KEYWORDS:
HYBRIDITY, DEEP FAKES, SYNTHETIC MEDIA
DOI:
https://doi.org/10.5771/2747-5174-2021-1-10

Miriam Meckel *is a Professor for Communication Management at the University of St. Gallen in Switzerland. A journalist and former State Secretary and government spokeswomen, she is also the Co-Founder and CEO of ada, a learning platform for innovative corporate training and development.*

Léa Steinacker *is a researcher at the University of St. Gallen in Switzerland with a research focus on the socio-technological implications of artificial intelligence. A journalist and practitioner of innovation in digital media, she is the Co-Founder and COO of ada.*

Hybrid Reality: The Rise of Deepfakes and Diverging Truths

AUTHORS: Miriam Meckel & Léa Steinacker

ABSTRACT:
While the manipulation of media has existed as long as their creation, recent advances in Artificial Intelligence (AI) have expedited the range of tampering techniques. Pictures, sound and moving images can now be altered and even generated entirely by computation. We argue that this development contributes to a "hybrid reality", a construct of both human perception and technologically driven fabrications. In using synthetic media involving deep learning, called deepfakes, as one manifestation, we show how this technological progress leads to a distorted marketplace of ideas and truths that necessitates a renegotiation of democratic processes. We synthesize implications and conclude with recommendations for how to reach a new consensus on the construction of reality.

https://doi.org/10.5771/2747-5174-2021-1-10

Every new technology brings forth its very own possibilities for manipulation. In the age of painting, parts of artworks were painted over if the statement of the picture was to be changed. Photography brought with it the possibility of historical snapshots and their manipulation. People could be removed again in retrospect. This is exactly what happened in a photograph from 1937 showing a group of people around Adolf Hitler in controversial filmmaker Leni Riefenstahl's Berlin garden. The picture exists in two variations, once with, once without Joseph Goebbels. He was simply retouched out of the photo as punishment for his unseemly enthusiasm for Riefenstahl. Editing options expanded considerably through the digitalization of photography. One well-known example went down in media history as „Reutersgate." In 2006, Lebanese photographer Adnan Hajj, on behalf of the Reuters news agency, documented an attack by the Israeli army on Beirut during the Lebanon War. As it turned out, the photos were manipulated. Hajj had intensified the columns of smoke over Beirut to make the images more dramatic (Usher, 2008).

Altering photos does not always require complicated technology. Sometimes it suffices to choose a deliberate section of the image. This was the case with a photograph taken of former U.S. President Barack Obama in 2010 while he examined the extent of the „Deep Water Horizon" oil spill on the Louisiana coast. The Economist published a clipping on its June 19, 2010 cover showing Obama alone - selected „as the ideal metaphor for a politically troubled president," as the New York Times interpreted it.[1] In the original photo, Obama can in fact be seen in conversation with other people.

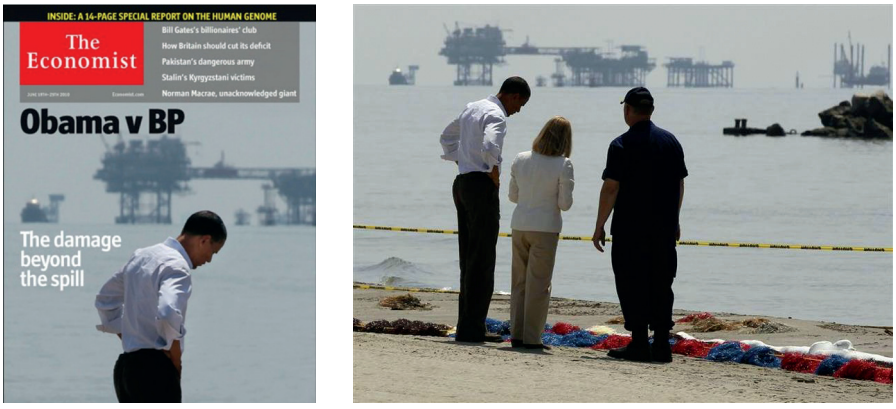


Figure 1 Collage by The Economist online, July 7th,2010

Evidently, technology is not the sole cause for the manipulation of images. Rather, altering reality has always accompanied the history of our cultural evolution in media as an option. But recent technological progress triggered the dawn of tampering with sound and moving images and with it a seismic change in the course of reality construction : Artificial intelligence (AI) has boosted the acceleration, expansion, and perfection of producing and disseminating manipulated and fake media content. It has ushered in the age of hybrid reality.

