



PARTNERSHIP ON AI

RESPONSIBLE
PRACTICES FOR
SYNTHETIC MEDIA
CASE STUDY

How CBC News decided against using AI to conceal a news source's identity

 **CBCNEWS**



This is CBC News's Case Submission as a
Supporter of PAI's Synthetic Media Framework.
[Learn more about the Framework](#)

1 Organizational Background

A contextual introduction to the case study.

CBC NEWS'S RESPONSE

CBC News is a publicly funded news organization in Canada. CBC News covers general-interest news and current affairs for local, national, and international audiences on video, audio, and text platforms. CBC News operates in English and is the counterpart to the French Radio-Canada News service.

CBC News employs roughly 2,000 people and is the largest journalistic organization in Canada. More information can be found [here](#). As a publicly funded organization, it is held accountable to very strict standards around accuracy, impartiality, fairness, balance, and integrity. Users rely on CBC News because they trust it to adhere to these standards.

The challenge described in this case study is an example of CBC News applying its current [Journalistic Standards and Practices](#) (JSP) to AI-related situations that did not exist when those standards were written. Our

objective is to determine whether our guidelines need to be updated to more clearly address such situations.

While CBC News views itself as a **Creator** and **Distributor** of media, it does not view itself as a **Creator** and **Distributor** of synthetic media as described in Partnership on AI's (PAI) [Synthetic Media Framework](#), except for occasional reporting on stories about synthetic media, where such content might be shown briefly as part of the news itself. Even then, CBC News has clearly committed not to use or present AI-generated content to audiences without full disclosure. No surprises: audiences have to be made aware of any AI-generated content before they listen, view, or read it. However, as described in the case below, CBC News briefly considered the implications of creating synthetic media for storytelling purposes, and then distributing it – but did not move forward for the reasons outlined here.

2 Challenge

Elaborate on the challenge being addressed in the case study, i.e. the issue to which your organization is applying the Framework.

CBC NEWS'S RESPONSE

Note: Some details of this case have been altered for privacy and business reasons.

As a large news organization, CBC News has rules and standards that pertain to almost every topic. However, like other organizations, we are still working to create standardized policies around the use of synthetic media in our reporting. This case pertains to efforts by one of our local news teams to enhance the balance between source protection and engaging storytelling.

Our team of reporters and producers was preparing a video report about victims of online romance fraud, including a teacher in his early fifties who had been scammed and lost tens of thousands of dollars but had not told his family and friends about his ordeal. He would only agree to speak to us on the record if we kept his identity concealed from the audience.

We agreed to move forward in accordance with the provisions described in our [JSP](#), which require us to identify all our sources and protagonists with their full name (or the name they're publicly known by), except in rare circumstances where someone's safety or livelihood might be at risk. We determined that this was one of those rare cases.

The potential use of an AI-generated persona raised two issues which were covered by PAI's Framework:

- It was determined that even with public labels conveying that content is synthetic, there was not sufficiently broad public understanding of the technology to prevent some audience members from being misled or misinterpreting what they saw.
- The workings of the proposed synthetic media technology were not well enough understood to allow

the team to be assured that the privacy of the subject being used to generate the synthetic persona would be properly protected.

The case does not contribute to the harms listed in [Appendix B](#) of the PAI Framework, but could be considered

as a slightly more detailed version of “Disinformation about an individual, group, or organization.” The harm might not be to the subject of the story, but rather to the accuracy and fairness of the story itself, and, therefore, to the audience.

3 Objective

Describe what your organization is attempting to accomplish by addressing this challenge and/or furthering the opportunities.

CBC NEWS’S RESPONSE

Our traditional methods to safeguard a confidential source’s identity consist of image blurring and basic voice alteration, such as audio speeding up or slowing down. Inspired by recent publicity around generative AI, our producers began investigating updated methods to hide a person’s features and found a software that fit the bill: it would turn footage of our source into a mirror persona of a virtual person of a similar age and with comparable features — but clearly different from our source, and with an alternate voice, making him unrecognizable.

The main draw for the production team was that the virtual character would display the same tone and convey the same emotions as the real source — a rich storytelling feature that is usually lost with more typical face and voice alteration techniques. This substitution of images would require us to inform the audience of the technical manipulation. That is consistent with the direction from PAI’s Framework — to directly disclose the use of synthetically created media.

