



THE JOCELYN HAY VOICE OF THE LISTENER & VIEWER LECTURE 2024

Thursday 8th February 2024

TRANSCRIPT

BARONESS BEEBAN KIDRON OBE

I feel very fortunate to be invited to give the Jocelyn Hay Lecture shortly after the ITV drama *Mr Bates vs the Post Office* hit our screens. *Mr Bates* was a drama that defies the drift toward algorithmically determined commissioning. It also pointed to something very wrong at the heart of our democracy. It took only four hours of prime-time telly to overturn 25 years of political and corporate negligence and malevolence. So it is a privilege to be speaking to an audience who have broadcast TV in their DNA.

TV, film, theatre, music, visual arts, architecture, and design are all clues to how we see ourselves both as individuals and in relation to each other. TV programmes that bring us together, whether the nightly news, *Mr Bates*, *Strictly* or – *Gogglebox* have a unique function in a world that has been ‘personalised’ or more accurately ‘atomised’ to an unhelpful degree – because they create a shared reality.

We live in a time of division, culture wars and disputed realities algorithmically pushed to highlight our difference or to mirror our certainties - until doubt and reason get rubbed out. The danger of these divisions—symbolised by our fragmentation north and south, Brexit or not, urban or rural, nationalism or union, present or future, foreigner or friend — is that they are all binaries that threaten to overwhelm our common interests.

But we also have a new game in town – the synthetic.

ChatGPT had an explosive landing, from one million users within days of its launch to 100 million users within two months. Set up as a non-profit in 2015 Open AI was as I typed these words, worth \$86 billion and reported to be fundraising at a \$100billion value. Its impact on the public consciousness AND the column inches in the press immeasurable. But perhaps most important for those who seek to develop technology that reflects societal values and

common interests was the boardroom battle in which the chief executive Sam Altman was fired and rehired in less than a week. A knockout blow to those who cautioned caution – it was not responsible innovation but commercialisation that won the opening `bout`.

The excitement of being able - with a few quick prompts – to have President Biden’s State of the Union Speech sung by the cast of *Hamilton*, or an ego prompt to find out what Baroness Kidron might say is wrong with the government’s approach to child online safety was palpable. It caught our imagination and suddenly we felt as if the computer knew something more than us – even as the problems inherent in the sophisticated ‘best guess’ game of generative AI were beginning to emerge.

And while commentators found increasingly humorous ways to use the new-found power to mimic human behaviour, a considerable and very reasonable anxiety began to emerge from the creative community. Their work, so fundamental to the ‘near magic’ outputs of Large Language Models – was being scrapped, sucked in, and regurgitated with no acknowledgement of its original purpose, value, or form. A life-time’s work whether text, image or moving image, whether creative or factual, whether academic or entertainment could, be surfaced and hidden simultaneously. Surfaced by emerging as a near likeness without acknowledgement of its creator. Hidden as an infinitesimal part of something new, that could not have been manifest without it, but no longer has the same purpose. To many it felt like a brilliant close hand card trick.

Along with the wave of awe and wonder, ChatGPT brought to town the tech bros warning that the monster they had unleashed would create misinformation at such scale that to **know** anything would be impossible, that most would be left **without** a job, and that robot weapons would **decide** who was the enemy. In short, that the scale of decision making that the new models allowed was on course to bring down society-as-we-know-it. They were simultaneously the authors and Cassandra’s of existential threat.

In one scene of *Mr Bates vs the Post Office*, an **unhelpful** voice manning the Horizon **helpline** instructs the distraught postmistress to click on the screen and doubles her putative debt from a little over two thousand pounds to more than four thousand. Nothing material had changed, no cash had been received or taken, no stamps given or received, no benefits demanded or dispersed. It was just the computer doing its thing.

As we have seen, and kudos to those who commissioned, produced, wrote, directed and played any part on or off screen, *Mr Bates* created a storm of empathy and outrage. In *Mr Bates*, we saw corporate entities enabled by governments asleep at the wheel, outsourcing the bugs of a technological system onto the least powerful - in this case the sub postmasters – and it felt horribly familiar.

Missing a precious GP appointment because the parking App doesn't work. Hanging on for hours for the HMRC to find that your question must be dealt with online but online there is no answer to your question. Trying to contact an airline from an airport when your flight is cancelled, only to realise that you are texting with a bot that doesn't do multiple-destination trips. Or in my case - a text every morning from a medical device that had been surgically removed from my chest in 2021. There was no one to ring, no one to contact, just the daily reminder of a near death experience in the form of distracting alert. Until very recently when someone suggested I text the word 'STOP'. Not subtle but – unbelievably – after more than a thousand messages – it responded, “you have opted out and you will receive no further messages – Everbridge Alerts”. That was the first time I had heard of Everbridge.

