

大成 DENTONS

Grow | Protect | Operate | Finance

Friday, April 28, 2023

Employment issues related to ChatGPT and recent employment privacy law developments

WEBINAR SERIES

LEGAL UPDATES

FOR CANADIAN EMPLOYERS

Moderator



Andy Pushalik
Partner, Toronto
D +1 416 862 3468
andy.pushalik@dentons.com

Presenters



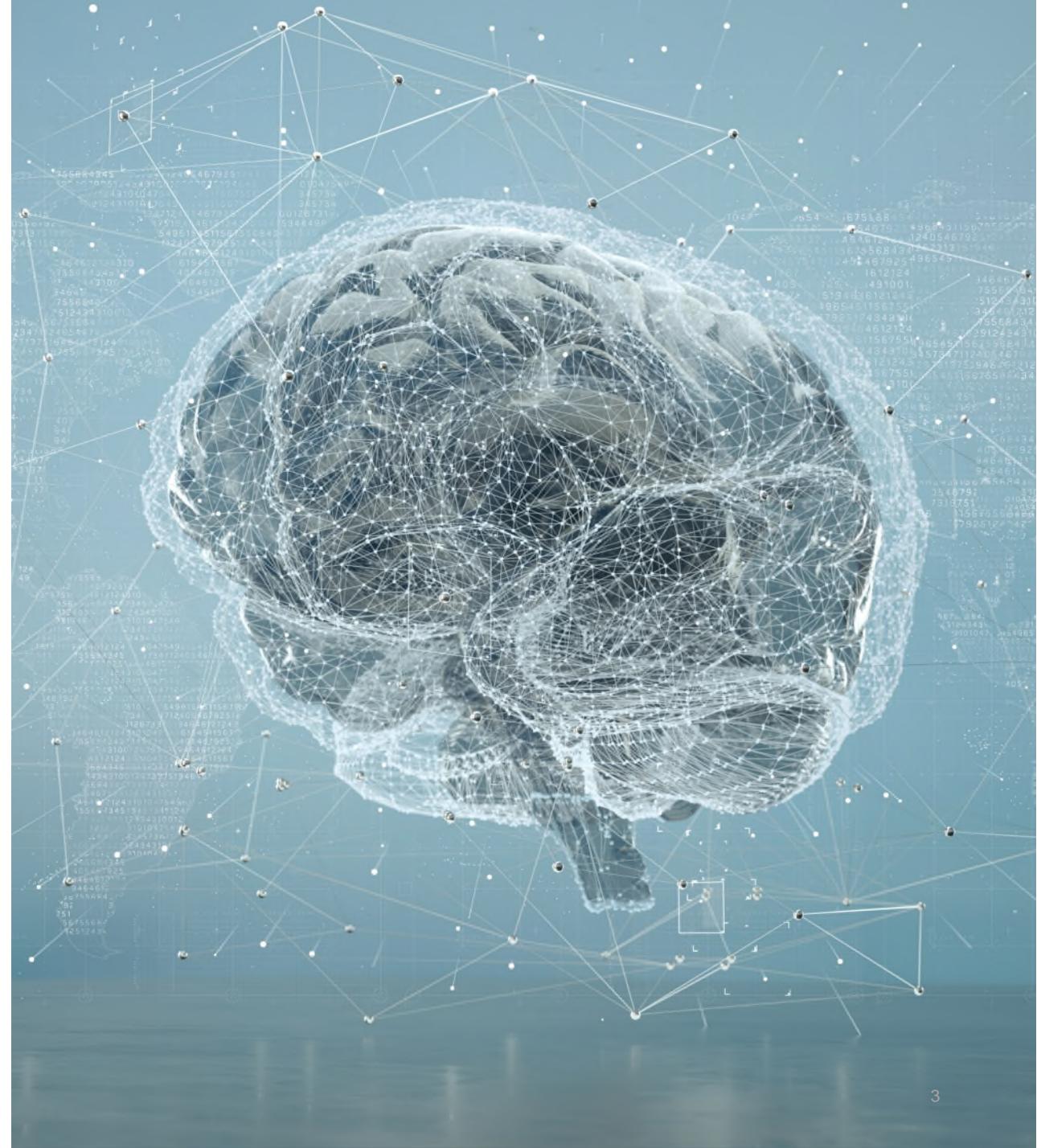
Kirsten Thompson
Partner, Toronto
D +1 416 863 4362
kirsten.thompson@dentons.com