HYBRIDIZATION: FROM GAMING TO REALITY

Originally, the term "hybrid" stems from the realm of biological evolution and describes an offspring of two different species, like a crossbreed of two different dog types or of two variants of roses. Meanwhile, it has quickly disseminated into social contexts describing blends of different cultural backgrounds as well as, in more recent times, blends of offline and online media cultures (Lindgren, 2014). Concepts of hybridity play a major role in Hybrid Reality Games (HRG), and research has provided some in-depth analysis on their impact on traditional concepts of human perception and social connection. HRGs, for example, change human perception of special environments and human interaction with them. They also contribute to blurring the boundaries between public and private spaces and increasingly impact how people interact regarding the difference of private and public role-playing, of investing mutual trust and bridging different online and offline communities (De Souza e Silva & Sutko, 2008). Zappi et al. (2011, p. 356) offer a definition for hybrid reality in art as a

performance in which "[…] real world and virtual world objects coexist, and real physical objects in the user's environment play a role in (or interfere with) the computer-generated scene".

Extending the concept of hybridity or, more precisely, hybridization to reality we move beyond the realm of gaming, media and arts to a general idea of how humans perceive reality. Based on the notion of technology as "a second self" (Turkle, 2005) we thus define hybrid reality as a construct concurrently composed by human perception and by different sources of material dynamics involving physical environments, cultural trajectories and technologically based features. This is, in the first place, a description of a technologically driven transformation process that impacts how humans perceive and construct reality. In a second step, it entails a range of consequences, which we will illustrate with the phenomenon of Deepfakes as one major manifestation. They redefine fundamental societal quandaries like: What is real? Can we share a joint version of the truth? Or do we have to adapt our notion of reality construction to allow for peaceful coexistence?

HYBRID REALITY: HISTORICAL EVOLUTION OR FUCKED-UP DYSTOPIA

Reality has always been an outcome of individual and socially synchronized perception. This insight runs like a thread through the history of humankind and has particularly manifested in the paradigmatic transformation in social sciences from positivism towards constructivist epistemology (Chompalov & Popov, 2014). This shift had ramifications for the assumption of a universally valid truth and indisputable reality. „The truth," writes author Salman Rushdie, "is that truth has always been a contested idea" (2018). Humbert, the protagonist of Vladimir Nabokov's „Lolita" tells us that „reality" is „one of the few words which mean nothing without quotes." Even within quotes reality has become a point of contention in an age when almost every word, gesture and action of a person can be technologically fabricated.

Nowadays, it is possible to alter media through digital technology, at times to the point of unrecognizability. With advanced software, pictures, sound and moving images can be manipulated or even created completely from scratch. Such developments may have far-reaching consequences for public discourse and the collective construction of reality. For example, in 2019, former U.S. President Richard Nixon delivered a video address to inform viewers about the failed moon landing. „These brave men, Neil Armstrong and Edwin Aldrin, know that there is no hope for their recovery," Nixon can be seen announcing in a grave voice, „but they also know that there is hope for mankind in their sacrifice". The following year, in her 2020 Christmas video, Her Majesty the Queen Elizabeth II performed a surprisingly agile dance routine. And that same year, North Korean dictator Kim Jong Un warned the world's population in a recorded speech that "democracy is a fragile thing, more fragile than you want to believe".

Of course, none of this ever happened. These videos belong to a new category of „synthetic media" (Witness & First Draft, 2018), also called „deepfakes", a portmanteau of the AI technique of "deep learning" and the inauthenticity marker "fake". Creators of deepfakes use AI applications that merge, combine, replace, and superimpose images and video clips to create fake videos that appear almost, sometimes even very authentic (Maras & Alexandrou, 2018). These artificially altered pieces of content are based on methods that „can superimpose face images of a target person to a video of a source person to create a video of the target person doing or saying things the source person does" (Nguyen et al., 2020, p. 1). Recent technological advancements have made it a lot easier to compose these increasingly deceptive videos using face and body swaps that leave little trace of manipulation (Chawla, 2019). Deepfake technology can generate, for example, a video of a person giving a speech, without the consent of the person whose image and voice are being used (Day, 2019; Fletcher, 2018).

Meanwhile, researchers and activists have also created a series of deepfakes with the aim of exposing the wide range of possibilities for manipulation and the related dangers for public communication and individual privacy rights. In 2018, the media website BuzzFeed published a video showing former US President Barack Obama seemingly speaking from the Oval Office. The first 35 seconds show only Obama's face. In the middle of his statement, Obama drops a bombshell: "President Trump is a total and complete dipshit." He briefly pauses, then continues, "Now, you see, I would never say these things, at least not in a public address, but someone else would, someone, like Jordan Peele." At this point, a split screen appears showing Obama on the left while on the right we see the US actor, comedian, and director, Jordan Peele. Obama's and Peele's facial expressions and

lip movements match almost perfectly. Using AI, Peele's production team has digitally reconstructed Obama's face to mirror his own and Peele impersonates Obama with his voice, who says at the end of the clip: "We are entering an era in which our enemies can make it look like anyone is saying anything in at any point in time". And he ends: "How we move forward in the age of information is going to be the difference between whether we survive or whether we become some kind of fucked-up dystopia." Is it survival versus dystopia in the era of hybrid realities? At least the much discussed early examples of deepfakes show that their origins involve some rather dystopian intent.

