the Code.

**Appendix A** provides examples to help you consider in more detail how bias, impact, and discrimination may manifest in educational AI use, and how the AIHRIAE can support educators in identifying, assessing, and mitigating these risks in alignment with the Ontario Human Rights Code.

**To get the most out of the AIHRIAE consider the following resources:**

- **Alignment with International Human Rights Frameworks: Appendix B:** For those outside of Canada, a detailed alignment table is provided which maps the AIHRIAE's core principles to international human rights and AI governance frameworks

- **Sample Use Cases: Appendix C** provides a set of ten sample use cases that illustrate how the AIHRIAE can be applied in real-world educational contexts. These scenarios represent a variety of perspectives: instructors, librarians, instructional designers, and accessibility advisors.  Reviewing these scenarios before beginning the In-Depth Guide can help you ground its application in a meaningful example that can help you better understand how the AIHRIAE can be applied in your context.

- **Glossary of Key Terms: A Glossary** is provided at the end of the document to support your understanding of the concepts used throughout this guide.

## PART A: Assess Purpose and Impact

### STEP 1: Define the AI Tool's Purpose

**Clearly articulate what the AI tool is intended to do and why it is needed or beneficial.**

In his book, *Teaching in a Digital Age: Guidelines for Designing Teaching and Learning,* Tony Bates[6] discusses how it can be tempting to embrace new learning technologies simply because they are new and exciting. However, he cautions educators that if these technologies are not used in a pedagogically sound manner and don't support achievement of preset learning outcomes, their use can hinder rather than enhance learning. The questions below can help you reflect on the purpose and potential benefits of an AI application and consider how it will help achieve learning outcomes.

| Q1 | Q2 |
|---|---|
| **What is the primary function of the AI application?** | **Why is this AI application needed or beneficial?** (See examples below) |
|  |  |

**Examples of potential benefits:**
1. **Personalized learning:** AI tools can adapt to the needs and pace of individual students and help mitigate the disparate academic impacts of the Covid-19 pandemic. [*]
2. **Accessibility:** AI can make education more accessible to students with disabilities through tools like speech-to-text, text-to-speech, and other assistive technologies. [*,***]
3. **24/7 support:** AI tutors can offer flexible, around-the-clock assistance, making it easier for students who are balancing academic responsibilities with work shifts, caregiving duties, or family obligations to access support when it fits their schedule. [*]
4. **Resource allocation:** AI tools can streamline routine tasks, enabling educators to dedicate more time to personalized instruction and other critical responsibilities, allowing for more efficient allocation of resources. [**]
5. **Modality:** By increasing learning modalities available to students (text, audio, visual, interactive), AI can support Universal Design for Learning (UDL) principles and create a more inclusive learning environment, promoting student engagement and success. [*,**,***]

[*]Relates to Section 1 of the OHRC: Ensures equal access to educational services (education is a service), including flexible learning support.
[**]Related to Section 11 of the OHRC: Ensures systemic barriers are not reinforced by resource limitations.
[***]Relates to Section 17 of the OHRC: Duty to accommodate disability to the point of undue hardship.

| Q3 |
| --- |
| **What preset learning outcomes will the AI application support or achieve?** |
| |

As the benefits listed above imply, AI has great potential to meet the needs of learners in ways that may not otherwise be feasible. When used thoughtfully and intentionally, personalized feedback and support can prompt students to engage in higher order thinking while allowing them to express themselves in ways most suited to their learning needs and preferences. Tailored learning can guide students to deeper levels of understanding when they are ready to do so.  Learning enhancement should be the primary motivation for incorporating AI applications into classrooms and courses.

| Q4 |
| --- |
| **How will the AI application enhance learning?** |
| |

## STEP 2: Consider the Intended User

**Determine who will benefit from the AI tool and who might be harmed if it fails or makes errors.**

In his SECTIONS model for selecting and using technology in education, Tony Bates[7] emphasizes the importance of ensuring that the technology meets the unique needs of the intended group of learners, including needs related to accessibility and privacy. Learners' needs vary depending on the learners' level of education, program of study, and past educational experiences, as well as their geographical location and socioeconomic status. Even within a specific cohort, learners will differ in their prior knowledge, language skills, and preferred modes of learning. Learners' unique needs can also be influenced by the mode of education (online, face-to-face, blended), their background, and their abilities. Diverse learners have diverse needs that are best met by offering multiple modes of learning[8]. For this reason, an AI tool should supplement rather than replace other modes of learning. When used effectively, AI can be a powerful tool for addressing learners' needs, but it is crucial to assess and mitigate any risks to learners.

| Q5 | Q6 |
| --- | --- |
| **Who are the intended users of the AI application and what are their unique learning needs?** Note: In a diverse classroom you may have multiple groups of users to consider. | **How can this AI application help meet the needs of these learners?** |
|  |  |

It is also important to also consider potential harms that could result if AI selection and use are not done well. Referencing the list of examples on the next page, take time to consider potential harms that might be associated with the AI application you are considering.

