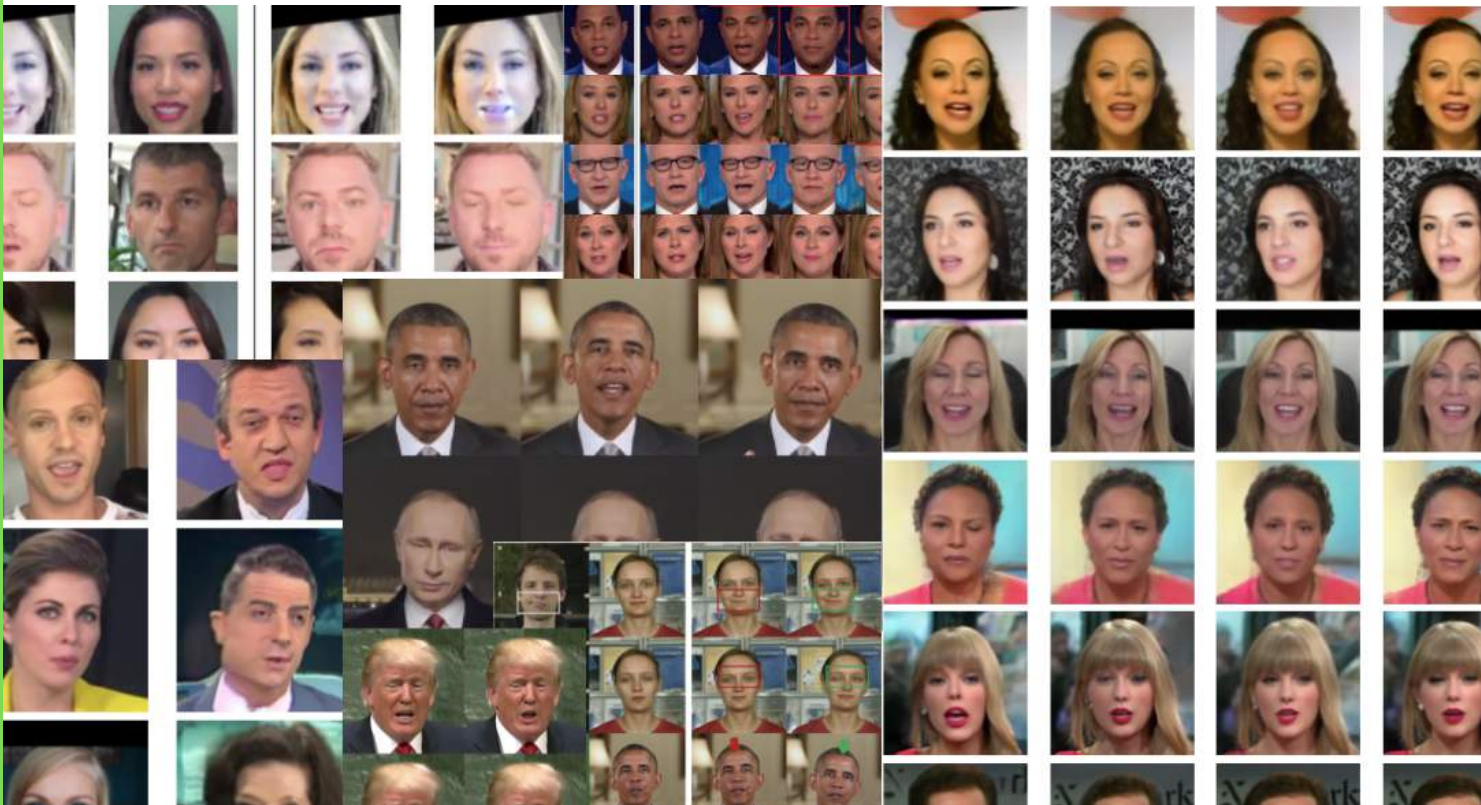


DESIRES DRIVING DEEPFAKES:

WHERE TECHNOLOGY ENCOUNTERS DEATH



UTRECHT UNIVERSITY
NM&DC MASTER'S THESIS
JULIA ALEXEENKO, 0468706
DR. IMAR DE VRIES
APRIL 28, 2021

Acknowledgements

Writing this thesis in the unprecedented times™, I was lucky to receive a lot of help, advice, and support.