SEX AND THE EVOLUTION OF DEEPFAKES

It all started with sex. In 2017, an anonymous user of the content aggregation and discussion platform Reddit with the user alias „Deepfakes" applied machine learning techniques to produce a series of sex videos involving the appearance of Hollywood stars like Taylor Swift, Emma Watson and Scarlett Johansson without their consent. By now, pornographic content accounts for more than 90 percent of deepfakes on the Internet (Deeptrace, 2019, p. 1). Reddit banned the eponymous user from its platform. But with this nefarious act, the craft of being able to fake the likeness of people in sound, image, and movement was born. Deepfakes went from a username to a brand for a new reality in which, at first glance, reality is made of what technology deceptively suggests.

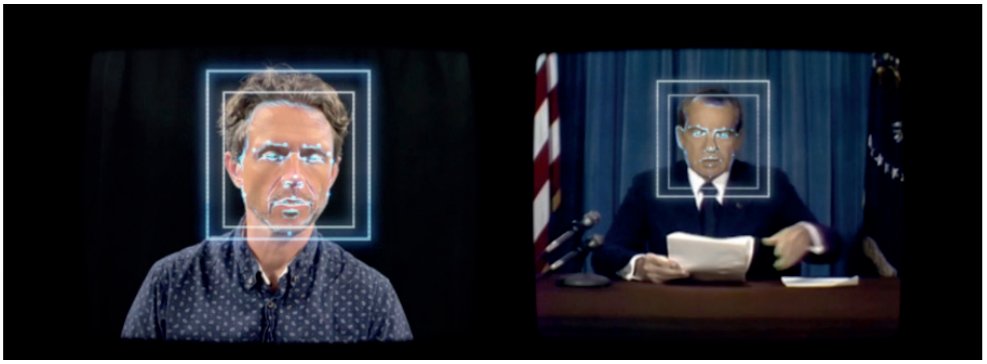
Plenty of freely available software and numerous apps can be used to manipulate photos, voices and videos in a way that pretends to reflect reality. With „FaceApp," an application from Russia, the image of a woman can turn into her likeness as a man and vice versa. A clean-shaven face suddenly features a full beard, or a person ages thirty years in either direction at the speed of a click. With the app „Avatarify" iPhone users can create Deepfakes on their phone. A synthetic video of a Hollywood actor can be used to congratulate friends on their birthday, or the user can sing „My heart will go on" in the voice of Celine Dion.

Even the human voice is no exemption from being altered in the form of deepfakes. Canadian company „Lyrebird" caused a sensation in 2017 with an audio recording in which Barack Obama, Donald Trump and Hillary Clinton talked about „fake news". The conversation had never taken place; nor did any of the sentences originate from the politicians themselves. The clip had not only been cleverly edited, it had been entirely artificially generated by a computer. This has wide-ranging implications for the trust we place in the authenticity of the individual human voice. In 2019, a German energy company's subsidiary in Britain reported one of the first cases of an AI-generated voice heist. A director-level employee received a call in which he heard an identical copy of his German boss's voice, instructing him to wire 240,000 USD to a Hungarian account. The voice turned out to be a synthetic imitation as part of a financial fraud scheme. The firm's insurance company's spokeswoman described that „the software was able to imitate the voice, and not only the voice: the tonality, the punctuation, the German accent" (Harwell, 2019).

GANS: THE COMPETITION OF NETWORKS

In the realm of deepfakes, the major game-changing factor is the potential of machine learning that allows for a broader scope, scale, and sophistication. The technology behind many of these applications makes use of Generative Adversarial Networks (GAN), artificial neural networks that evolve in parallel at many levels. A GAN uses two neural networks simultaneously: One learns to analyze images and videos, the other compares the original data with the results of the first network. Its objective is for the first network to become so accurate that the second can no longer detect differences between the original and the fake. So even for algorithms, competition may increase quality. And while humans give the impetus for this construction of reality, the networks continuously do their bidding: they learn and improve on their own.