| Q7 |
|---|
| **What are the potential harms that could arise for these learners?** (See examples below) |
| |

**Examples of potential harms:**
- **Providing students with inaccurate information** can affect their academic performance and future understanding of the subject. (Learn more from Resource #4: Truth and AI).
- **Providing students with biased information** can lead to a distorted view of the subject matter, which can perpetuate stereotypes and misinformation. (Learn more from Resource #4: Bias and Discrimination). [*,**]
- **Sharing students' private Information** is a breach of privacy which can lead to identity theft, cyberbullying, or other forms of exploitation. (Learn more from Resource #4: Privacy. [*]
- **Failing to accommodate students' diverse needs:** AI that does not account for the needs of students with disabilities may lead to students being unable to fully participate in educational activities, widening the achievement gap. [*,***]
- **Excluding students** of lower socioeconomic status, those with limited connectivity, or those with disabilities. [*,**]
- **Sharing students' intellectual property** making their ideas available to others jeopardizes the originality of their work. (Learn more from Resource #4: Copyright and Intellectual Property) [*]
- **Unduly persuading or influencing students' opinions or actions:** Increasingly realistic AI applications can lead students to attribute inappropriate levels of trust in AI technology and potentially undermine their capacity to think critically and exercise independent judgement. (Learn more from Resource #7: Why We Think AI Has Feelings) [9, 10, 11]. [*,***]

*Relates to Section 1 of the OHRC: Ensures equal access to educational services (education is a service), including flexible learning support.
**Related to Section 11 of the OHRC: Ensures systemic barriers are not reinforced by resource limitations.
***Relates to Section 17 of the OHRC: Duty to accommodate disability to the point of undue hardship.
****Relates to the OHRC Preamble: Dignity, autonomy, and freedom from coercion.

Once potential harms have been identified, consider whether steps can be taken to mitigate them. Referencing the list provided below, consider potential mitigation strategies. Note that in Part B you will be provided with more detailed guidance on assessing risks and implementing mitigation strategies.

| Q8 |
|---|
| **Can these potential harms be mitigated?** (See examples below) |
| |

**Possible Mitigation Strategies:**
- **Choose accessible technologies:** Ensure technologies selected are free or low cost, do not require large downloads, minimize connectivity needs, and can accommodate the needs of diverse learners. [*, ****]
- **Ensure transparency:** Verify that AI decision-making processes are explainable and ensure that students have access to those explanations. [**,***]
- **Identify and address bias:** Regularly audit AI systems to identify and correct biases. [*, **]
- **Protect privacy:** Implement strict data privacy policies and limit data collection to what is necessary. Review data anonymization processes. [*]
- **Maintain human oversight** to ensure AI tools are used appropriately and effectively.
- **Solicit feedback** from students and educators and address identified issues promptly.
- **Framing and presentation:** Present AI applications as supplemental tools used to enhance learning, reminding students of their fallibility and limitations and emphasizing that AI is not a substitute for professors, experts, or mentors. Encourage students to critically evaluate AI-generated content and seek guidance from their instructors or peers when needed. Highlight the importance of ethical use, including respecting privacy, intellectual property, and maintaining academic integrity. [*, ****]

[*] Relates to Section 1 of the OHRC: Ensures equal access to educational services (education is a service), including flexible learning support.
[**] Related to Section 11 of the OHRC: Ensures systemic barriers are not reinforced by resource limitations.
[***] Relates to Section 5 of the OHRC: Equal treatment in employment and services includes the right to be free from opaque or discriminatory decision-making.
[****] Duty to Accommodate: Education is considered a service under the OHRC and providers must accommodate students with disabilities to the point of undue hardship.

After evaluating the potential harms associated with a specific AI application and determining whether these harms can be mitigated, it is important to consider whether alternate technologies are available that will better meet students' needs. When considering alternate technologies, include AI and non-AI applications.

| Q9 |
| --- |
| **What alternatives exist?** |
| |

## STEP 3: Consider the Educational Context

**Consider the context in which an AI tool operates and the population it affects in order to assess whether it is likely to impact human rights significantly.**

An educational context becomes high risk for human rights issues when AI tools are used to make decisions, assign a grade, or draw conclusions about a student. In these cases, there is a much higher potential for discrimination, inequality, and privacy violations[12]. Consider the following questions:

| | | | |
| --- | --- | --- | --- |
| **Q10** | **Does the AI application make a decision, or provide information or a score that may influence a high stakes decision?** AI tools used for grading or admissions decisions can have significant long-term impacts on students' academic and professional futures. If the AI system is biased or inaccurate, it could unfairly disadvantage certain groups (and also leave the educator or educational facility in a precarious legal position). | **YES** | **NO** |

| | | | |
| --- | --- | --- | --- |
| **Q11** | **Does the AI application analyze the student's face, facial expressions, fingerprints, voice, vocal patterns, movements, or behavior for the purpose of making predictions or drawing conclusions about the person?** AI applications that monitor students' online activities or physical behaviors during exams or at other times can infringe on students' privacy rights and create a stressful, intrusive environment. | **YES** | **NO** |