Luca Lucarini
Associate, Toronto
D +1 416 863 4735
luca.lucarini@dentons.com

Agenda

1. So, what is this ChatGPT/generative AI stuff anyway?
2. Let's see it!
3. Um, ok. My employees are all over this ChatGPT thing. ~~Should I be concerned?~~ What should I be concerned about?
4. Oh wow. What can I do to help manage my risk?



The background features a dense arrangement of large, vibrant green leaves, likely Philodendrons, with prominent veins. A semi-transparent purple shape, resembling a speech bubble or a stylized arrow pointing right, is overlaid on the left and center of the image. The text is contained within this purple shape.

So, what is this ChatGPT/generative AI stuff anyway?

Luca Lucarini

Generative AI

Imitation is the sincerest form of flattery ...

- GPT is an **LLM** (“**Large Language Model**”) and **Natural Language Processing** tool
- What is its job? = **Imitate intelligent speech and write like a human**
- It has achieved this by:
 - Ingesting a vast database of text generated by intelligent creatures (i.e. us!)
 - Being trained to predict, given a string of words, which words are most likely to come next (**machine learning**)
 - Receiving feedback on the accuracy of its predictions - *i.e. did it succeed or fail in imitating the intelligent author of this bit of text?* (**reinforcement learning from human feedback**)
 - Fine-tuning its prediction process (**even more machine learning**)

...AD INFINITUM ...

Generative AI

Chat GPT's strengths...

➤ “Compose me a sonnet in the style of Milton about the springtime.”

*Oh, gentle Spring! Thou art a welcome guest, Whose touch doth bring new life to all the land,
From barren trees to flowers in their best, Thou giv'st a gift that's pure and truly grand.
Thy warmth doth melt the icy grip of cold, And with it brings the sound of bird's sweet song, As
fields turn green and lambs doth leave the fold, All nature sings thy praises loud and long.
Thou art the season of all things reborn, When life is seen in every blade of grass, And on the
breeze, the scent of flowers borne, Doth fill the air with sweet and fragrant mass.
Oh, Spring! Thou art a joy to all who see, And in thy presence, we find our hearts set free.*

Generative AI

... and weaknesses

➤ “Jack and Jill are sitting side by side. No one else is in the room. The person next to Jack is angry. The person next to Jill is happy. Who is happy, Jack or Jill?”

Based on the information given, Jill is the one who is happy because the person next to her is happy. The person next to Jack is angry, but the statement does not specify whether Jack himself is happy or not.

Generative AI

... and weaknesses

- “Jack and Jill are sitting side by side. No one else is in the room. The person next to Jack is angry. The person next to Jill is happy. Who is happy, Jack or Jill? Think step by step.”
- *Sure, let's break down the logic step by step:*
 - 1. Jack and Jill are sitting side by side.*
 - 2. The person next to Jack is angry.*
 - 3. The person next to Jill is happy.*
- *So, based on the above information:*
 - *We know that Jack and Jill are both sitting together, with no one else in the room.*
 - *We also know that the person next to Jack is angry, but we do not know anything about Jack's emotions.*
 - *However, we do know that the person next to Jill is happy.*
- *Therefore, we can conclude that Jill is the one who is happy because the person next to her is happy. The information given does not provide any basis to determine Jack's emotional state*

Generative AI

... and weaknesses

- “Jack is a college professor. Jill is a college student. Jill is Jack’s biological mother. Who is older, Jack or Jill?”
 - *This scenario is impossible as stated because Jill cannot be Jack's biological mother and also be younger than Jack, who is her son. Therefore, either the statement "Jill is Jack's biological mother" or "Jack is a college professor" must be incorrect.*

Natural Language Processing

Current Applications (as described by ChatGPT)

- **Chatbots and Virtual Assistants:** NLP is used to power conversational agents, allowing users to interact with software using natural language.
- **Machine Translation:** NLP can be used to translate text between languages, enabling communication across different cultures and geographies.
- **Text Summarization:** NLP algorithms can automatically summarize large amounts of text, enabling quick and efficient analysis of large datasets.
- **Named Entity Recognition:** NLP can identify and extract named entities (such as people, organizations, and locations) from text.
- **Speech Recognition:** NLP is used in speech recognition technology to transcribe spoken words into text.