I would like to thank my supervisor, Dr. Imar de Vries, for thorough feedback, understanding, artistic choice of MS Team backgrounds, and a plentiful of great books and film references.

I express my gratitude to Dr. Iryna Makhouskaya for her enthusiasm and believing I could do this after boiling in History in Ethnography for a very long period.

Thank you, Nicky Heijmen, for proofreading, *harde taal*, personalized recommendations for *Moist memes Maison*; thank you, Hester Julia Voddé for all the great teamwork; and you both for the Newer Media Team, tunes, and books. You two have messed up my Spotify forever, which is a beautiful thing. Thank you, Renske van der Wal, for tea, cakes, random rooftop philosophy talks, making me happy with Foucault, and reading this text with a handful of colored pens.

I cannot even express how grateful I am to Anna Mahlysh for *being* and for reminding me of the importance to be gentle to ourselves; and to Herbert Bitter for his support, sleep policing, listening to my whining, and willingly being lectured on postphenomenology for as long as I needed to lecture to understand it myself.

I thank Emmaus Haarzuilens, as an idea, a place, and a great gathering of individuals, for showing me how hard work makes lots of things possible. I also thank my grandparents for encouraging me to do anything. I thank my mother, Irina, for keeping on asking what I want to be when I grow up.

Thesis summary

In this thesis, I argue how the discourse around deepfakes is drenched with fantasies of limitless mastery of a subject upon serviceable objects. Employing postphenomenological optics, I access the technology and the fantasies around it through the published accounts of mediation (technical and non-technical publications on deepfake), discovering the ideological material in how the technology is discussed. I conduct a Foucauldian discourse analysis (FDA) to discover and describe the objects formed in the discourse through explicating or concealing aspects of deepfake use. In the course of an FDA, I locate fantasies that align with a utopian program of streamlined communication, limitless mastery, and control, along with dystopian fantasies of societies disrupted by the deceiving powers of technology. By employing the media-archaeological method of digging, I discover and examine the (dis)continuities in how political-economic configurations of deepfakes are tied in ideological material and trace this entanglement in the stabilities, marginalized and dominant, of deepfake.

This thesis brings the deepfake instance to the academic debate about the fantasies and desires that drive technological development. I conclude that the ambiguous character of fantasies about deepfakes closes the span of (design) choices and that the dominant in the existing academic debate view of the deepfakes (as a new level of improving deceiving media) does not allow for discussing this technology as it is (pre)configured in terms of political economy. The inquiry contributes to changing the ideological material that constructs the meaning of deepfakes.

Keywords: deepfake, digging, Foucauldian discourse analysis, GAN, media fantasies, (socio-)technical imaginary, resurrection.

Table of contents

Introduction	1
Chapter 1. Theoretical framework: naming the warm dough of desires	4
1.1. Situating deepfake resurrection in academic debates	4
1.2. Fantastic fantasies and where to find them	7
1.3. Enchantments of technology and imagined rationality	8
1.4. On being made of code, desire, sunlight, and ideology	9
Chapter 2. Methodology: a blue sky and a full tank of postphenomenology	10
2.1. The meaning we make of things before they make us	10
2.2. Chasing authoritative texts beyond the “cool-or-creepy”	11
2.3. Getting to the implied: method	14
2.3.a. Anything is a Foucault when you are brave enough	14
2.3.b. Hovering between media theory and speculative work: digging	15
Chapter 3. Who gazes, who controls the gaze?	17
3.1. Crossing the uncanny valley: harder, better, faster, stronger cheaper	17
3.2. We’ve lost control (again, or maybe never had it in the first place)	20
3.3. Reconciling the found and the envisioned	23
Chapter 4. Choose your unanswerable questions	27
4.1. Where deepfakes encounter death of social groups	27
4.2. One cannot abuse quantified identities—deepfake machine goes brrrr	32
Conclusion	37
References	42
Appendix 1. Corpus of technical research publications	51
Appendix 2. Corpus of non-technical expert publications	77