| Q12 | **Does the AI application make decisions on topics that impact vulnerable or protected populations or groups?** Could decisions made by the AI application negatively impact individuals in a protected category (age, race, sex, religion, disability, etc.) or those in a vulnerable circumstance (unwell, unemployed, or unhoused)? For example, an AI application that makes decisions impacting students with disabilities may fail to accommodate their specific needs or perpetuate existing biases. | YES | NO |
|---|---|---|---|

| Q13 | **Does the AI present itself as a teacher, expert, mentor, or authority?** <br> AI agents, especially embodied AI agents, have the potential to influence and persuade[13].  Risk associated with this ability are higher when the agent is designed to instruct, direct, or mentor students [14, 15]. | YES | NO |
|---|---|---|---|

| Q14 | **Does the AI impersonate a real person (celebrity, influencer, medical or academic expert, politician, etc.)?** <br> AI agents that impersonate real-life individuals not only violate the human rights of the person being impersonated but are deceptive and mislead those using the AI application. | YES | NO |
|---|---|---|---|

**Answering "yes" to any of Questions 10-14, indicates that AI use in this context may be at high risk for human rights concerns.**

Note that the European Union (EU) AI Act describes all educational use of AI as high risk and outline specific requirements for educational AI use in "Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators"[16].  Be sure to consult these guidelines if they are relevant to your context.

## STEP 4: Seek Feedback, Review, and Revise as Appropriate

**Solicit feedback on an ongoing basis, ensure it is reviewed regularly, and concerns are addressed.**

In his chapter in emerging technologies, Bates[17] discusses the importance of iterative design, where the learning activity is developed and tested. Feedback is gathered from learners, and that feedback is used to refine and revise the learning activity. An iterative development process is essential to ensure that preset learning outcomes are achieved, and learners' needs are met.  Use the questions below to plan how you will obtain, review, and respond to feedback.

| Q15 | Q16 | Q17 |
|---|---|---|
| How will feedback be solicited? | Who will review the feedback and how often? | How will concerns be addressed? |
| | | |

## PART B: Risk Mitigation Strategies

The following strategies can help you consider more carefully how to mitigate risks associated with this AI application and create a plan for ongoing reassessment.

### STRATEGY 1: Assess for Biased Outputs

AI models are developed and trained by individuals who have their own innate biases. The data used to train AI systems will reflect pre-existing societal biases as well.  Consequently, when AI is used without proper assessment, monitoring, or mitigation strategies, AI can exacerbate these biases[18]. However, when properly designed and used, AI has the potential to improve educational access and inclusion for all students and especially for learners with disabilities and special needs[29]. Therefore, it is important to monitor an AI application to determine whether it generates biased output that may negatively impact users, especially protected groups or individuals [20], and take steps to mitigate identified biases.

Use the questions and examples provided below to create a plan for identifying and mitigating biases in the AI applications you are hoping to use.

| Q18 |
|---|
| **What steps can be taken to assess the AI application to determine whether it currently produces biased or discriminatory output?** |
| |

**Examples:** Assessing biased outputs could involve:

- **Creating simulated scenarios** that mimic real-world use cases to observe how the AI application performs in diverse situations. This can help identify any unintended biases in its outputs.
- **Collecting feedback from diverse groups** of students and educators who interact with the AI application.

*Relates to Section 1 of the OHRC: Ensures equal access to educational services (education is a service), including flexible learning support.
**Related to Section 11 of the OHRC: Ensures systemic barriers are not reinforced by resource limitations.

| Q19 |
|---|
| **How can biases be mitigated?** (See examples below) |
| |

**Note:** The *Framework for Accessible and Equitable Artificial Intelligence (AI) in Education*[21] is an open access resource that looks more deeply into AI as a tool for eliminating educational barriers rather than reinforcing them or creating new ones. This framework also elaborates on the "human-in-the-loop" model that has informed this guide. A link to this resource is provided in the "Additional Resources" section.

> **Examples:**
> - **Choose AI models that have been trained on diverse data sets** and are continuously monitored and improved to minimize biases.
> - **Clearly document** how AI models are trained, the data sources used, and the steps taken to mitigate biases. Ensure transparency in AI development processes.
> - **Review fine-tuning materials for biases** before finetuning AI agents and, when possible, limit the AI agent's responses to content within these materials to ensure accuracy and fairness
> - **Train educators** using the AI application to look for an identify biased outputs.

*Relates to Section 1 of the OHRC: Ensures equal access to educational services (education is a service), including flexible learning support.
**Related to Section 11 of the OHRC: Ensures systemic barriers are not reinforced by resource limitations.

