



PARTNERSHIP ON AI

RESPONSIBLE  
PRACTICES FOR  
SYNTHETIC MEDIA  
CASE STUDY

# How the BBC used face swapping to anonymize interviewees

**BBC**  
**R&D**



This is Adobe'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.

---

## BBC'S RESPONSE

The BBC is a global media organization producing news, sports, children's entertainment, and factual and educational content across radio, TV, and online outlets (including one of the world's most-visited news websites). It has very high engagement across the United Kingdom (UK) and, through output such as the BBC World Service, across the globe.

Some of the BBC's content is brought in from external providers (such as the independent producer referenced in this case), but much of it is created in-house.

The BBC has a Research and Development (R&D) team which builds, adapts, adopts, and focuses on technology in partnership with wider BBC teams — it is R&D which has driven our work with Partnership on AI's (PAI) [Synthetic Media Framework](#). R&D teams also liaise with other teams across the BBC on responsible AI policies and practices.

All BBC content and services are governed by [Editorial Policy Guidelines](#), a formal document shared with all staff and the public. The guidelines spring from the BBC's Royal Charter which sets out the BBC's public purposes:

1. To provide impartial news and information to help people understand and engage with the world around them.
2. To support learning for people of all ages.
3. To show the most creative, highest quality, and distinctive output and services.
4. To reflect, represent, and serve the diverse communities of all of the United Kingdom's nations and regions and, in doing so, support the creative economy across the United Kingdom.
5. To reflect the United Kingdom, its culture, and values to the world.

The BBC also operates according to its editorial values of trust, fairness, accuracy, and editorial integrity, balancing the need for impartiality, respect for privacy, and care against potential for harm and offense.

The use case we have chosen for this document is the BBC's use of face swap avatars to preserve privacy. In this case, per PAI's Framework, the BBC was the **Creator** (though working with an independent company through commission) and **Distributor** of a set of synthetic faces designed to protect the identity of interviewees in a documentary in which anonymity was required.

Our objectives were to make the best possible program while preserving the anonymity of key contributors. Anonymizing is an age-old challenge as traditional techniques are clumsy, and can adversely affect the relationship between anonymized subjects and the audience.

The program was produced by the independent production company, Daisybeck Studios. The synthetic media work was conducted by graphics house Glassworks. There is a brief explanation of the process in [this article](#).

*Our objectives were to make the best possible program while preserving the anonymity of key contributors.*

## 2 Challenge

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

---

### BBC'S RESPONSE

In 2022, the BBC commissioned a program on Alcoholics Anonymous (AA) called **I'm An Alcoholic: Inside Recovery**. It aired on December 7th, 2022 with the following description:

*A look inside the recovery group Alcoholics Anonymous, with the first ever access to an AA meeting. How does an organization rooted in 1930s American evangelicalism still work in the modern world?*

The program is on the BBC's iPlayer [here](#). (Note: This link may not work outside the UK due to geographic restrictions and rights constraints.)

A key challenge for the BBC when embarking on this storytelling project was the preservation of interviewee anonymity. AA has a tradition that those who take part do not disclose their identity outside of meetings. Since recovering alcoholics were the main interviewees, delivering anonymity posed a significant challenge. Traditional techniques such as pixelation, blurring, or shots of non-facial body parts (hands, backs of heads, etc.) have always been problematic, not least because the most "effective" methods are also the most obscuring.

The BBC's Editorial Guidelines provide guidance on how to deal with source anonymity in sensitive contexts, stating:

"Footage and photos of people who wish to remain anonymous can be used provided the person's face and hair are thoroughly blurred or obscured. Blurring is preferable to pixelation as the latter can be reversed."

We judged that using these techniques on a program in which so many interviewees required anonymity would have greatly affected its impact on audiences. There is also the risk that the very act of anonymizing can, to some, suggest "perpetrator" and the program team were keen to avoid this notion. Face swap technology helped to give the program much more visual appeal, as well as to avoid any potentially negative framing.

Commissioning Editor Daisy Scalchi noted that offering access to the interviewees' testimony via synthetic media technology helped bring out the emotions of the first-hand testimony being shared by contributors, making it as impactful as possible. However, in order to ensure that audiences were at no point misled about the techniques being used, voiceover and on-screen text indicated every time an interviewee appeared with face masking. There is more detail about this in the subsequent sections.