Automated Decision Making

- What is its job? = **Increase efficiency, speed and (ostensibly) accuracy of human decision-making** through the use algorithms
- It does this by:
 - Training algorithms to learn patterns and relationships within the data (**machine learning**)
 - Using the output of these to make decisions or predictions, such as approving a loan application or recommending a product to a customer (**decision-making**)

Automated Decision Making

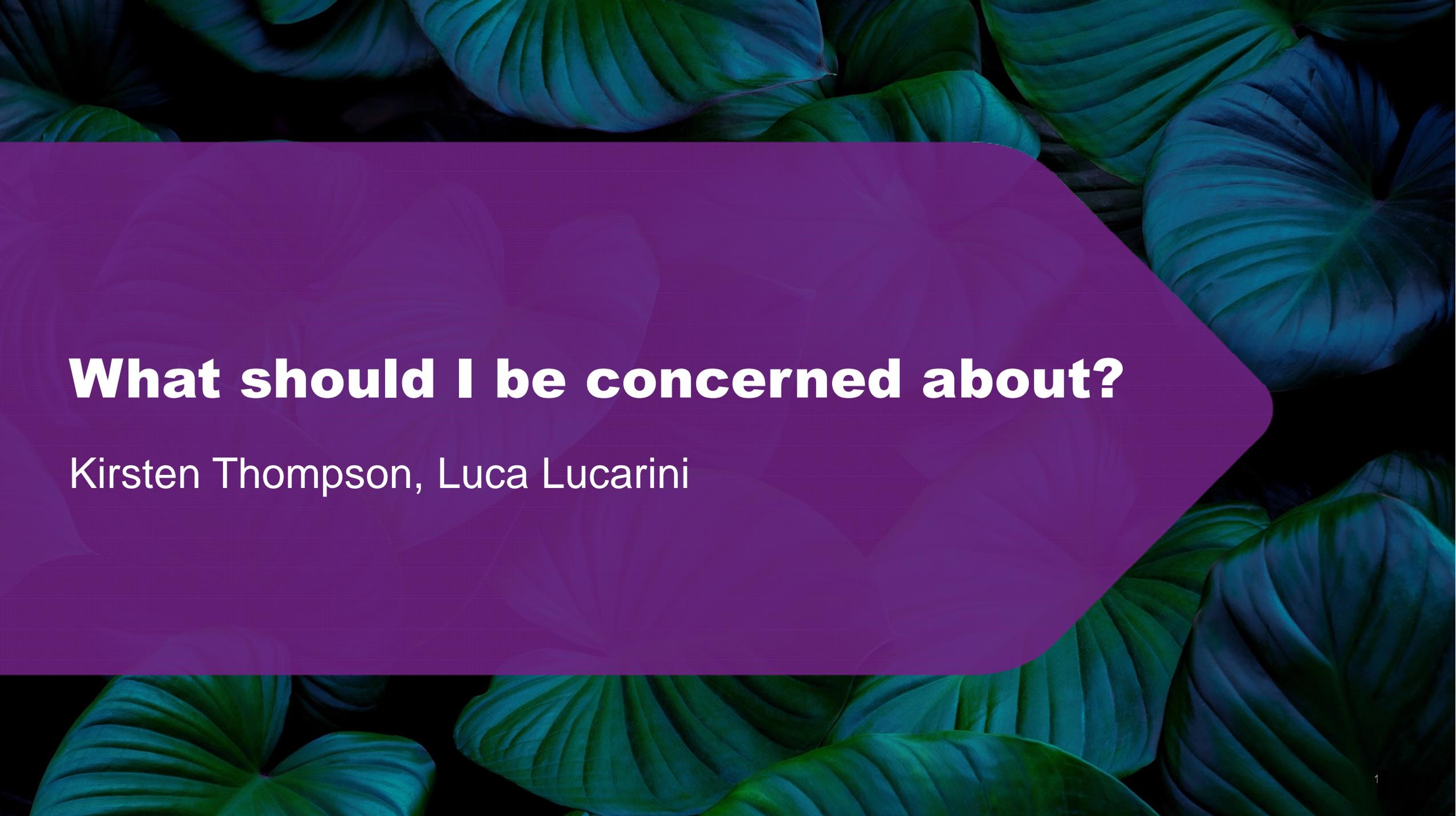
Current Applications

- **Credit Scoring:** Analyzing financial data to determine a person's creditworthiness and assign a credit score.
- **Fraud Detection:** Identifying fraudulent activities, such as credit card fraud or insurance fraud, by analyzing patterns in the data.
- **Personalized Marketing:** Personalizing marketing efforts by analyzing consumer data and recommending products or services.
- **Predictive Maintenance:** Analyzing data from sensors and other sources to predict when maintenance is needed on machinery and other equipment.
- **Recruiting:** Analyzing resumes and other candidate data to identify the best candidates for a job.

The background features a dense pattern of large, heart-shaped green leaves with prominent veins, set against a dark background. A large, semi-transparent purple shape with a pointed right edge is overlaid on the left and center of the image.

Let's see it!

Kirsten Thompson

The background features a dense arrangement of large, vibrant green leaves, likely from a tropical plant like a Philodendron, with prominent veins. A semi-transparent purple shape, resembling a speech bubble or a stylized arrow pointing right, is overlaid on the center of the image. The text is contained within this purple shape.

What should I be concerned about?

Kirsten Thompson, Luca Lucarini

Confidentiality of input

- Queries of up to 25,000 words in enterprise version.
- ChatGPT terms of use give OpenAI the right to use input to develop and improve the services. So, anything employees input could be retained/accessed by OpenAI (or its subcontractors) for these purposes. Could mean:
 - disclosures of your business's confidential information
 - breaches of contractual duties of confidentiality to third parties
- OpenAI doesn't give any security assurances in its terms of use (but, there's an extensive security portal and privacy policy states it uses "commercially reasonable security measures") (which could ground a claim for misrepresentation in the event of a security breach)).