AI tools are constantly changing.  Even if you continue to use a single tool, training data or LLM use behind the scenes may take place without your awareness.  Therefore, it is important to continuously monitor AI systems to identify and address biases that may emerge over time. Solicit feedback regularly from educators and students to identify any biases in AI tools. Use this feedback to implement effective mitigation strategies.

| Q20 | Q21 |
|---|---|
| **How frequently will a bias audit occur?** | **Who will be responsible for acting on the recommendations arising from the bias audit process?** |
|  |  |

## STRATEGY 2: Consult with Interdisciplinary Experts

As AI technologies become more complex and provide more sophisticated responses, their ability to persuade and influence students will continue to grow [22, 23, 24]. The potential for harm will become more complex as well. To ensure that all potential risks are identified, and optimal mitigation strategies are implemented, it is important to consult others with diverse expertise. Choose experts from disciplines most relevant to your application and your learners. You may consider those with expertise in computer science, communications, data privacy, human

rights, ethics, psychology, sociology, educational technologies, teaching and learning, librarians, and those with expertise specific to your educational setting and student group. Invite the experts you feel can provide the most relevant input to share their concerns and propose strategies for mitigating potential risks.

| Q22 |
|---|
| **What experts have been or will be consulted? What concerns have they expressed and how will those concerns be addressed?** |
| |

## STRATEGY 3: Ensure Transparency

When AI tools have opaque decision-making processes[25], it becomes difficult to challenge or appeal unfair outcomes, such as those from automated proctoring systems.

| Q23 |
|---|
| **What can be done to make the AI tool's decision-making process more explainable and transparent?** |
| |

| Q24 |
|---|
| **How will the purpose and function of the AI application be communicated to students? instructors? others?** |
| |

The Artificial Intelligence Disclosure (AID) Framework[26] guides transparent attribution of the use of artificial intelligence in research or educational work.  Where appropriate, model transparency by including an Artificial Intelligence Disclosure Statement in your course materials.

| Q25 |
|---|
| **Draft your Artificial Intelligence Disclosure Statement.** |
| |

## STRATEGY 4: Plan for Regular Audits

AI technologies are constantly changing and so are the needs of learners. Create a plan to regularly review this document, the AIHRIAE, and use it to reassess the AI technology periodically. Review the data sources and/or fine-tuning resources used by the AI application to ensure they are accurate, complete, and free from biases. Review each of the questions in the AIHRIAE and review new alternate technologies, considering whether they could provide equal benefit to learners with fewer risks. Be prepared to make changes to the AI tool based on feedback and new insights to continuously improve its alignment with pedagogical and human rights principles.

| Q26 | Q27 |
|---|---|
| **What will future audits look like?** | **Who will be responsible for ensuring audits are done?** |
| | |

## Additional Resources

1. [Online Learning | Ontario Human Rights Commission](#) - Nine eLearning modules pertaining to human rights in Ontario.
2. [European Union (EU) AI Act](#) aims to foster the development of responsible and trustworthy AI in Europe and around the world. The [Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators](#) support the achievement of these goals.
3. [Framework for Accessible and Equitable Artificial Intelligence (AI) in Education](#) – A practical guide focusing equity and accessibility of AI in educational settings.
4. [Teaching AI Ethics: The Series – Leon Furze](#)  - Ten excellent blog posts pertaining to teaching AI ethics in an education context.

5. [The Artificial Intelligence Disclosure (AID) Framework](#) guides transparent attribution of the use of artificial intelligence in research or educational work
6. [An Inquirer's Guide to Ethics in AI](#) - Matthew Silk and Ian MacDonald explore the ethical dimensions of artificial intelligence through an interdisciplinary lens.
7. [The AI Consciousness Illusion: Why We Think AI Has Feelings - Deep Learning Dialogues](#) A 2024 study on public perceptions of AI consciousness, revealing that 67% attribute some level of consciousness to LLMs, highlighting that frequent use increases this perception.
8. [AI through an EDI lens - The need for human-centered design – AI Dialogues](#) - A podcast about engaging with generative AI through an EDI lens, highlighting the importance of human-centered design to mitigate bias, empower educators, and maintain transparency.
9. [The Ontario Human Rights Code | Ontario Human Rights Commission](#)
10. [Canadian Charter of Rights and Freedoms](#)
11. [Universal Declaration of Human Rights | United Nations](#)

# Acknowledgements

In the AIHRIA for Educators, I advocate for seeking interdisciplinary input when selecting or developing AI applications for education, so it is only fitting that I did the same as I developed this tool. Thanks to the team of interdisciplinary experts who shared their thoughts and provided feedback:

- Kari Weaver: Learning, Teaching, and Instructional Design Librarian, University of Waterloo
- Ken McKay: Management Sciences Professor, University of Waterloo
- W. Jim Jordan: Ethicist and co-author of *With a Clear Conscience: Business Ethics, Decision-Making, and Strategic Thinking* (Oxford University Press); senior software test developer, SkillsWave Corporation, Kitchener, ON.

## Limitations

This guide focuses on helping educators evaluate and decide whether to use AI tools in their classrooms, ensuring that human rights considerations are central to their decision-making process

This document does not constitute legal advice and does not provide a definitive legal answer regarding any adverse human rights impacts, including violations of federal or provincial human rights law or other relevant legislation.