4 Framework Scope and Application

Identify which Framework principle was used to help address the challenge/opportunity, how it was chosen and implemented, and describe how it was applied.

CBC NEWS’S RESPONSE

The principles applied in this case are entirely contained in CBC’s [JSP](#), a publicly available set of guidelines that define CBC News journalism.

These comprehensive journalistic standards extend well beyond the scope of PAI’s Framework. In this case, the existing standards were more constraining than the Framework. For example, our standards indicate that “Once we have undertaken to protect a source, we ensure no details that could lead to identification are revealed.” As discussed in the next section, not being able to guarantee source protection using that technology contributed to CBC News deciding not to move forward with the synthetic content.

It is difficult to answer the other questions here given that the model **didn’t** make it to air (more below). With more experience with the technology, and with growing public understanding of the nature of synthetically generated media, future approaches to similar cases may lead to a different result.

Relevant portions of CBC’s [JSP](#) include the following, under “Production”:

“We are clear and open about the production methods we use, so the audience can put our images, sound and statements in their proper context. We advise the audience of the use of certain techniques, for example

the reenactment of a scene, the use of archival material in scenes of current events, or the use of clandestine methods.”

“A reenactment of an event must match the reality as closely as possible. When a reenactment is necessary for a proper understanding of the subject, we take care to be factually accurate, using transcriptions, minutes, or official documents. We may use a transcript word for word or set the reenactment in the location where the actual scene occurred. To eliminate any risk that the audience will confuse the reenactment with the reality, we will ensure that the audience can clearly identify the reenacted scenes.

“Other methods of illustrating a subject may attempt to describe a situation in general terms without pretending to be a precisely accurate rendering of reality. Such methods can be used, subject to certain conditions.

“A simulated scene aims to evoke or give an impression of an event, its protagonists, their actions, and the place where the event occurred. A simulated scene is produced and presented in a way that makes clear it is an evocation rather than a precise depiction of reality. If a risk of confusion remains, we advise the audience that the scene is simulated and not real.”

5 Obstacles

Elaborate on any internal or external obstacles intrinsic to the Framework that were overcome.

CBC NEWS'S RESPONSE

CBC News standards leaders decided to reject the project for the following reasons:

1. The lack of transparency behind the software company's usage of private data. In the case of this specific story, it is unlikely that anyone would have spent money and effort to try and obtain the original video from the company and reveal the source's identity, but this is definitely a concern that may emerge in other files, with more controversial or legal-related outcomes.
2. Our lack of knowledge around the technology itself and the level of randomness in the image and voice generation process. What are the odds of the virtual image looking like a specific real person out in the world? What combination of artificial features might reveal a real combination of features that would in turn point to the person's real identity?
3. Our general reluctance to create fake identities. We rarely use pseudonyms, since a name, a face, a voice, all carry human characteristics that reach the audience's conscious and subconscious perception of reality. While we can be very transparent about the fact that we're showing an AI-generated character, the audience will still be subject to a subconscious impression of a personality that we cannot control and that doesn't necessarily match reality. This is one area where future iterations of the Framework may be helpful in providing

guidance — how should Creators take into account user perception and possible accompanying bias when being transparent with their synthetic content?

Decision making on whether to consider implementing this technology came down to the following tradeoffs:

PROS (TRUTH-UPHOLDING):

- Preserve tone/prosody
- Convey original emotions of the source
- Labels could help mitigate risk of misinterpretation by explaining that the content is AI-generated to audiences transparently

CONS (TRUTH-DECAYING):

- Lack of broader AI literacy might lead to confusion, in contrast to more ubiquitous methods like face blurring
- Novel facial features might prompt inaccurate personality judgments

Tension emerged between synthetic content allowing for a more truthful rendition of a testimony but, on the other hand, introducing a degree of “falsehood” that might mislead the viewer. CBC's commitment to accuracy and public trust means that it cannot compromise on even a small possibility of misleading the audience, and the decision was made not to proceed.

The fact that we did not have readily available data to inform our thinking on the possibility of misleading audiences emphasizes the need for case study

submissions such as this one and others from the Framework launch cohort. We hope they highlight the need for innovation and research into specific use cases of the rapidly developing synthetic media space.