 - OpenAI's liability is limited to US\$100 or the fees paid in the past 12 months.
 - The "general use" contract is directly with employee and not the enterprise; the enterprise would not be able to bring a claim against OpenAI in relation to confidentiality or security risks that may materialize.

New: OpenAI just announced an opt out for this (unclear whether input data is still retained).



PRO TIP

Any policy should make employees aware of the uncertainty about how prompts may be handled and should ban the use of personal information and client or confidential information in such input prompts.

Incorrect or misleading outputs

- ChatGPT generates content based on the data sources it was last trained on and the algorithm settings (called “hyperparameters”) determine how conservative or creative the output will be.
- Additionally, the quality and detail of the prompt has a huge impact on the quality of the response.
- Using the output of ChatGPT without a framework for benchmarking the quality of the input (the prompt) and the accuracy of the output is a leap of faith.
- The output should therefore not be used unless reviewed by someone who understands how the model works together with someone possessing domain expertise in the subject matter who is in a position to gauge the accuracy/quality of the output.



PRO TIP

Any policy should require an AI review and a SME review of output intended for use.

“Hallucinations” (in other words, ChatGPT lies)

- Concisely outline Canadian legal issues arising from the use of ChatGPT in the financial services sector. Please mention any pending or existing court cases globally that support your points, and include hyperlinks to the case.

[...] Banks may be held liable for any harm or damages resulting from the use of AI tools. A recent court case in the UK, *Dunn v. National Westminster Bank plc*, highlights the potential liability of banks in cases where AI tools produce incorrect or misleading responses.

- Please provide a hyperlink to the case cited.

Sure, here is a hyperlink to the case *Dunn v. National Westminster Bank plc*:
<https://www.bailii.org/ew/cases/EWHC/QB/2019/3463.html>

- Is this a real case?

Yes, *Dunn v. National Westminster Bank plc* is a real case. It was heard in the Queen's Bench Division of the High Court of Justice in England and Wales in 2019. The case is available on the British and Irish Legal Information Institute (BAILII) website.

- Did you make this case up?