# References

1. Law Commission of Ontario, & Ontario Human Rights Commission. (2024). *Human rights AI impact assessment*. https://www3.ohrc.on.ca/sites/default/files/Human%20Rights%20Impact%20Assessment%20for%20AI.pdf

2. The Ontario Human Rights Code and system. The Ontario Human Rights Code | Ontario Human Rights Commission. (2024). https://www3.ohrc.on.ca/en/ontario-human-rights-code

3. *Canadian Charter of Rights and Freedoms*, s 7, Part I of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), 1982, https://laws-lois.justice.gc.ca/eng/Const/page-12.html

4. United Nations. (2024). *Universal declaration of human rights*. United Nations. https://www.un.org/en/about-us/universal-declaration-of-human-rights

5. UNESCO. (2024). *Harness the benefits of AI while ensuring protection & promotion of human rights.* https://core.unesco.org/en/project/32106B3011

6. Bates, A. W. (2022). *Teaching in a digital age: Guidelines for designing teaching and learning* (3rd ed., Chapters 9 and 10). Tony Bates Associates Ltd.

7. Bates, A. W. (2022). *Teaching in a digital age: Guidelines for designing teaching and learning* (3rd ed., Chapters 9 and 10). Tony Bates Associates Ltd.

8. The UDL guidelines. UDL. (2022). https://udlguidelines.cast.org/

9. Colombatto, C., & Fleming, S. M. (2023). Folk Psychological Attributions of Consciousness to Large Language Models. https://doi.org/10.31234/osf.io/5cnrv

10. Divekar, R. R., Drozdal, J., Chabot*, S., Zhou, Y., Su, H., Chen, Y., ... & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332-2360. https://doi.org/10.1080/09588221.2021.1879162

11. Lim, S., Schmälzle, R., & Bente, G. (2024). Artificial social influence via human-embodied AI agent interaction in immersive virtual reality (VR): Effects of similarity-matching during health conversations. *arXiv preprint arXiv:*2406.05486.

12. Silk, M. S. W., & MacDonald, I. J. (2024). An Inquirer's Guide to Ethics in AI. Chapter 6. Retrieved from https://platform.virdocs.com/read/2878065/2/#/4/2

13. Colombatto, C., & Fleming, S. M. (2023). Folk Psychological Attributions of Consciousness to Large Language Models. https://doi.org/10.31234/osf.io/5cnrv

14. Divekar, R. R., Drozdal, J., Chabot*, S., Zhou, Y., Su, H., Chen, Y., ... & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332-2360. https://doi.org/10.1080/09588221.2021.1879162

15. Lim, S., Schmälzle, R., & Bente, G. (2024). Artificial social influence via human-embodied AI agent interaction in immersive virtual reality (VR): Effects of similarity-matching during health conversations. *arXiv preprint arXiv:*2406.05486.

16. European Commission: Directorate-General for Education, Youth, Sport and Culture, *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*, Publications Office of the European Union, 2022, https://data.europa.eu/doi/10.2766/153756

17. Bates, A. W. (2022). *Teaching in a digital age: Guidelines for designing teaching and learning* (3rd ed., Chapters 9 and 10). Tony Bates Associates Ltd.

18. Silk, M. S. W., & MacDonald, I. J. (2024). An Inquirer's Guide to Ethics in AI. Chapter 4. Retrieved from https://platform.virdocs.com/read/2878065/2/#/4/2

19. Furze, L. (2023). Artificial Intelligence and education: The role of librarians in promoting ethical and responsible AI. *Synergy, 21*(1).

20. Silk, M. S. W., & MacDonald, I. J. (2024). An Inquirer's Guide to Ethics in AI. Chapter 4. Retrieved from https://platform.virdocs.com/read/2878065/2/#/4/2

21. Lo, L., Mitchell, J., Scheuhammer, J., Treviranus, J., & Spence, R. (2024). Framework for accessible and equitable artificial intelligence (AI) in education. Framework for Accessible and Equitable Artificial Intelligence (AI) in Education

22. Colombatto, C., & Fleming, S. M. (2023). Folk Psychological Attributions of Consciousness to Large Language Models. https://doi.org/10.31234/osf.io/5cnrv

23. Divekar, R. R., Drozdal, J., Chabot*, S., Zhou, Y., Su, H., Chen, Y., ... & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332-2360. https://doi.org/10.1080/09588221.2021.1879162

24. Lim, S., Schmälzle, R., & Bente, G. (2024). Artificial social influence via human-embodied AI agent interaction in immersive virtual reality (VR): Effects of similarity-matching during health conversations. *arXiv preprint arXiv:*2406.05486.

25. Silk, M. S. W., & MacDonald, I. J. (2024). An Inquirer's Guide to Ethics in AI. Chapter 4. Retrieved from https://platform.virdocs.com/read/2878065/2/#/4/2

26. Weaver, K. (2024). The artificial intelligence disclosure (AID) framework. *College & Research Libraries News 85*(10). https://crln.acrl.org/index.php/crlnews/article/view/26548/34482

## Artificial Intelligence Disclosure Statement:

Artificial Intelligence Tool: Microsoft Copilot (University of Waterloo institutional instance); Conceptualization: Microsoft Copilot was used to suggest previously unidentified benefits, risks, and mitigation strategies; Writing – Review & Editing: Microsoft Copilot was used to generate synonyms, refine wording, and improve flow.