Nowadays, a skilled technologist with enough computing power and the necessary data can fabricate videos or develop synthetic media that are practically indistinguishable from authentic content for a human eye (Fletcher, 2018). While early examples of deepfakes focused on political leaders, like the North-Korean leader Kim Jong Un mentioned above, and actresses, like Scarlett Johansson, being turned into a porn star (Hasan & Salah, 2019), in the future they might have a



MIT Center for Advanced Virtuality: In Event of Moon Disaster

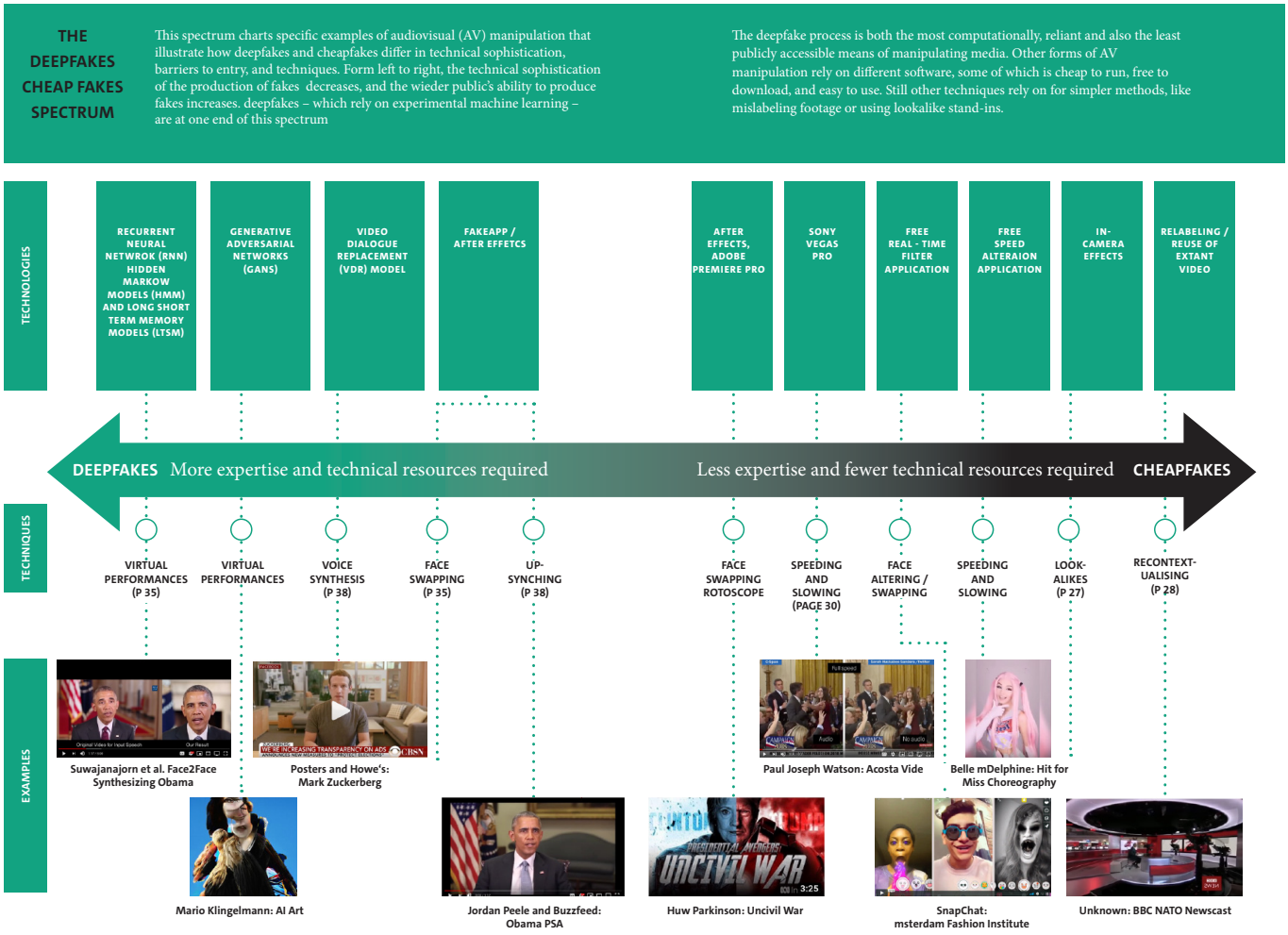
range of other malicious purposes: tweaking political messages, faking video evidence in courts, social sabotage, terrorist propaganda, market manipulation, and fake news (Maras & Alexandrou, 2019). In the 2016 presidential campaign Republican Senator Marco Rubio likened Deepfakes to modern day nuclear weapons. "In the old day," he said to an audience in Washington, if you wanted to threaten the United States, you needed 10 aircraft carriers, and nuclear weapons, and long-range missiles. Today, you just need access to our internet system, to our banking system, to our electrical grid and infrastructure, and increasingly, all you need is the ability to produce a very realistic fake video that could undermine our elections, that could throw our country into tremendous crisis internally and weaken us deeply."[2]

While machine learning techniques have fostered the generation of artificial realities, our susceptibility to them is further fueled by peculiarities of human psychology. For example, the „illusory truth effect" states that repeated exposure reinforces our tendency to believe something, even if it is incorrect. That is to say, the mere repetition of information can systematically bias people to believe it is true (Dechene et al., 2010). In the 2016 US presidential elections, a BuzzFeed News analysis found that fake news stories about the election generated more total engagement on Facebook than the top factual election stories from 19 major news outlets combined. [3] The sheer repetition of fake stories seems to have turned them into believable or at the very least shareable factoids irrespective of their actual factuality.