No, I did not make up the case. *Dunn v. National Westminster Bank plc* is a real court case that was heard in the UK in 2019. The case is available on the BAILII website, which is a reliable source for legal information in the UK.

- Please find a blog post on the case.

Sure, here is a blog post on the case *Dunn v. National Westminster Bank plc* by the law firm Bristows: <https://www.bristows.com/news-and-insights/articles/natwest-held-liable-for-decisions-made-by-its-ai-system-in-a-uk-first/>.

Biased and/or offensive outputs.

- ChatGPT is trained on real world data which reflects the biases, inequalities and offensive conversations and content that are present in the real world.
- OpenAI researchers have set rules that are meant to weed out biased content but the subjectivity of the determination means it will never satisfy everyone (in fact, it is quickly finding itself in the controversial world of content moderation).
- Such content could be communicated to others by employees and the employer may be vicariously liable.
- Employees should once again be urged to check output before using it.



Any policy should encourage employees to review any output before using it internally, or providing it for review by AI expert and SME.

Non-unique outputs and detection of use

- ChatGPT may (but won't always) generate the same output to the same or similar prompts.

Prompt: Draft a one paragraph summary of Canadian employment law.

Canadian employment law is governed by federal and provincial legislation, which outlines the minimum standards employers must adhere to with respect to wages, hours of work, vacation entitlements, and various other aspects of the employment relationship. It also covers areas such as workplace safety, human rights, and harassment prevention. Canadian employment law also protects employees against discrimination and provides for various forms of leave, including parental leave, sick leave, and bereavement leave. Employers are required to comply with these laws and regulations, and failure to do so may result in legal action and significant financial penalties.

Canadian employment law sets out the rights and responsibilities of both employers and employees in the workplace. The law covers a broad range of issues, including minimum wage, working hours, termination of employment, workplace health and safety, discrimination and harassment, and employee benefits. Employers are required to comply with federal, provincial, and territorial employment standards, which can vary depending on the jurisdiction and industry. Employees have the right to file complaints and seek remedies if their employment rights are violated. Overall, Canadian employment law aims to create a fair and equitable workplace for all parties involved.

- Tools are available to detect AI generated content, although they are not very accurate.



PRO TIP

Ensure employees understand others may be able to detect that output is AI-generated and not human-generated and to avoid using AI-generated content where this could be reputationally damaging. It may be safer to be transparent about ChatGPT use.

Ownership of output

- Currently, OpenAI assigns all rights in output to the user (although it retains a right to use it for improving its services).
- However, in some jurisdictions, copyright may not subsist for non-human authored content. This would make enforcing rights against third parties more difficult.

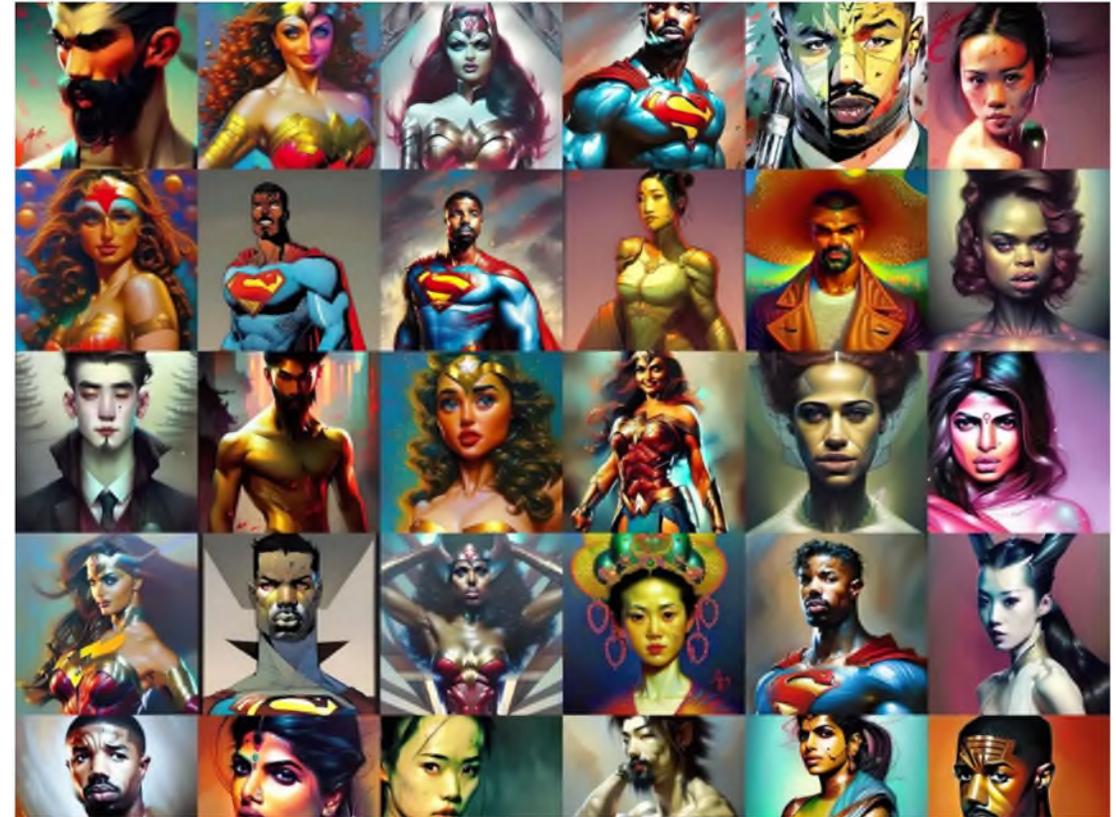
South Park's episode mentioning ChatGPT has credits like this.