**Share your feedback on this tool by completing a short survey:**



Cite as:
Long, L. (2024). AI Human Rights Impact Assessment for Educators.
https://uwaterloo.ca/conflict-management-human-rights/ai-human-rights-impact-assessment-tools

## Glossary of Key Terms

**A**

**Accessibility:** Designing AI tools to be usable by all students, including those with disabilities.

**AI Application:** A software system that uses artificial intelligence to perform tasks such as tutoring, grading, or content generation.

**AI Disclosure Statement:** A formal declaration explaining how AI was used in educational content or decision-making.

**Algorithm:** A set of rules or instructions used by AI systems to process data and make decisions. Algorithms can introduce bias if not carefully designed and tested.

**Artificial Intelligence (AI):** The simulation of human intelligence in machines that are programmed to think and learn.

**Audit (Bias Audit / Risk Audit):** A systematic review of an AI system to identify and address potential biases or risks in its outputs, training data, or implementation.

**B**

**Bias:** Systematic favoritism or prejudice in AI systems that can lead to unfair outcomes.

**Bias Audit:** A structured evaluation to detect and mitigate bias in AI systems.

**Biased Output:** AI-generated results that reflect or perpetuate unfair prejudices or stereotypes.

**C**

**Contextual Analysis:** A human rights-based approach that evaluates the full impact of AI use on individuals or groups, especially in protected social areas like education.

**D**

**Data:** Information collected, stored, and processed by AI systems, including student performance, behavior, or personal identifiers.

**Data Privacy:** The protection of personal information collected and used by AI systems, especially sensitive student data.

**Discrimination:** Adverse treatment or impact on individuals based on protected characteristics, as defined by human rights legislation.

**E**

**Educational Context:** The specific learning environment in which an AI tool is used, including its pedagogical purpose and user population.

**Equity:** Ensuring fair access and treatment for all students, particularly those from marginalized or protected groups.

**Explainability:** The ability to understand and interpret how an AI system makes decisions.

**F**

**Fairness:** The principle that AI systems should treat all users equitably, without discrimination or favoritism.

**Feedback Mechanism:** A process for collecting and responding to input from students and educators about an AI tool's performance.

**Framing and Presentation:** The way AI tools are introduced and contextualized for learners, emphasizing their limitations and ethical use.

**H**

**Hallucination:** When an AI system generates factually incorrect, fabricated, or misleading information.

**High-Risk Context:** Situations where AI tools influence significant decisions (e.g., grading, admissions) or affect vulnerable populations.

**Human Oversight:** The involvement of educators or administrators in monitoring and guiding AI use to prevent harm.

**I**

**Informed Consent:** Ensuring users understand and agree to how their data will be used by AI systems.

**Intellectual Property:** Original student work that must be protected from unauthorized use or reproduction by AI systems.

**Interdisciplinary Consultation:** Engaging experts from multiple fields (e.g., ethics, law, education, computer science) to assess AI risks and benefits.

**L**

**LLM (Large Language Model):** A type of AI model trained on vast amounts of text data to generate human-like language (e.g., ChatGPT, Copilot).

**M**

**Mitigation Strategies:** Actions taken to reduce or eliminate potential harms caused by AI tools.

**P**

**Pedagogy:** The theory and practice of teaching, especially as it relates to how AI tools align with instructional goals.

**Prohibited Grounds:** Characteristics protected under the Ontario Human Rights Code that cannot be used as a basis for discrimination. These include race, ancestry, place of origin, colour, ethnic origin, citizenship, creed (religion), sex (including pregnancy and breastfeeding), sexual orientation, gender identity, gender expression, age, marital status, family status, disability, and receipt of public assistance (in housing).

**Protected Areas:** Social areas where individuals are legally protected from discrimination under the Ontario Human Rights Code. These include services (such as education), employment, housing, contracts, and membership in unions or professional associations.

**R**

**Risk Assessment:** The process of identifying and evaluating potential harms associated with AI use in education.

**T**

**Transparency:** Openness about how AI systems function, make decisions, and use data.

**U**

**Universal Design for Learning (UDL):** An educational framework that promotes inclusive teaching by offering multiple means of engagement, representation, and expression.

**Appendix A**

**Understanding Bias, Impact, and Discrimination Under the Ontario Human Rights Code**

AI systems used in education can unintentionally produce biased outcomes that disproportionately affect students from protected groups. These biases may stem from the data used to train the AI, the design of the system itself, or the context in which it is deployed. While these impacts may not be deliberate, they can still result in discrimination under the **Ontario Human Rights Code**.

The Code outlines a three-part test to determine whether discrimination has occurred:

1. The individual has a characteristic protected under the Code (e.g., disability, race, sex, age, etc.);
2. The individual experiences adverse treatment or impact in a protected social area (such as education, which is considered a "service" under the Code); and
3. The protected characteristic is a factor in the adverse impact.