As complementary effects, confirmation bias, selective exposure, and a lack of analytical thinking skills can impact how people perceive various forms of misinformation (Wardle & Derakshan, 2018, p. 46). The combination leads to individuals pursuing, finding and accepting information that aligns with their pre-existing attitudes. This psychological mechanism, on the one hand, simplifies information-seeking, on the other hand it renders specific groups of individuals prone to believing fake news, pictures or videos (Frimer et al. 2017; Tandoc, 2019).

Technological opportunity combined with psychological susceptibility is a particularly difficult fusion. Additionally, the Internet offers a dissemination platform built for virality that allows to make deepfakes accessible to a broad public within seconds. Network effects further elevate these synthetic media as tools in a specific form of political communication: propaganda. An example of this form of disinformation is a deepfake video showing former U.S. President Donald Trump in the Oval Office not only absolving himself of any guilt in the impeachment process, but also announcing that financier and sex offender Jeffrey Epstein had not taken his own life in prison.[4]

Especially in the course of Donald Trump's second presidential campaign, deepfakes took on considerable significance and became a part of political communication. This does not always require complex technical manipulation, as a video of a 2020 press conference involving House Speaker Nancy Pelosi shows. In it, Pelosi appears to be slurring her words. In fact, the video, initially posted on TikTok, had been slowed down significantly, with individual portions of Pelosi's responses edited out to make her appear incoherent. While Deepfakes represent the technologically advanced side of audio and video alterations, „cheap fakes" can be fabricated with simple software or no software at all (Paris & Donovan, 2019, p. 2). In light of these developments, the Economist warned of „the art of the lie" in times of „post truth politics" (Economist, Sept. 10, 2016), the Financial Times of the „post factual age" (Barber, 2017). From a cynical point of view, one could even call this the democratization of lies, as



Source: Paris & Donovan 2019, p. 10 f.

everyone can use available technologies to doctor audio and video material for the sake of supporting one's own position and view of reality.

Deepfakes, therefore, are a means of distorting public opinion formation that is aimed at drowning truthful content and images of reality in a sea of fabricated audio-visual content. That is part of a communication strategy meant to at least distract people from engaging with factual content, often rather influencing them to believe in conspiracy narratives (Westerlund, 2019), a phenomenon media sociologist Zeynep Tufekci calls "whistle-drowning".[5] This process, hence, contributes to deteriorating digital literacy and citizens' trust toward authority- and expert-provided information. „Indeed, deepfake videos, pornographic or merely propagandistic, may well finish off photography as what we wanted it to be since 1839 - a largely trustworthy documentation of the actual", writes Rebecca Solnit (2018, p. 9) adding: "Perhaps all this will lead to an era in which no one believes anything, and everything solid that hasn't already melted into air liquefies into slime." Harvard's history professor Jill Lepore (2016, p. 93) writes: "The era of the fact is coming to an end: the place once held by 'facts' is being taken over by 'data.' This is making for more epistemological mayhem." The former New York Times' critic, Michiko Kakutani, even reaches a step further. In her book „Death of Truth" (2018) she writes that communication strategies impelled during the presidency of Donald Trump can be interpreted as "emblematic of dynamics that have been churning beneath the surface of daily life for years, creating the perfect ecosystem in which Veritas, the goddess of truth […] could fall mortally ill."

HYBRID DEMOCRACY: A NEW MARKETPLACE OF IDEAS - AND TRUTHS

Throughout history, truth has never had an easy run. But if it ever had an identity, with the help of AI, humans are currently engaged in the greatest identity theft imaginable. Various manifestations

of the technology are fundamentally changing our perception and interpretation of reality. When you can no longer distinguish fact from fiction, the real from the imaginary, and a lie from the truth, everything can be called into question. US technology expert Aviv Ovadya coined the term „infocalypse" for this development in 2016 (Warzel, 2018). It describes a process in which mis-, is- and mal-information increasingly influence public discourse up to a threshold at which the distinction between fair and fake information will be difficult to draw. As a potential result, society might be overwhelmed and people might retreat to „reality apathy". In other words: In a hybrid reality everything seeking the truth becomes much more complicated and relative.