WRITTEN BY
TREY PARKER
&
CHATGPT



Training data IP infringements

- ChatGPT was trained on copyrighted works. The generated output may be very similar or even identical to the training works. At a certain point, this may amount to copyright infringement by OpenAI and by the user.
- Infringement cases have already started. Notably, Getty Images has brought copyright infringement proceedings against Stability AI in the UK High Court for the use of images from its library (including the reproduction of the Getty Images watermark in some of the generated images).
- Where output is going to be valuable, widely reproduced or disseminated, this latent IP infringement risk may make using it too risky.



Ensure your employees declare whether output is generated by AI so that these types of risks can be evaluated before such use is made

Training data privacy infringements

- The Italian data protection authority has temporarily banned ChatGPT in Italy because the individuals whose information was in the training data set were not given notice by OpenAI that their information was held and being used for training.
- Canadian Privacy Commissioner has launched a similar investigation.
- AI models have previously been shown to be particularly adept at re-identification of data subjects even when the source data set was supposedly de-identified. Use of ChatGPT in ways that potentially impact or involve personal data should not be undertaken in the enterprise without review of the use-case (preferably by someone with knowledge of AI-based re-identification attacks) and approval of legal after considering same.



Consider a policy that prohibits making queries about individuals via ChatGPT so that your organization isn't at risk of privacy violations.