Introduction

Can deepfakes put flesh on the bones of artificial consciousness? This question is a quote from an article by deepfake enthusiasts speculating, among other things, about the possibility of deepfake-facilitated resurrection (Deepware 2020). *Deepfake* is a process and result of training neural networks to realistically swap individuals' faces (Wagner & Blewer 2019). Given the most obvious affordance of deepfake—manipulating video/audio material to make people “say” and “do” things they have not said or done—the phenomenon is often viewed as a direct political and societal threat (Westerlund, 2019; Whyte, 2020). Yet it is expected to have positive applications, too: enhance film production, diminish communicative barriers (Whittaker et al. 2020), and allow the digital resurrection of people (Westerlund 2019; Kwok & Koh 2020). In the entertainment industry, whys and why-nots of “resurrecting” deceased celebrities through deepfake are a hot item (e.g., Dimitropoulos 2020; Singer & Connolly 2019). Celebrities, including politicians, are easy “prey” for deepfake manipulation as there is enough free-access footage of them. If we are to believe trend-watchers and journalists, we all become more deepfakeable as we store and distribute our photos and videos (e.g., Deep Nostalgia™ by MyHeritage 2021; Future Funeral 2021; Kneese 2021; Linder 2021) and as the technology gradually becomes cheaper and more productive (Poulopoulos 2020). Reanimating people (in some form) by technology is thus seen as a near-future possibility.

The emergence of deepfakes lends itself willingly to be analyzed as a part of a linear development of fake news becoming less traceable and more convincing. While experts address deepfake resurrection as one of the positive applications of the technology, the academic debate circles primarily around deepfake forensics (e.g., Westerlund 2019; Whyte 2020) and shaping a legislative response to the phenomenon (e.g., Brown 2020). Another ongoing academic discussion concerns the consequences of an opportunity of *resurrection by tech*—I here define it as fantasies about and practices of reanimating deceased persons with the use of technology as they figure in academic debate, journalistic publications, and cultural

texts. This opportunity invites reflection upon the rights, agency, authenticity of the resurrected (Leaver 2019), perception of real and fake (Whittaker et al. 2020), and commercialization of death (Öhman & Floridi 2017). Very practical, these debates are aimed to act upon the future imagined from the standpoint of technological determinism. Predicting the effects here overshadows understanding how the technology develops in the first place. The technology is then viewed, following Mark Ratto's reading of Bruno Latour's notion of "matter of fact/concern," as something we appear to have no control over and thus should not care about (Ratto 2011). Such a viewpoint narrows the space for rethinking configurations of the technology that could yield different sets of social and cultural practices. It also heavily influences how we frame the complexities of being human, inflicting it with the "clichés of consumerist-oriented industrial design" (Hertz 2015: 6).

Is it possible to analyze deepfake without the lens of technological determinism and linear progress; as a technology that yet can be (re)configured? To do so, an inquiry into what comes before this technology, what informs its development and application is necessary. The fantasies of, i.a., *deepfake resurrection*—I here define it as possible results of manipulating the images of or content created by the deceased to achieve an effect of posthumous presence or performance—offer an entry point to analyzing the irresistibility of the technology. Fantasies, myths, and desires inform *(socio-)technical* imaginary (Marvin 1990: 4–6), that is a particular case of *social imaginaries*,¹ holding socially-shared visions of how technology could/should be used by social groups, along with anticipated consequences of the uses (Thurlow & Sawyer 2019: 171). Critical examination of such fantasies in the discourse around deepfake is, therefore, at the heart of my attempt to initiate an academic conversation on what makes the deepfake technology possible and compelling.

The research question of this thesis is, therefore, *What fantasies can be found in justifications of the development of the technology allowing for deepfake resurrection, and how can they be traced in the workings of the technology?* The following two sub-questions are to guide my analysis:

¹ I elaborate on these concepts in chapter 1.

- What fantasies are voiced in the legitimations of deepfake technology development in technical research and non-technical expert publications touching upon deepfake resurrection?
- What interests and hopes are furthered in these fantasies through explaining the (mis)uses of the technology and imagining its (im)possibilities?