Yet the search for factuality and truth was conceptualized differently by the pioneers of democracy. In 1644, in a „Speech for the Liberty of Unlicensed Printing to the Parliament of England," John Milton (1890) for the first time developed the concept of a contest of forces in public: In a free and open encounter of arguments, truth wins out, Milton assumed, trusting in the positive forces of competition. John Stuart Mill (1991) further sharpened the idea in his book On Liberty that first was published 1859. In it he writes that information and opinion compete in a free market and that truth emerges from competition among them. He reasons that even dissenting opinion by a single individual offers great value to society because it may turn out to be true. Humans are error-prone and if a minority opinion is not protected humanity will be „deprived of the opportunity of exchanging error for truth." Interestingly, Mill also asserts that even false opinions might be valuable because they require that accepted truths need to be continually defended and would otherwise be taken for granted. This, eventually, will contribute to creating a „clearer perception and livelier impression of truth."

Mill and Milton never heard of deepfakes, though. From today's perspective they might have been concerned. Recent developments like the ones described above let experts doubt whether the marketplace of ideas might have failed supporting the establishment of truth. „John Stuart Mill's notion that a ‚marketplace of ideas' will elevate the truth is flatly belied by the virality of fake news" writes Zeynep Tufekci (2018). Personalization algorithms, misinformation and disinformation on social media and the option to create deepfakes with an increasing perfection propel the distortion of reality.

One might come to conclude that reality is currently behaving a bit like money in the financial crisis: All of a sudden, a bundle of reality derivatives is emerging, and hardly anyone can identify their origins, components and values. Is that still a market, or rather a bubble of illusory realities deliberately pumped up by various groups? Perhaps the concept of a marketplace of ideas has indeed failed. Or perhaps it simply needs some differentiated adjustments. Markets do not always function perfectly. The state responds to such „market imperfections" through regulation, e.g. antitrust and competition law. In an analogous application, is disinformation also just another word for the distortion of competition in the marketplace of opinions (Lee, 2010)? The British economist Ronald Coase (1974, p. 384) pointed out an essential difference in approach: „In the market for goods, government regulation is desirable, whereas, in the market for ideas, government regulation is undesirable and should be strictly limited." There are good reasons for this difference: Freedom of speech and freedom of the press are constitutionally protected goods that the state should interfere with as little as possible.

While the choice and competition between competing goods and services can be described and quantified quite clearly, the situation with competing ideas or truths is more complicated. Those who want to take regulatory action against those who distort the competition of information on the Internet quickly get into conflict with „freedom of speech." Some deepfakes, however, might be an exception here. Under certain conditions the technical production of deceptively real audio and video is not necessarily covered by the liberties of free speech. If they are produced and used to deceive people and influence their perception of reality, e.g., in favor of certain political goals, it can be argued that regulation in the sense of „distortion of competition" in the marketplace of ideas can take effect here. Much public debate is needed to figure out the extent and approach of regulation towards the fabrication of Deepfakes as a challenge for facts, democracy and a functioning marketplace of ideas. The European Commission (2021) recently proposed rules to ban „AI systems considered a clear threat to the safety, livelihoods and rights of people". In the document the Commission warns "that technology can also be misused and provide novel and powerful tools for manipulative, exploitative and social control practices" (EU Commission, 2021, p. 21). However, the term Deepfakes does not appear once in the document.

TO UNMASK TECHNOLOGICAL MANIPULATION WE NEED … MORE TECHNOLOGY

How can we deal with a reality that offers us a growing number of technically perfected variants and derivatives of itself, made possible by AI's progress in generating synthetic voices, images, and videos resembling real people, often those in public life?

First, this development will entail an adaptation of certain „realism heuristics" (Frenda at al., 2013). Up to now, people consider audio and video material as solid representations of a real world experience. We will have to partly unlearn this conceptualization and get acquainted with a concept of hybrid reality that requires a much more differentiated and detailed interpretation of markers of reality as many of the conventional notions can now be technologically fabricated, altered and distorted.

Second, this development will pose far-reaching challenges to the educational system. The technical possibilities for the production of deepfakes and their possible social, political and economic implications need to be part of the curriculum of every school and secondary educational institutions. Awareness of the phenomenon of deepfakes, the ability to question presented realities and to use technological tools to verify the authenticity of videos must become a part of digital literacy.

Third, relatedly, technologically based techniques of detecting synthetic media are becoming more ubiquitous. They seek to automatically determine whether different image sources have been connected in a video, additional pixels have been inserted, or whether elements have been used that demonstrably originate from other sources and have been embedded in a video. For some detection methods, the same technology can be used as was employed to create the deepfakes: GANs can also be used to anonymize faces to generate highly realistic faces with a seamless transition between the generated face and the existing background (Hukkelas, Mester & Lindseth, 2019).