A notable paradox in a case with such intricate implications is that the “no” decision was actually quick and easy to make. This speaks to the intentionally slow and risk-averse pace at which an organization like CBC News updates its guidelines – with a decision-making structure designed to limit risk taking.

Production teams across the network are free to make creative editorial decisions and enhance their storytelling as long as they follow the strict rules laid out in the [JSP](#). In other terms, thinking outside of the box is encouraged, but it can’t affect accuracy, fairness, or the other core principles of CBC journalism. The audience doesn’t come to us because they like our politics – we don’t have any. They come to us because they trust us to follow those rules.

In that spirit, exceptions to CBC News standards are only allowed in very specific circumstances defined in the core JSP document itself (such as the use of a hidden camera, for instance) and require approval by the Standards Office on behalf of the Editor-in-Chief. There are no exceptions to the core principles.

The case of the proposed avatar described here touched a nerve: it would have implied showing an image that was not accurate by any reasonable definition. Accuracy means a reflection of reality. One could imagine that in a distant future the audience might come to approach it differently, perhaps with synthetic media becoming a widespread feature of media content with its own codes. That is not the case now and CBC News does not have a mandate to experiment with its principles. While we are keenly interested in observing the evolution of generative AI in public perception, we remain committed to a rigorous definition of accuracy.

6 Benefits

Identify the opportunities created for your organization by utilizing the Framework to address the challenge.

CBC NEWS’S RESPONSE

We are not closed to exploring more possibilities around the potential use of generative AI for source protection. While teams directly involved in TV production were showing enthusiasm to test the technology, the CBC’s Standards office makes the final call on all journalistic practices, including this one – it declined to approve the use of the technology to create a digital avatar of the romance fraud victim.

Subsequent to the publishing of the PAI Framework and after the case described here had unfolded, CBC News’ Editor In Chief provided [public guidance on the use of AI tools](#) by CBC news. Those guidelines are based on CBC’s long standing principles of accuracy and transparency and make PAI’s Framework easy to implement in its current

form. Two points in particular are relevant here:

- “We will not use or present AI-generated content to audiences without full disclosure. No surprises: audiences will be made aware of any AI-generated content **before** they listen, view, or read it.”
- “We will not use AI to generate voices or a new likeness for confidential sources whose identity we’re trying to protect. Instead, we will continue practices well understood by audiences, such as voice modulation, image blurring, and silhouette. In all cases, we are transparent and clear with audiences about how we’ve altered original content.”

7 Conclusion/Key Takeaways

A description of how implementing the Framework ended for your organization, including any lessons learned.

CBC NEWS'S RESPONSE

While there is significant potential for the use of generative AI in investigative journalism, CBC News did not feel moving quickly to roll out an anonymous digital avatar as a tool for balancing source protection with engaging storytelling was worth the risk of potentially misleading viewers.

The proposed tactic for leveraging AI to preserve source anonymity was not approved, but we committed to including similar questions in upcoming reviews of our standards, which will imply more research and consultation around the companies providing AI technology, their use of data, and more visibility into how the technology functions. We should be able to convey this information to the audience, in our promise for transparency, and getting there implies more control and clarity around the technological processes that we would use.

Reviewing CBC's journalistic standards is a formal process that only takes place every five or six years through the work of various consultation committees and is subject to approval by CBC's Board of Directors. The

last time that process occurred [was in 2018](#). Any future review of CBC standards would likely study the relevance of adding guidelines around synthetic media, both around gathering and producing news content.

One takeaway from this exercise is that disclosure, whether direct or indirect, is only a part of the solution to communicating the presence and nature of synthetic content. Equally important is a well established understanding of the implications of the presence of synthetic media by the general population. The sophistication of the synthetic generation technology currently far outstrips the comprehension of many people who make up the audience we serve – although it would be useful to gather data on how different categories of people in the audience perceive the use of synthetic media. Until there is wider understanding of what is technically possible, and greater media literacy, any labeling of synthetic content risks increased confusion and possible inadvertent deception to some segments of the audience. The importance of increased media literacy efforts should be highlighted in future releases of PAI's Framework.