To answer my research question, in chapter 1 of this thesis, I present my readings of the existing research concerning deepfake in general and deepfake resurrection in particular, along with the theoretical framework that guides my research. In chapter 2, I elaborate on the methodological foundations of this thesis, including postphenomenological optics that enable me to access the technology through the published *mediation* accounts. In this chapter, I also explicate the employment of Foucauldian discourse analysis and media-archaeological method of *digging* to examine the research corpus consisting of a) technical papers² and b) non-technical expert publications on deepfake resurrection.³ In chapter 3, I conduct discourse analysis to explore the fantasies around deepfake technology development in technical research and non-technical expert publications. Finally, chapter 4 presents my interpretations (obtained by the media-archaeological method of digging) of how fantasies expressed in discussions around the technology further particular interests and desires through imagining deepfake's (im)possibilities.

The aim of this research is not to describe the exact cause-and-effect mechanics between technology and fantasies. Because fantasies cannot be excavated quite literally, this research is an experiment in excavating them metaphorically, as given in language, “externalized in the public symbolic space” (Žižek 1997: 212). By conducting this experiment, I aim to eschew the technological-determinist optics and instead render visible how the fantasies around deepfake set ideological coordinates for collective expectations towards technology (Jørgensen & Tafdrup 2017: 98) and thus

² Technical research features (often partially and implicitly) justifications of the development of technologies making deepfakes possible, disclosing its ideal state, eventual uses and abuses, and, accordingly, fantasies as to why the technology is needed.

³ This includes interviews with experts who create this kind of software (e.g., Ajder 2019) and publications explaining the phenomenon of deepfake and its uses to wider audiences (e.g., Roose 2019; Cole 2018).

its political-economic configurations. I argue that this happens precisely because the technological-determinist position allows for a blind spot where all reflection upon the changing perception of real and fake (Whittaker et al. 2020) and perils of death commercialization (Öhman & Floridi 2017) become a matter of fact, making the efforts to prepare legislative and ethical responses to deepfakes inevitably lag behind its reality.

Besides bringing the deepfake instance to the discussion about the fantasies and desires that drive our technology, I aim to contribute to creating (however fragmented and particular) media histories telling of how technologies are configured in certain ways in terms of political economy, and do this rather now, before the stabilities of this one medium would become history.

Chapter 1. Theoretical framework: naming the warm dough of desires

In this chapter, I provide an overview of how the phenomenon of deepfake is reflected upon in the existing academic research (1.1) and proceed to build a theoretical framework for answering my research questions by defining the notions of *(socio-)technical imaginary*, *media fantasies*, and *enchantments of technology* (1.2, 1.3). I also shortly touch upon some theoretical limitations of this thesis (1.4).

1.1. Situating deepfake resurrection in academic debates

Deepfake videos in their current form are possible since 2014 (Goodfellow et al. 2014), and the technology develops rapidly. In academic research, a major discussion of the possible maltreatments of the technology is taking place, revolving around the concern that synthetic media become more “true to life,” blurring the boundaries between genuine and fake (Whittaker et al. 2020). Explicitly addressing the issue of deepfake resurrection, authors resort to a rather slippery notion of *respect* as a

condition for ethically-responsible resurrection by tech (e.g., *ibid.*; Kwok & Koh 2020). While the dichotomy of *real* versus *not real* is addressed excessively, authors refer to the famous cases of deepfake resurrection as a very *real* profit generator. A discrepancy when it comes to what makes deepfakes both an “opportunity” and a “danger,” i.e., being true to life, is generally ignored in the existing research.