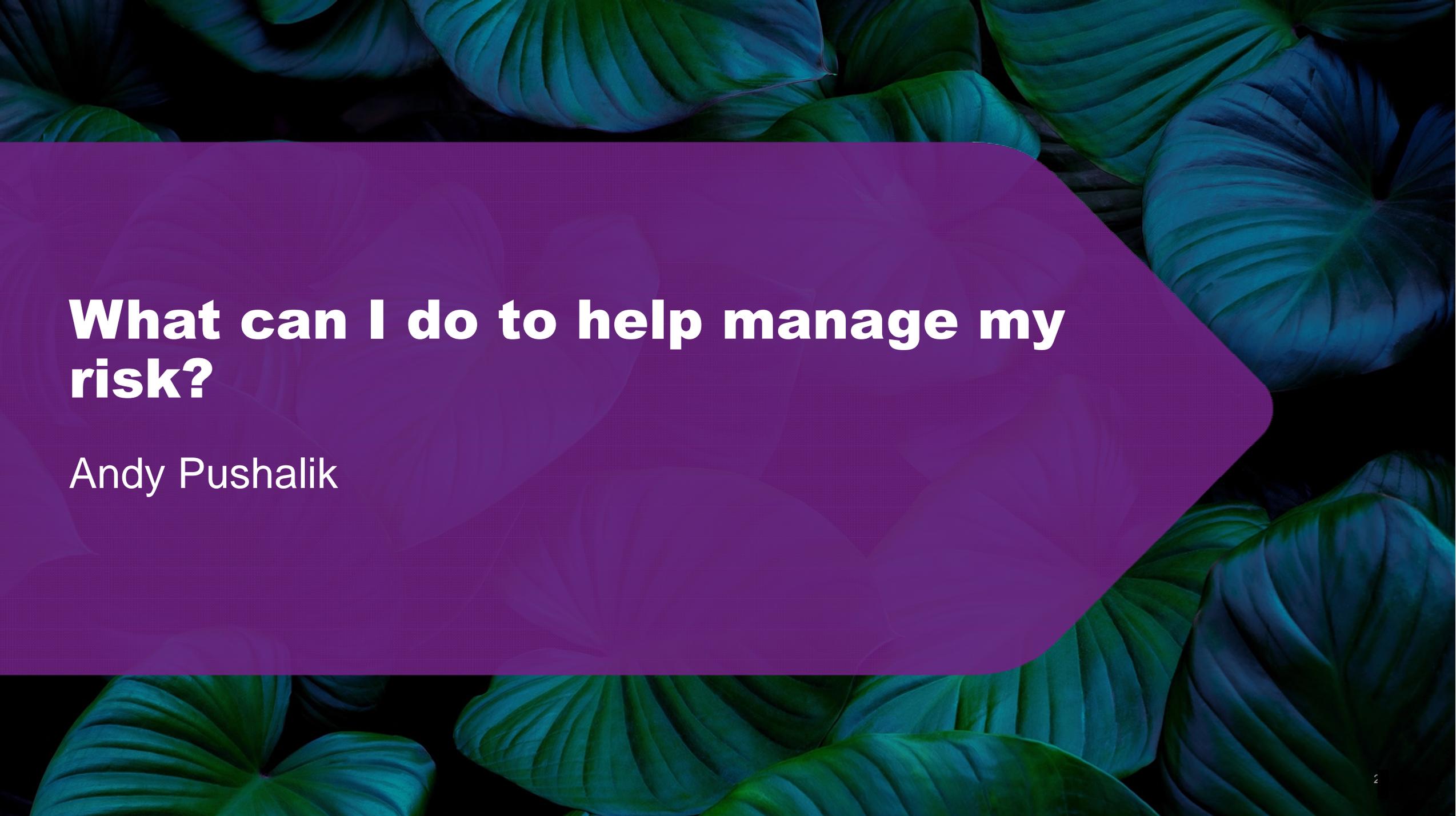


Explainability

- Although there are explanations of how ChatGPT was trained, it is a proprietary technology and not fully transparent as to the underlying processes that were used to build it, or how the model works or responds to any particular prompt.
- In Europe (and in Canada, under Bill C-27 if it passes, and Quebec's Law 25) where the application is used to make decisions about individuals with a significant impact, the user organization will need to provide an explanation of:
 - how the automated AI works;
 - the key factors it will take into account;
 - the mitigations applied against bias and inaccuracy risks; and
 - a right to contest or complain about (depending on the law) the decision with a human.