The absence of overt masking techniques like blurring or pixelation allows the interviewees to be better integrated into the rest of the program. Synthetic media therefore has a significant benefit related to the creative impact of the interviews.

None of the harms of the Framework's [Appendix B](#) (such as extortion, fraud, the deliberate influencing of public opinion etc.) were relevant. This was a question of personal anonymity for individuals in line with the BBC's Editorial Guidelines.

### 3 Objective

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

---

#### BBC'S RESPONSE

As described above, we sought to improve the safety of the interviewees and maintain the principles applied in meetings of Alcoholics Anonymous. Daisybeck Studios and the team making this film were rigorous in their attention to detail and duty of care to all participants to ensure everyone understood the processes being used.

The aim was to use newly available synthetic media/ deepfake technology that could “face swap” documentary interviewees in order to preserve anonymity, while avoiding

techniques which would detrimentally affect the audience experience watching the interviewees. We wanted to make sure the use of “face swapping” would not cause our viewers to perceive the subjects negatively and would allow the interviewees to share their story without being subject to preconceived notions traditionally associated with anonymizing techniques such as face blurring.

This objective remained the same throughout the process.

### 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.

---

#### BBC'S RESPONSE

Prior to beginning shooting the documentary, we first ensured we had the consent to be “face swapped” from all the subjects to be interviewed. Their experience is not the main topic of this case study, rather, we are concerned here with the ways audiences perceived what was presented.

In considering the audience experience, we focused on those parts of the Framework concerned with disclosure and transparency when carrying out this project. As the **Creators** of content that included AI-generated videos, our responsibility under the Framework was to:

“Disclose when the media you have created or introduced includes synthetic elements especially when failure to know about synthesis changes the way the content is perceived.”

We employed two direct disclosure mechanisms at different times throughout the program – one auditory and the other visual. At the beginning of the program, the narrator stated:

“All the recovering alcoholics we interviewed have had their faces changed by a cutting-edge computer program. This deepfake technology allows us to capture these emotional stories while maintaining

the anonymity that AA requires and those sharing their stories feel comfortable with.”

Later on in the program, one of the interviewees underscored our assessment of the utility of using generative AI technology to bring out the positive message of being in recovery:

“It’s to protect us because there can be a lack of understanding of alcoholism, and it can be judged. The important thing is stories and recovery and the road to freedom from alcoholism; it doesn’t matter that it’s not my face.”

The second direct disclosure mechanism occurred whenever interviewees appeared on the program. When they did, we used an on-screen caption that said:

“Faces have been digitally altered.”

The production team observed that it was important to them to very clearly signpost “wherever someone’s face had been deepfaked” and this was in line with the BBC editorial guidelines described above.

While we have been able to assess the impact of the synthetic media effort on the subjects who were face

swapped, this is a very new development and we have not assessed audience attitudes towards the use of synthetic media for face swapping, or their understanding of the direct disclosures we chose to leverage. As this technology becomes more widely used we will conduct follow up work on this.

In terms of transparency, the disclosure offered was felt to be enough. Going deeper with audiences on the potential limitations of this kind of technology (as laid out in section 7) would have been inappropriate and irrelevant given the directness and clarity of the disclosure and how it was woven into the program narrative.

## 5 Obstacles

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

---

### BBC'S RESPONSE

Most of the challenges we faced were technical or workflow based. For example, key interviews had to be done earlier than would normally be the case in a documentary of this nature (given their importance to the narrative and need to respond to any emerging issues). This workflow step was frontloaded to ensure that there was time for the synthetic media to be created. Further, all key interviews had to be filmed on the same day because the conditions needed to be identical for effective face swaps. Again, this is not normal practice.

One technical challenge we faced was that one of the interviewees had a beard. This made his “face swap” more complicated for the production team, although it was ultimately overcome.

Although the face swap work added an extra cost to the production, disclosure did not add any costs, nor did it elongate the production process significantly. This would be an important consideration for future productions, where budget pressures are always intense.