I realise I sound like a tech detractor. Not so. I am an early adopter in so much as I had a webcam so early, that having set it up, I suddenly realised I had no one to video call. And I am a tech optimist in so much as I see and believe that digital technology will play a starring role in helping us solve many of the issues society faces. Which is not to say, I don't regret – hugely regret – that technology that could contribute so much to human flourishing, is so often cannibalised by those who seek vast profit and even vaster power.

Much of the work I have done in Parliament and beyond, is about technology as it intersects with children. Specifically designing and regulating it to respect their privacy, account for their development journey and observe their rights - **in full**. In that capacity I have been involved with others in drafting and passing the General Comment 25 that sets out how children's existing rights under the United Nation's Convention of the Rights of the Child apply to the digital world, the Age Appropriate Design Code, which requires greater privacy from digital services likely to be accessed by children, and the UK's Online Safety Act which has brought in a swathe of new legal and regulatory demands that should – if implemented effectively make a significant difference to children's lived experience. And, I have contributed to legal and treaty frameworks

internationally in the knowledge that while the first billion children online are predominantly in the global north – the next billion will come from the south – particularly Africa.

My focus and that of the team at the 5Rights Foundation, is on product design. How to, and why we must, make safety, privacy, and rights by design an industry norm. This focus on up-stream, or as the lawyers prefer ex-ante design, and governance involves a broad range of strategies to ensure that children remain participants in the digital world, but with recognition of their rights, development stage and mitigations for the vulnerabilities associated with their age.

Among the issues that persistently comes up is that of Child Sex Abuse Material (CSAM).

The Creation, distribution, consumption of CSAM is illegal almost everywhere. Here in the UK CSA is covered by at least four separate laws¹ - the oldest of which is from 1978 and the newest from October last year. ***But as yet the models, or ‘plugs ins’ – sometimes referred to as LoRa’s, trained on and specialising in the creation of CSA material - are not.*** Allowing for an eye-watering rise in AI-generated Child Sexual Abuse **Material**.

The creation of child sexual abuse **material** is now unfettered by the friction of real life. A child’s photograph, scraped from a school website, social media or advert, combined with code built on pornography, can be individually created to match the bespoke wishes of a disordered sexual imagination.

Images proliferate that may be any combination of a real child in a synthetic act, a real act with a synthetic child whose physical attributes have been moulded from a cacophony of prompts – bigger bluer sadder eyes - with tears – for example. The images may be unreal action in a real environment or the exact opposite. For a shockingly small amount of money, the police see people ordering up scenarios using a picture of their neighbour’s child or children from their own family. And while some claim, as they did in the recent case of a child’s rape in the metaverse, that there is not victim. I would ask you to consider if you would like your child’s image to be used in this way?

¹ [section 1 of protection of children’s act 1978, section 160 of the criminal justice act 1988 and section 62 of the coroners and justice act 2009 and OSA 2022]

This is not a victimless crime. CSAM is being created in vast quantity, the numbers of people engaging it is rising, it is being normalised. And the police believe that rehearsing these scenarios online, emboldens perpetrators and shortens the time it takes them to take those same actions offline.

My point here is that, in this highly legislated area of digital content of child sexual abuse, we have once again allowed those that own the means of production – in this instance image creation companies - to avoid responsibility for what they enable; we have been slow to understand that code itself is not neutral; in the name of innovation we have failed to account for the social cost. We are – again - allowing tech companies to scrape our data, build powerful systems with no corresponding societal responsibility.

The reason that I am setting this miserable scenario in some detail is because it offers three important lessons in a period in which the shouts of existential threat from those who are creating the threat drown out the more orderly voices that call for technology to be deployed in a way that is focused on social benefits.

First, we have laws in many of the areas that are causing concern. So before raising the alarm about existential threats of the future it would seem prudent to look at the present and see how our existing rights and laws do apply or could be updated. I have just applied this thought to the CSAM context, but I could make the same argument about intellectual property, data protection rights, employment laws such as the collective bargaining – as the Hollywood writers have just done - consumer rights and or safety standards, or what about Human rights? children’s rights?

How would the sub-postmasters’ lives have played out if we routinely made companies who sell, supply, or deploy technical systems responsible for their impacts? **Every complex computer system has bugs.** That is not hyperbole or rhetoric – it was said to me a couple of weeks ago in a meeting with a colleague - a Professor of Computing at Oxford – who in his spare time acts a consultant to some of the biggest tech companies in the world – his specialisation: finding bugs!