According to Tama Leaver, the “true-to-lifeness” can be read not as realness but as a promise of disembodied immortality (Leaver 2019). Drawing upon a wide range of cultural narratives, Leaver provides an account of how reanimation endeavors have for long occupied human minds, exemplifying his findings with cases of three start-ups⁴ that (planned to) monetize the ever-present fantasy of resurrection by technology. Not touching specifically upon deepfakes, Leaver nonetheless formulates extremely apt points of critique towards synthetic media. He emphasizes that such endeavors depart from what N. Katherine Hayles calls “fantasies of unlimited power and disembodied immortality” (Hayles 1999: 5) and asserts that tech start-ups tend to treat their desired outcomes as a justification for the means of achieving. Consequently (and somewhat disappointingly), Leaver emphasizes the need for preclusive ethical evaluations of technological possibilities. The most curious and essential for my research remark in the cited text refers to the fact that using digital technology for resurrection (or at least fantasizing about it) is a “natural” extension of the human urge to apply communication technologies to both communicate with the dead and to reach immortality (Leaver 2019: 8).

This urge to use technology to communicate with the dead is examined by Tony Walter in his article “Communication Media and the Dead: From the Stone Age to Facebook.” The most interesting by-product of his research is tracing how concerns about the deceased have historically served political and economic goals, regularly resulting in popularizing new technologies (Walter 2015). From this perspective of collective and personal memory (e.g., Brubaker & Hayes 2011), the deepfake technology can be seen as able to change how “the dead communicate with the living and how the living (...) remember their dead” (Walter 2015: 216). Being significant for ethical inquiries, such analysis does not allow to meaningfully

⁴ LivesOn, Eterni.me, and Humai.

distinguish between different media, as they all can influence the processes of remembering. It also all too willingly renders the origins of any medium obvious: sure, we try to resurrect our deceased to communicate with them; sure, any communication technology serves an utterly human need to remember the dead. This viewpoint does accentuate the political economy aspect of both technological developments and attempts to communicate with the deceased—even though it is not furthered by Walter, leaving space to speculate whether it matters *how* exactly the political economy functions behind each attempt to communicate with the dead.

The aspect of political economy is essential to take into account when talking about deepfake.⁵ In the case of celebrities performing posthumously, this is especially evident. As Carl Öhman and Luciano Floridi argue in their analysis of the political economy of the digital afterlife industry, the industrial commercialization of the dead is not groundbreaking for the celebrity culture studies (Öhman & Floridi 2017). If resurrection by tech will bring disembodied immortality, it might very well be the immortality of endless posthumous labor. Öhman and Floridi set out to analytically structure an economic approach to the *digital afterlife industry*, understood as production of commercialized goods/services that involve online usage of digital remains, whereby “services that use personal data to generate new content replicating a dead person’s social behaviour” is one instance of the industry (ibid.). The authors resort to Marx’s notions of *living and dead labor*, further relying upon his concepts of *inorganic body* and its *alienation*, extending this metaphor to the *informational body*. Exploring economic mechanisms behind digital afterlife industry undertakings, Öhman and Floridi conclude that the discourse produced by enterprises commercializing death in the digital domain results from, not precedes, the mechanisms of death commercialization. The authors provide a valuable map of the economics of the digital afterlife industry. However, primarily ethical aspects of the digital afterlife occur in it, disclaiming the very possibility that commodification

⁵ Besides the fact that deepfake technology allowing for resurrection, such as Deep Nostalgia™, is commercialized, profiles of the deceased stay economically profitable for social media platforms (Meese et al., 2015). Resurrected users (would) provide even more profitability as they can be engaged in the mechanics of profit-making intrinsic for social media platforms.

discourse might steer practices of death commercialization and technologies used for it. I, on the contrary, set out to see *if* and *how* commercialization—or at least fantasies about endless performance and profit generation—is a part of the justifications of the development of deepfake technology.

Given how the deepfake (and deepfake resurrection, in particular) is reflected upon in academic debate, I argue that lack of criticism towards the ideological (technologically-determinist, techno-optimist, or capitalist, or all in one) foundations of this technology create boiler rooms where the limits of this technology are preconfigured. Only how to excavate these ideological premises of the technology?