One challenge did emerge due to the nature of the Framework itself. Our application of the principles felt well-aligned with our editorial policy and instincts for this production, but because the BBC was both a **Creator** (via a commission) and **Distributor** of synthetic content, we had to consider both aspects because they are treated separately in the Framework. Perhaps future versions can provide more specific guidance for organizations that identify as both – in other words, are there any unique implications beyond what is written that relate to those who play dual roles? Is there increased responsibility on the creation front if **Creators** are also going to be distributing material incorporating synthetic media?

*Although the face swap work added an extra cost to the production, disclosure did not add any costs, nor did it elongate the production process significantly.*

## 6 Benefits

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

---

### BBC'S RESPONSE

The PAI Framework has given us a mechanism to think about how synthetic material should be addressed in our output. We can combine these insights with the “real life” experience of the program to help us build expertise and policy around synthetic media.

Determining how to provide transparency when content has been edited or augmented, whether with AI or not, is not new within the BBC, particularly when it comes to areas like natural history programs where there can be questions around authenticity and editing of material captured in

the wild. The Framework has provided a valuable reference point for our discussions across all editorial and audience-facing areas of the BBC. This has been made even more timely by the emergence of easy to use generative AI tools to create synthetic images and video.

An additional benefit is exposing BBC teams, independent producers, and audiences to the promise and nature of this kind of technology for storytelling which, in itself, usefully raises awareness for the technology more broadly in society.

## 7 Conclusion/Key Takeaways

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

---

### BBC'S RESPONSE

We consider our use of generative AI to create synthetic media in order to “face swap” interviewees to be a success, and also view it as a meaningful implementation of the Framework. Individuals for whom privacy was crucial were able to appear in a program, convey their thoughts, and display their emotions without the use of traditional anonymizing methods, and they did so in a way that provided users with engaging content.

In the context of the Framework, there were a number of relevant aspects relating to this production:

- As far as disclosure is concerned, it felt appropriate to tell the audience in two ways during the program about the synthetic nature of content: at the beginning, when the narrator informed audiences that interviewees were having their faces “changed,” and then again with a caption that appeared whenever a “face swapped” interviewee appeared stating that the faces has been digitally altered. Having both auditory and visual disclosure was a comprehensive way of providing transparency in an audio-visual medium.
- We may, in the future, offer this kind of technical work in-house (in fact, since this program aired, we

have provided in-house generated face swap support for another production). This could potentially offer additional benefits, such as an additional layer of protection to the actual likeness of subjects as their data would not be shared with a third party organization outside the production company/BBC.

- This kind of disclosure is often used to signal artifice – for example, a reconstruction in a crime documentary. However, since these techniques are new, thorough disclosure felt especially important. An open question is whether, over time, disclosure may be done differently – for example, simply with an on-screen caption at the relevant point in the program.
- Another open question is how organizations can identify what “proper” or “responsible” use cases are for face-swapping. This program was an ideal deployment, but are there risks in other instances of anonymity provision that might prove problematic, where the ease of providing an “alternative” face means that we grant anonymity to interviewees who should be pressed harder to speak in public.

- There are other issues too. We worked on the assumption that face swapping is a less intrusive form of anonymization, but it may be worth commissioning further work to clarify this. For example, might a generated face make us feel differently about an interviewee with respect to their “real” face – we know that humans have responses to perceived “attractiveness” and other factors. What is the impact of this on the way a story is put across? Would it be appropriate to change characteristics such as race in a swap? Could a generated face look too “similar” to a real person – either the one in focus or another human, and what are the risks in that? All these issues and more need further consideration and research.
- This work combined with the PAI Framework leads us towards providing clear guidance to teams on the use of synthetic media so that together we can address some of the issues outlined above.

Finally, we want readers of this case study to reflect on a comment made by Commissioning Editor Daisy Scalchi who says it’s important to interrogate any future proposals to use synthetic media to disguise identity on screen. She says, “there has to be a real purpose behind it, and it has to bring something authentic to the storytelling.” This case provides an excellent example of using these emerging technologies to deliver content with journalistic integrity, and we hope it will help guide others making decisions about when there is “real purpose” behind their use in storytelling.