Second, the language of existential threat that AI will replace humans – is something that disempowers most of us. **We feel it’s just too big, too amorphous a problem.** But ask us if we want to supercharge the creation of Child Sexual Abuse material? I would hazard a guess that the answer is ‘No’. If

it is ok to have facial recognition trained on white faces so a black head teacher, visitor, parent, child is not considered human and cannot pass a school security scanner (I have that picture the black head teacher!) – again ‘No’. Whether we think accounting systems should be designed with a backdoor by which an unknown unauthorised person can change the entries with impunity – ‘No’. You all in this room are communicators - we have language that provides for shared human values and that language gives us agency and possibility of a critique which allows legislative change and societal rules. The language of existential threat gives us none.

Nuclear, biological weapons, disease contagion, or even climate change all have the capacity to bring the world as we know it to an end. On the first two, the global community curtailed both development and spread to a degree of success that has at a minimum prevented global annihilation. **So far.** The pandemic saw the human agency at scale – as every part of the world moved to contain the virus. Perhaps climate change is simultaneously the best and worst example – in that we see a struggle for human agency over vested interests – in which the enormous equity disparities between polluters and polluted, *between the natural world that provides Co2 and human behaviour that gobbles it up*, between the short-term politicians and businesses and the longer-term interests of the young. This battle is in full swing and offers a glimpse of how it is possible to make a question so big that it creates an environment in which the immediate and practical actions that might really contain the threat are overlooked in favour of a yet unidentified silver bullet that will save us when the time comes.

And third, in this election year. And I say this as a crossbench peer, who neither has a party or a vote – because prisoners and peers do not get a vote. In this election year, we must pay particular attention to AI’s extraordinary ability to create synthetic truth. Because the story I told at some length about CSAM is as true of mis- and disinformation as it is about abuse scenarios.

And here let me tell you a short story.

I often do workshops with children and before Christmas one group of young people – about 12 children between the age of 13 and 19 - asked me to do a workshop on AI. Towards the end of our session, we had a brief discussion about the structure of Parliament, the government, the official opposition the minority parties, civil service etc – during which I explained that Keir Starmer

was ahead in the polls by 21% and could become the next Prime Minister. We discussed at some length the attributes they felt a prime minister should have. **At which point I said that I had seen a video from his private office in which he was berating his staff, cursing, and humiliating them.**

Their condemnation was fast and furious. Their vision of leadership, which had been thoughtful, sophisticated, and generous, was smashed by this new information. And they withdrew – **and this is important** – they withdrew in a direct sense in that they did not think a man who behaved like that should be Prime Minister. But **they also withdrew emotionally**, from being excited at determining what good might look like, into a feeling that things were (forgive my language) shit again. Which is when I told them about the Pope’s puffa, Biden’s fall, the London Mayor’s voice – and we had a conversation about mis and disinformation and the implications for the elections in the US, UK, and India.

In election year, or in any year, it is dangerous to have an information environment in which the synthetic passes for truth unchallenged – not only because those that are wrongfully accused will suffer - **but because those who believe will suffer**. And as someone who routinely sees synthetic material of every possible kind, including pictures of myself as a terrorist, at demonstrations, as a showgirl, as superman – I would like you to believe me when I say that you can not tell the difference. If I did not know I wasn’t there, I would not know it wasn’t true.

Which is why it is urgent to reclaim our public space - our communication space. Our narrative space. And our truth.

The language that suggests that AI is too late and too difficult for us to deal with – is part of a decades-long deliberate strategy of tech exceptionalism that has privatised the wealth of technology and outsourced the cost to society. We do have agency and we must exercise our democratic power to ensure that our common interests are served. **It is tech exceptionalism that poses an existential threat to humanity not the technology itself.**

AI is built, used and purveyed by business, governments, civil society and, as I have already pointed out, criminals. It is part of the story of the sub-postmasters, part of the convergence of media, and it is part of the choices over which we still have agency. ***Who owns the AI, who benefits, who is responsible and who gets hurt – is at this point – still in question.***

Which brings me back to why we in the UK, with our unique ecosystem of broadcast TV, with our national obsessions of house hunting and watching other people cook; with our incredible history of drama production, must continue to value, and insist on the opportunity to watch *I am Ruth*, *Small Axe*, *Peaky Blinders*, *Mr Bates vs the Post office*; to hear the nightly news from sources that acknowledge there is a common reality and reflect our identity as a people. A people that is defined by place and history, defined by respect for different values and, most importantly, defined by a commitment to a democratically arrived at, perpetually affirmed and reaffirmed, common reality. There IS such a thing as society.