Source: Hukkelas, Mester & Lindseth, 2019

In the future, it will also be possible to hide clues in picture frames that may indicate the authenticity of a video. This variant, called "visual steganography", operates by producing visual watermarks to unobtrusively conceal a secret message within other data. The visual steganography technique hides a full-sized color image or video within another (Weng et al. 2018).

There are also opportunities of manual forensic detection of video manipulation, like the " Eulerian Video Magnification", a technology that allows to see the visible pulse rate of a person that would not be detectable in a deepfake video. Recently, research has shown that the Eulerian Video Magnification framework can be used for visualizing the human pulse or tiny movements in video material to detect fakes (Das, Negy & Smeaton, 2021).

Whatever technological progress will allow for in terms of detecting, deciphering or even preventing deepfakes that distort individual or public perception of reality, the progress in detection will parallel that in creation and vice versa, which might be called „an ongoing arms race between manual and automatic synthesis of media, and manual and automatic forensic approaches." (Witness & First Draft, 2018, p. 8). This prognosis might also entail a new technological divide between those able to cope with the advancements of deceptive uses of technology and those left to be deceived. "Small communities of the technologically skilled may be able to discern the glitches introduced by GAN software and report the fakery online," write Vaccari & Chadwick, "though in

the long term there is also the problem that AI-based methods of detection will become victims of their own success because the training data sets will be used by malicious actors to further refine production of deepfakes" (Vaccari & Chadwick, 2020).

1 https://mediadecoder.blogs.nytimes.com/2010/07/05/on-the-economists-cover-only-a-part-of-the-picture/
2 https://www.matellio.com/blog/deepfakes-a-modern-day-nuclear-weapon-or-a-passing-it-prank/
3 https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook
4 https://www.youtube.com/watch?v=aPp5lcqgISk
5 https://www.npr.org/2016/10/22/498954190/wikileaks-dump-method-destroys-privacy-sociologist-says-not-all-leaked-pass-publ?t=1618921838890

CONCLUSION:

Reality has never been the one single concept we could possibly all believe in but a composition of representations of physical reality as well constructions of individual perception. Deepfakes as fabricated illusions of reality are just one technical option to influence our take on reality. They will be one central element in an evolving composition of hybrid reality. Merely regulating deepfakes by banning technological applications for production and dissemination will rarely be successful. Specific requirements, e.g. that of a human-like chatbot self-identifying as such, might work be a role model for a possible mixture of legal rules and corporate self-regulation. Using technological methods such as watermarking or Eulerian Video Magnification to detect deepfakes might be even more promising. Service providers specialized in this field should focus on informing the public and boosting a broader range of digital literacy amongst citizens and users in general. Reality has never been and never will be naturally given. In the process of historical evolution and technological progress we are now entering the era of hybrid reality.

REFERENCES:

Barber, L. (2017). Fake news in the post-factual age. Lecture to Oxford Alumni Festival, Oxford university. Financial Times, September 16th.

Chawla, R. (2019). Deepfakes: How a pervert shook the world. International Journal of Advance Research and Development, 4(6): 4–8.

Chompalov, I. M. & Popov, L. S. (2014). Sociology of Science and the Turn to Social Constructivism. Social Sciences (3)2: 59-66.

Coase, R. H. (1974). The Market for Goods and the Market for Ideas. The American Economic Review (64) 2: 384-391.

Das, R., Negi, G. & Smeaton, A. F. (2021). Detecting Deepfake Videos Using Euler Video Magnification. arXiv:2101.11563

Day, C. (2019). The Future of Misinformation. Computing in Science & Engineering, 21(1): 108–108. https://doi.org/10.1109/MCSE.2018.2874117

Dechêne, A., Stahl, C., Hansen, J., Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. Personality and Social Psychology Review, 14, pp. 238-257.

Deeptrace (2019). The State of Deepfakes. Landscape, Threats, and Impact. https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

De Souza e Silva, A., Sutko, D. M. (2008). Playing Life and Living Play: How Hybrid Reality Games Reframe Space, Play, and the Ordinary. Critical Studies in Media Communication, 25:5: 447-465, DOI: 10.1080/15295030802468081

Dolhansky, B. et al. (2019). The Deepfake Detection Challenge (DFDC) Preview Dataset. arXiv:1910.08854

European Commission (2021). Proposal for a Regulation of the European Parliament and of the council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. Brussels.

