Challenging Algorithmic Bias through Video Game Design

The idea for this game jam comes from a project I'm working on titled 'Challenging Algorithmic Racism through Digital Cultures in Brazil' (for more information, see https://creativetechnologies.blogs.bristol.ac.uk/2023/11/19/contesting-algorithmic-racism-through-digital-cultures-in-brazil/). As part of the project, I'm working with Brazil-based partners in the creative industries and non-profit sector to develop a video game designed to raise awareness of the embedding of racial biases in software systems. So, by way of a brief introduction, I thought I'd describe the project and what we mean by algorithmic bias in that context before thinking a bit about what the medium of video games can bring to this discussion.

The project interrogates how cultural practices, including video game design, can be used to challenge the ways in which, despite their sheen of neutrality, new technologies often reproduce existing social biases and power hierarchies. I'm particularly interested in the intersections between the creative industries and activism, the roles that culture plays in forming and unsettling understandings of the social impacts of technological systems.

Although algorithmic racism is a global concern, it has become the subject of particularly urgent debate in Brazil. A combination of high rates of smartphone ownership and social media usage with low levels of digital literacy, coupled with the normalisation of racist discourse by the far-right Bolsonaro movement, has made it especially visible in the country. At the same time, cultures of opposition to algorithmic racism have been particularly quick to mobilise in Brazil due to two main factors: 1) the prominence of media activism that is the legacy of government-supported digital inclusion initiatives of the 2000s; and 2) the explosive confluence of both national anti-racism movements (including the Movimento Negro Unificado) and transnational campaigns such as BlackLivesMatter.

A researcher from Brazil called Tarcízio Silva has constructed a 'timeline' of algorithmic racism in Brazil (https://tarciziosilva.com.br/blog/destaques/posts/algorithmic-racism-timeline/), which includes the following examples: the racist assumptions of search engines employed by photographic databases such as GettyImages, which was exposed by Brazilian NGO Desabafo Social in 2017 when they showed that the results for the search term 'family' were images of exclusively white families; the blocking of a cartoon of a black child in a favela in 2019 because Instagram's image-recognition algorithms mistook its skateboard for a gun; and the fact that, by November 2019, more than 90% of prisoners caught by facial recognition technologies in Brazil were Black.

There are, of course, plenty of examples closer to home. For instance, more than 100 schools in Bristol are currently using an app called Think Family Education which gives them easy access to information about their students' and families' contacts with police, child protection and welfare services. It also provides an analysis of which children could be at risk of exposure to criminality. Campaigners have been quick to point out that the app risks producing more discrimination against pupils from minority or working-class backgrounds. As Griff Ferris of the charity Fair Trials put it in an article in the Guardian published in September (https://www.theguardian.com/education/2023/sep/21/calls-to-shut-down-bristol-schools-use-of-think-family-education-app-pupils-and-families): "Schoolchildren

should not be monitored, profiled and criminalised by secretive police databases. Surveillance is not safeguarding. Systems like this uphold existing discrimination against children and families from minoritised ethnic and more deprived backgrounds."

To come back to our project, we're currently in the development stage of our game, which is basically an extended (and more complicated) version of a game jam. We're running a consultation or 'Co-Creation Lab' with representatives of non-profit organisations working on digital literacy or digital inclusion initiatives across the country. This will culminate in the design of a blueprint for the game, which will be developed by the Aoca Game Studio (https://aocagamelab.games/), producers of the excellent historical adventure game *Árida: Backland's Awakening*. This is building on our experience of having worked together last year to develop the demo of 'Future Call', a narrative game that raises awareness of the dangers of misinformation on social media (https://aocagamelab.games/futurecall/).

So, why are we so convinced that video games are an appropriate medium for exploring these issues? An artistic project called 'Genetic Automata', which is currently exhibiting at the Wellcome Collection in London, is useful for thinking it through (https://wellcomecollection.org/exhibitions/ZAW0PxQAACcG-pX8). The exhibition comprises a series of four short films by the artists Larry Achiampong and David Blandy, which, to quote them, 'explore race and identity in an age of avatars, videogames and DNA ancestry.' The films use video games as a metaphor for how social reality continues to be mediated by the categorising logics that date back to the racist pseudo-sciences of the 19th century including the work of figures such as Carl Linnaeus, whose taxonomies both drew on and reinforced a belief in racial hierarchies. And this influence at the level of conceptual structures is intensifying as our lives become increasingly mediated by digital systems that employ this categorising logic. As the voice-over puts it: 'The grid holds us all in place. They define the rules of play – then show how we lose, switching all the variables in their favour, playing gods.'

The first film, titled 'A terrible fiction' (2019) (https://wellcomecollection.org/works/wys2bdym), points out how the logics of racial categorisation are mirrored in the infrastructures of many video games: the processes of avatar selection, including the use of default options; and the reduction of identity to clear-cut categories that are often ranked and pitted against each other in struggles for survival or supremacy. The latest instalment of 'Genetic Automata', '_GOD_MODE_' (2021) (https://wellcomecollection.org/works/bbbwbh85), is a little more hopeful about the promise of video games to expose and undermine these processes. It asks: 'can't we shapeshift too – work together to undermine these fantasies – to create a new game engine – to break the webs of myth held in zeroes and ones?'

In his 2007 book *Gamer Theory*, McKenzie Wark points out that while playing a video game we're caught in a tension between the level of representation, symbols and identity on the one hand and the algorithmic system that makes the game work on the other. Gameplay entails an attempt to figure out the connection between the two – which objects and actions have meaning or value within the algorithmic system of the game. So, the gamer – or at least Wark's idealised figure of the 'gamer as theorist' – is in the ideal position to interrogate the increasingly intimate interrelations between identities and code in what passes for the real world: the way that the social order is increasingly regulated by algorithmic systems.

I'm looking forward to seeing what everybody comes up with!