And perhaps, I can divert for just one second to mention my own interactions with our host - Channel Four. I started my career here with one of the very first film fours that was also the first film of both David Thewlis and Clive Owen – *Vroom* – at a time when I could count female directors on one hand – in fact barely more than a thumb. I made a documentary about sex workers for religious programming that was (at the time) so worrying to the channel that they asked me to record an onscreen disclaimer taking personal responsibility for it. I made a crazy version of Cinderella for New Years' Day 2000, not knowing if the century would begin with the predicted Y2K Computer meltdown or Kathlyn Turner in full flood as the wicked stepmother – both quite terrifying in prospect. And almost a decade later a documentary about the sculptor Antony Gormley – at the time the C4 controller wrote to me saying that the film **stopped time** – I believed that was a compliment but, now I think about it, it may have been a worry. And in more recent times I have been a contributor on C4 flagship News, a clear example of how an hour of thoughtful news can trump 24 hours' worth of chasing the latest. My experience provides a fleeting glimpse of the extraordinary opportunities the channel has provided to successive generations of creators, performers, and participants, to reflect and comment on diverse subjects in multiple formats. Taking risks **without** abdicating responsibility. Giving **opportunity** on both sides of the camera to many – including me – who may have struggled to be heard. So as we move, towards a convergence of media – perhaps the question is 'How much of the culture and quality of the legacy broadcast media we should fight to retain in the brisk march to digital first?'

The traditional broadcast sectors, subject to regulation, chartered to public service and regulated for the common good are being pitted against an offshore cartel unfettered by anything but their share price. Protected from

responsibility for their impact by a US law 'Section 230', weaponised by the US courts and ruthlessly defended by eyewatering levels of lobbying. Tech spends more money than the gun lobby, tobacco or big pharma to keep legislation and regulation at bay: not a fair fight.

Meanwhile, 'Social' media and its associated businesses – counterintuitively - are built to address individuals, at the cost of our common experience. And while the sector vaunts the liberatory uses of 'tech' for the individual, the public space has been rapidly privatised and held in the guardianship of a small handful of gatekeepers who have brilliantly managed the institutions of state, law, regulation, to ensure profits are privatised and the costs regressively democratised.

Streamers and video-on-demand services are deliberately designed to offer a personalised world. When I choose content based on my interests or characteristics, I am offered more of the same. While it feels comforting to be reflected, if it automatically demotes content based on other interests and alternate characteristics then I am separated from my fellow viewers. Video, images and text on digital products and services do the same but supercharged. Even hesitating for a moment over an image changes the algorithm which determines what you see next, which then determines what you might see in the future – you don't even have to click to be captured.

In the fractured world in which we have outsourced decisions to automated systems that do not understand – but powerfully act on the information they hold. It is ever more important that we all see ourselves in the context of each other. And, occasionally, even at the same time, as each other - a task in which broadcast TV still plays a pivotal role and, I would argue, excels. A task that in the world of information that is being extracted and monetised away from its role in describing our common experience – holds a different version of us.

I do want to briefly acknowledge the role of the BBC, which has suffered an unwarranted ideological assault on its core purposes. And while I do not see a straight road forward, I would like to put to bed the notion that the battle for control of our attention is between the BBC and commercial TV and radio, local press, or even the streamers. It is, of course, YouTube, Facebook, Instagram, X, TikTok, Snap – or the pushy demands of e-commerce – that dominate our cultural and information technologies. We may disagree on the specifics of funding the BBC, or feel disappointed about what is missing, who is underrepresented or poorly served. But let us not misunderstand the purpose

of a national broadcaster or the PSB system more widely. Neither culture nor politics is a zero-sum game. It does not follow that, if social media or streamers have content, we need none in our collective hands. Nor does it follow that, because this generation of the young have been highjacked by the persuasive design strategies of an advertising business model, we should plan that future for the next.

The PSB system offers the opportunity of a contemporary and collective vision of what binds us. This is a crucial time in which money rules, politics discredited, nations states weakened, and the international community divided by layers of self-interest and proxy wars. This is a time in which something that can be shared may also – at best - allow us to discern a collective path.

Technology is not a replacement for human decision making. It is a brilliant powerful tool for crunching information. It can do accounts and mess up accounts, it can identify break-through patterns that will transform medicine and mis-identify faces as criminal because of the colour of their skin. It can transform the imaginative worlds in which the stories you – in this room - make and create child sexual abuse with no limit to the horror depicted. It can capture the truth of an injustice on our smartphone, and it can create a lie that is unjustly popularised on our smartphone.

We have two jobs. One is to insist that democracy runs technology and not the reverse – that is the job of lawmakers and the electorate. The other is to hold up a mirror to the world. A mirror that builds a collective experience, contributes to a shared reality, speaks of a common truth and reflects our multiple paths – that is your job – and it could not be more important.

Thank you