Fletcher, J. (2018). Deepfakes, Artificial Intelligence, and Some Kind of Dystopia: The New Faces of Online Post-Fact Performance. Theatre Journal, 70(4): 455–471. Project MUSE, https://doi.org/10.1353/tj.2018.0097

Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. Journal of Experimental Social Psychology, 49(2): 280–286.

Frimer, J. A., Skitka, L. J. & Motyl, M. (2017). Liberals and conservatives are similarly motivated to avoid exposure to one another's opinions. Journal of Expertimental Social Psychology 72: 1-12.

Harwell, Drew. (2019). An artificial-intelligence first: voice-mimicking software reportedly used in a major theft. Washington Post, September 5th: A15.

Hasan, H. R., & Salah, K. (2019). Combating Deepfake Videos Using Blockchain and Smart Contracts. IEEE Access, 7: 41596–41606. https://doi.org/10.1109/ACCESS.2019.2905689

Hukkelås, H., Mester, R. & Lindseth, F. (2019). DeepPrivacy: A Generative Adversarial Network for Face Anonymization. In: Bebis G. et al. (eds). Advances in Visual Computing. ISVC 2019. Lecture Notes in Computer Science, vol 11844. Springer, Cham: 9-44. https://doi.org/10.1007/978-3-030-33720-9_44

Kakutani, M. (2018). The Death of Truth: Notes on Falsehood in the Age of Trump. New York: Tim Duggan Books.

Lee, S. P. (2010). Hate Speech in the Marketplace of Ideas. Glolash, D. (ed.), Freedom of Expression in a Diverse World. Dordrecht: Springer, 13-25.

Lepore, J. (2016). After the fact. In the history of truth, a new chapter begins. New Yorker, March 21: 91-95.

Lindgren, S. (2014). Hybrid Media Culture. Sensing Place in a World of Flows. London: Routledge.

Maras, M. H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. International Journal of Evidence & Proof, 23(3): 255–262. https://doi.org/10.1177/1365712718807226

Mill, John S. (1991). On Liberty and Other Essays. Oxford: Oxford University Press.

Milton, J. (1890). Areopagitica: A speech of Mr. Milton for the liberty of unlicensed printing, to the Parliament of England. New York: The Grolier Club.

Nguyen, T. T., Nguyen C. M., Nguyen, D. T., Nguyen, D. T. & Nahavandi, S. (2020). Deep Learning for Deepfakes Creation and Detection: A Survey. https://arxiv.org/pdf/1909.11573.pdf

Paris, B. & Donovan, J. (2019). Deepfakes and Cheap Fakes. Data & Society, September.

Rushdie, S. (2018). Tuth, Lies, and Literature. New Yorker, May 31st. https://www.newyorker.com/culture/cultural-comment/truth-lies-and-literature
Solnit, R. (2018). Driven to Distraction. Harper's Magazine, May: 7-9

Tandoc Jr., E. C. (2019). The facts of fake news: A research review. Sociology Compass (13)9: 1-9. https://doi.org/10.1111/soc4.12724

Turkle, S. (2005). The Second Self. Computers and the Human Spirit. Cambridge/Mass: The MIT Press).
Tufecki, Z. (2018). It's the (Democracy-Poisoning) golden Age of Free Speech. Wired, June, 1st.
Usher, N. (2008). Reviewing Fauxtography: A blog- driven challenge to mass media power without the promises of networked publicity. First Monday (13)12, December, 1st.https://firstmonday.org/ojs/index.php/fm/article/download/2158/2055

Vaccari, C. & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. Social Media + Society, January 2020. doi:10.1177/2056305120903408

Wardle, C. & Deraksham, H. (2018). Thinking about 'information disorder': formats of misinformation, disinformation, and mal-information. UNESCO (ed.). Journalism, 'Fake News' & Disinformation. Handbook for Journalism Education and Training, Paris: 44-56.

Warzel, C. (2018). Believable: The Terrifying Future Of Fake News. Buzzfeed, February 11. https://www.buzzfeednews.com/article/charliewarzel/the-terrifying-future-of-fake-news

Weng, X., Li, Y., Chi, L. & Mu, Y. (2018). Convolutional Video Steganography with Temporal Residual Modeling. https://arxiv.org/pdf/1806.02941.pdf

Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review (9)11: 40-53.

Witness & First Draft (2018). Mal-uses of AI-generated Synthetic Media and Deepfakes: Pragmatic Solutions Discovery Convening. http://www.mediafire.com/file/q5juw7dc3a2w8p7/Deepfakes_Final.pdf/file

Zappi, V., Mazzanti, D., Brogni, A. & Caldwell, D. (2011). Design and Evaluation of a Hybrid Reality Performance. Proceedings of the International Conference on New Interfaces for Musical Expression, Oslo, Norway: 355-360.

https://doi.org/10.5771/2747-5174-2021-1-10