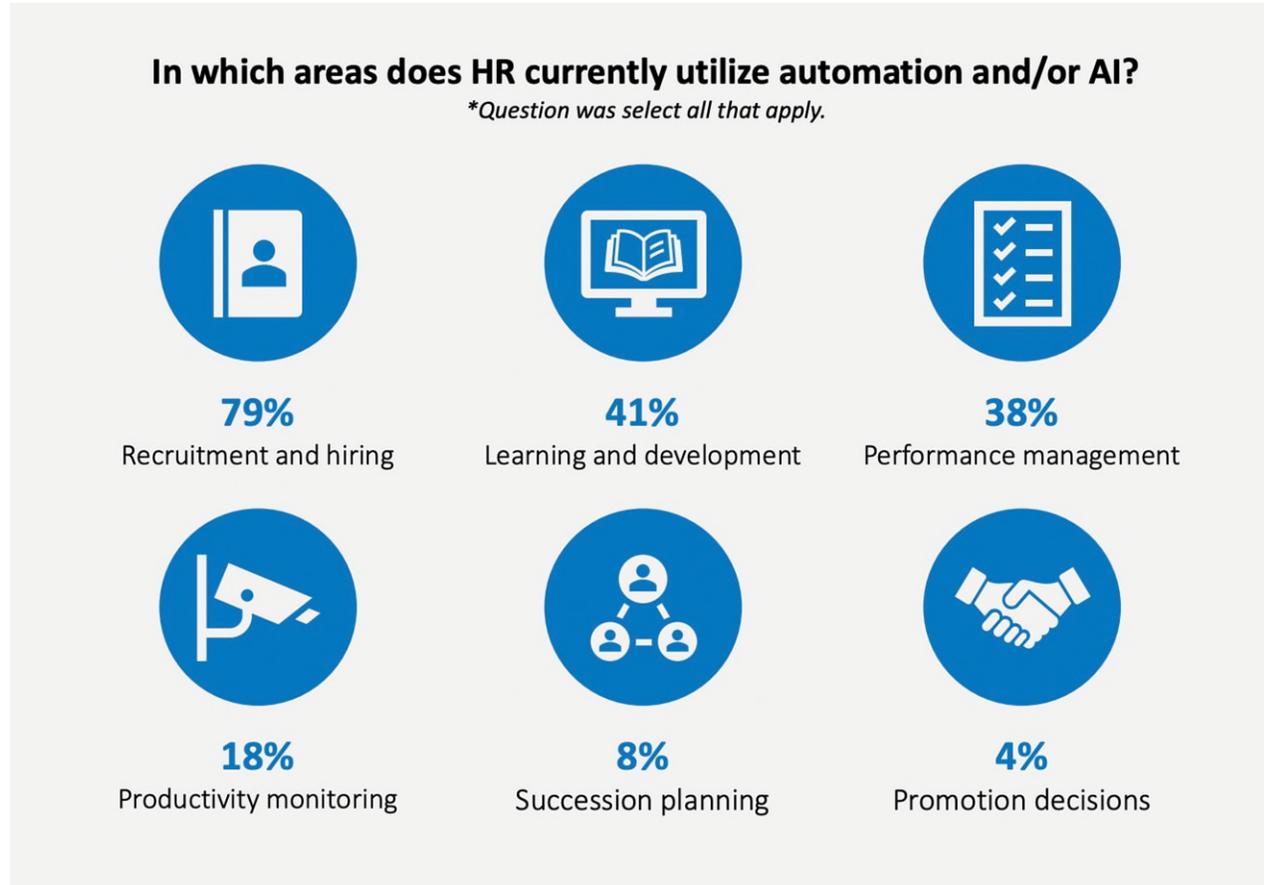
A data protection impact assessment is advisable (and may be legally required) before rolling out such an implementation. Consider having this form part of an overarching AI impact assessment program.



What can I do to help manage my risk?

Andy Pushalik

Where does HR use AI in the workplace?



Are there already employment laws about AI in the workplace?

- Currently there are no Canadian laws that specifically address use of AI in employment law decisions

US Regulation of Employers' Use of AI

- Currently no federal laws that specifically address use of AI in employment law decisions
- Some state and local laws have been passed regarding the use of AI in employment law decisions
- New York City Local Law 144
 - Regulates the use of AI-driven hiring tools
 - Law limits employers or employment agencies from using an automated employment decision tool to make an employment decision unless:
 - the tool is audited for bias annually;
 - the employer publishes a public summary of the audit; and
 - the employer provides certain notices to applicants and employees who are subject to screening by the tool.

Workplace AI Policy

- Important elements of any workplace AI policy
 - Employees should not provide any confidential, sensitive or personal information to AI
 - Employees must use AI in accordance with company policy and applicable law
 - Training
 - Highlight risks of using AI
 - Requirement for human review

Thank you

Moderator



Andy Pushalik
Partner, Toronto
D +1 416 862 3468
andy.pushalik@dentons.com



Kirsten Thompson
Partner, Toronto
D +1 416 863 4362
kirsten.thompson@dentons.com



Luca Lucarini
Associate, Toronto
D +1 416 863 4735
luca.lucarini@dentons.com