Human rights law requires a **contextual analysis**—educators must consider not only the AI tool's intended function but also its real-world effects. Since education is a protected area under the Code, educators have a legal and ethical responsibility to assess whether AI tools may cause harm, and to take steps to mitigate those harms. This includes identifying potential biases, evaluating their impact, and ensuring that AI use supports equitable and inclusive learning environments.

**Example 1: Personalized Learning Platform**

**Scenario:** An AI platform uses historical performance data to recommend learning paths. Students from underrepresented communities are disproportionately steered toward less challenging content.

**Discrimination Analysis:**

1. **Protected characteristic:** Race, socioeconomic status (linked to systemic disadvantage).
2. **Adverse impact:** Students are denied access to more advanced learning opportunities.
3. **Causal link:** The AI's reliance on historical data reflects and reinforces systemic inequities.

**Conclusion:** This may constitute discrimination under the Code.

**Example 2: AI-Based Proctoring System**

**Scenario:** An AI proctoring tool flags students for "suspicious behavior" based on facial expressions or movement. Students with tics or neurodivergent behaviors are disproportionately flagged.

**Discrimination Analysis:**

1. **Protected characteristic:** Disability (e.g., Tourette's, autism).
2. **Adverse impact:** Students are subjected to increased scrutiny or disciplinary action.
3. **Causal link:** The AI interprets disability-related behaviors as misconduct.

**Conclusion:** This may constitute discrimination under the Code.

**Example 3: AI-Generated Study Plans**

**Scenario:** An AI tool generates study plans based on time availability and learning goals. It assumes all students have flexible schedules and access to quiet study environments, disadvantaging students with caregiving responsibilities or limited home resources.

**Discrimination Analysis:**

1. **Protected characteristic:** Family status, socioeconomic status.
2. **Adverse impact:** Students are given unrealistic or unhelpful plans, leading to academic stress or failure.
3. **Causal link:** The AI fails to account for diverse life circumstances in its recommendations.

**Conclusion:** This may constitute discrimination under the Code.

**Appendix B**

**International Alignment Table: AIHRIAE and Global Human Rights Frameworks**

This table illustrates how the AI Human Rights Impact Assessment for Educators (AIHRIAE) aligns with key principles found in international human rights and AI governance frameworks. It is intended to support users outside of Canada in understanding the tool's relevance and applicability across diverse legal and ethical contexts. The table maps AIHRIAE's core principles, such as human dignity, transparency, bias mitigation, and accessibility, against corresponding standards from the Ontario Human Rights Code, the Canadian Charter of Rights and Freedoms, UNESCO's AI ethics recommendations, and the European Union's AI Act. This alignment reinforces the AIHRIAE's commitment to global best practices in the responsible and ethical use of AI in education.

| Principle | AIHRIAE | Ontario Human Rights Code/ AIHRIA | Canadian Charter of Rights and Freedoms | UNESCO Recommendation on the Ethics of AI | UNESCO Generative AI Guidance in Education | EU AI Act |
|---|---|---|---|---|---|---|
| Human Rights & Dignity | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |
| Transparency & Privacy | ☑ | ☑ | | ☑ | ☑ | ☑ |
| Bias Mitigation | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |
| Identification & Consideration of High-risk Contexts | ☑ | ☑ | | ☑ | ☑ | ☑ |
| Human Oversight | ☑ | ☑ | | ☑ | ☑ | ☑ |
| Inclusivity & Accessibility | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |

**Descriptions of Key Principles**

- **Human Rights & Dignity:** Focuses on upholding the inherent dignity, autonomy, and equality of all individuals. In the context of AI in education, this means ensuring that technologies do not undermine students' rights or well-being.
- **Transparency & Privacy:** Emphasizes the need for clear, understandable information about how AI systems function, make decisions, and handle personal data. This includes ensuring informed consent and protecting student privacy.
- **Bias Mitigation:** Involves identifying, assessing, and mitigating the impact of algorithmic bias that could lead to unfair treatment or discrimination. This principle supports equity and fairness in AI-supported educational environments.
- **Identification & Consideration of High-Risk Contexts:** Encourages proactive assessment of situations where AI use could significantly impact learners' rights, such as grading, admissions, or behavioral monitoring. These contexts require heightened scrutiny and safeguards.
- **Human Oversight:** Reinforces the importance of maintaining meaningful human control over AI systems. Educators should remain the final decision-makers and ensure that AI tools support, not replace, instructor interactions.
- **Inclusivity & Accessibility:** Promotes the design and use of AI tools that are accessible to all learners, including those with disabilities or from marginalized groups. This principle aligns with Universal Design for Learning (UDL) and equity-focused practices.

**International Human Rights and AI Governance Frameworks**

- Ontario Human Rights Code
- Canadian Charter of Rights and Freedoms
- UNESCO Recommendation on the Ethics of Artificial Intelligence
- UNESCO Guidance for Generative AI in Education and Research
- European Union Artificial Intelligence Act (EU AI Act)

**Appendix C**

**Sample Use Cases: AIHRIAE in Action**

The following scenarios are provided to help AIHRIAE users explore the use of AI tools in educational contexts while critically reflecting on their potential human rights impacts.

**Scenarios 1–4** reflect the perspective of a course instructor. The tools featured are freely available through the University of Waterloo and are designed to protect user privacy and intellectual property. These include the institutional instance of Microsoft Copilot and Contact North's AI Tutor Pro. These tools offer a secure environment for experimentation and learning. Each scenario illustrates how the **AIHRIAE** can support thoughtful planning and decision-making.

**Scenario 5** takes a different approach. In this case, users are asked to select any AI tool they believe could be used to solve a real-world problem. However, they do not actually use the tool to generate a solution. Instead, they conduct a critical assessment of the tool using the **AI Human Rights Impact Assessment for Students (AIHRIAS)**, focusing on its potential human rights impacts and ethical considerations.

**Scenarios 6–10** expand the focus beyond classroom instruction to include other roles in educational settings, such as instructional designers, librarians, teaching assistants, educational technologists, and accessibility specialists. These scenarios illustrate how AI tools may be used in support roles and institutional decision-making, and how the **AIHRIAE** can be applied in those contexts.

These scenarios are intended to help users situate the AIHRIAE in a real-world context. You are encouraged to select the scenario that feels most relevant to your own work, discipline, or interests as you complete the exercise.

**Scenario 1: Human vs. AI Solutions: Group Problem-Solving with Microsoft Copilot**

As a course instructor planning for a course you will be teaching this fall, you want students to work in groups to propose a solution to a problem using a method you have taught for several years. After that, you want them to use the UW institutional instance of Microsoft Copilot to propose a solution to the same problem. Finally, you want students in their groups to reflect on and critique the AI-generated solution. Before implementing this activity, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 2: Foundations First: Personalized Review with Contact North's AI Tutor Pro**

As a course instructor, you have identified that some students coming into your first-year course do not have a solid understanding of certain foundational concepts. Over several years, you have developed resource sheets to help students review these concepts. You want to upload these resource sheets into Contact North's AI Tutor Pro. Then you will be able to provide students with a link to their own private tutor that can quiz them on those materials. Before implementing this tool, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 3: Personalized Study Planning with Microsoft Copilot**

You've noticed that students in your second-year course often struggle with managing their time and reviewing prerequisite material. To help them plan more effectively, you are planning to ask students to use the UW instance of Microsoft Copilot to generate a personalized study plan based on the course syllabus, their learning goals, and key assessments. They will revise the plan mid-term and reflect on how Copilot supported their learning at the end of the course. Before introducing this activity, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 4: Structuring Group Work with Microsoft Copilot**

Students in your third-year design course often find it difficult to coordinate group work and stay on track during long-term projects. To support collaboration, you are considering asking each group to use Microsoft Copilot to plan three structured work sessions, including agendas, roles, and deliverables. They will submit their plans for feedback and reflect on how the sessions supported their final presentation. Before implementing this activity, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 5: Comparing Human and AI Solutions with a Human Rights Lens**

In class discussions, it has become clear that students in your senior ethics course sometimes struggle to critically assess the ethical implications of AI tools. To build this skill, you ask them to identify a real-world problem they believe AI could help solve. They then choose an AI tool they would like to use to generate a solution. Finally, they use the AIHRIAS tool to complete a human rights impact assessment and reflect on the ethical implications of using AI in their chosen context.

**Scenario 6: Collaborative Course Enhancement with Microsoft Copilot**

As an instructional designer, you have been asked to support an instructor who is redesigning a large undergraduate course. The instructor wants to improve how students engage with weekly readings and prepare for assessments. You are considering suggesting using the UW instance of Microsoft Copilot to co-develop study guides and learning activities. Before doing so, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 7: AI Research Support with a Privacy Lens**

As a university librarian, you are exploring how the institution's instance of Microsoft Copilot can support students in developing information literacy skills. You plan to demonstrate how Copilot can assist with brainstorming research questions, locating relevant sources, and organizing search strategies during library instruction sessions. Before integrating it into your workshops, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 8: AI-Generated Study Guides for Student Support**

As a TA in a first-year biology course, you notice that students are struggling to synthesize lecture content. To support their learning, you use Microsoft Copilot to generate weekly study guides based on lecture slides and textbook chapters. These guides include summaries, key terms, and sample questions. You complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 9: Evaluating Campus-Wide AI Tools**

As an educational technologist, you have been piloting a new AI-based discussion moderation tool for use in large online courses. The tool is designed to summarize threads, flag inappropriate content, and suggest follow-up questions. Before recommending broader adoption, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.

**Scenario 10: Reviewing AI Accessibility Tools**

As an accessibility advisor, you are evaluating an AI-powered note-taking tool that transcribes and summarizes lectures in real time. The tool is being considered for use by students with learning disabilities. Before presenting it at your team meeting, you complete the AIHRIAE to assess benefits, evaluate risks, and develop mitigation strategies.