1.2. Fantastic fantasies and where to find them

Formulated by Canadian philosopher Charles Taylor to explore social norms characterizing Western modernity, the concept of *social imaginaries* indicates a common understanding that potentiates common practices and a widely shared sense of legitimacy (Taylor 2004: 23). The concept shows productive for analyzing the mainsprings behind technologies and exploring what drives technological development (e.g., De Vries 2012; Flichy 2007; Mosco 2005; Marvin 1990). Defined as current and recurring ideas and practices constituted in relation to meanings and practices of the past (James 2019: 41), (socio-)technical imaginaries are dominant (rather than totalizing) stories and utopian hopes circulating in social groups. To be able to grasp particular instances of these imaginaries, I here resort to the concept of *fantasy*, defined, after Slavoj Žižek's interpretation of Jacques Lacan, as a matrix of coordinates for subject's desires (Žižek 2008: 45). As Stina Hasse Jørgensen and Oliver Tafdrup argue, fantasy is a theoretical concept sensitive to the political economy aspect of technologies (Jørgensen & Tafdrup 2017: 98–99).

Carolyn Marvin uses the term *media fantasies* in her interpretation of social imaginaries with specific regard to media technologies. She defines *media fantasies* as an essential cultural and social product legitimizing the development and appropriation of particular technologies and creating space for the making of new elites. Marvin asserts that media fantasies are grounded in perceived reality, reflecting cultural conditions and social stakes (Marvin 1990: 8–9). They both

determine and are determined by how technologies lend themselves as new media, which implies that new media embody a possibility of subversion of the status quo (ibid.), disregarding the efforts to contain or discipline it. Following this line of reasoning, media fantasies are a great source to tap into the stories systematically told within a given group as to what this group should be and how its technologies should work. If deepfakes are given shape reflecting (perceived) cultural conditions and social stakes of communities, fantasies about it should be telling of how the reality of this technology comes about, independent of a posteriori legislative and ethical considerations.

1.3. Enchantments of technology and imagined rationality

In his book *The Enchantments of Technology*, Lee Worth Bailey uses the term *enchantments*—socially-set “imaginative, mythic, and unconscious dimensions” having major collective implications (Bailey 2005: 3)—to refer to the foundations of technological culture. Bailey demonstrates how technical solutionism is rooted not primarily in reason, despite actively promoting rationality in collective discourses, but rather in technological enchantments that originate in the Cartesian distinctions of spirit/matter, mind/body, subject/object, and shape the development of technologies and fueling utopian programs. Reflecting upon Bruno Latour’s idea of humanity never having been modern, Bailey points to several umbrella technological enchantments⁶ to retrace how the seemingly rational position is based on myths about technology and desires for an almighty consciousness.

While *media fantasies* can be read as paving the way from a technology to a medium, *enchantments of technology* provide a glimpse of how technology itself comes to be. Situationally defining fantasies as cultural/social products legitimizing the development and appropriation of technologies and grounded in broader enchantments (myths about and desires for a technology) allows me to critically inquire how the phenomenon of deepfake is shaped. This, because the regard to

⁶ Such as *Bottomless subject* (presupposition that a subject, as opposed to objects, has unlimited power to comprehend and shape the world), *Absolute machine* (the yearning to make clever machines and in so doing to disentangle “the mystery of making life” (Bailey 2005: 155), and *Sublime speed* (“the lust for the highway rush” (ibid., 103).

how the limits of technology are envisioned exposes the social and cultural requirements to this technology (Jørgensen & Tafdrup 2017: 99), social stakes, and thus political and economic interests given in language when technologies are discussed.⁷

1.4. On being made of code, desire, sunlight, and ideology

As Lee Worth Bailey points out, technological enchantments are grounded in neatly separating spirit and matter, characteristic for some premises of Transhumanist thought.⁸ Resorting to a *posthumanist perspective* is one way to approach the enchantments of technology (and thus my research question) critically to examine fantasies of human exceptionalism. As technology becomes more ubiquitous, the ambiguity is expected to grow within the dichotomies of mind/body, real/unreal (Haraway 2006: 153), resulting from the ontological characteristics of *cyborgs*—assemblages of data, code, and computational power. This theoretical point of reference would suit perfectly for analyzing ontological characteristics of people resurrected by deepfake.⁹ It, however, falls beyond the scope of this thesis. From a posthumanist perspective, deepfake resurrection can be read to posit questions to the identity, personhood, and agency of humans and non-humans. Fascinating and well-researched,¹⁰ such questions would imply a high relevance of Actor-Network Theory to inquire into the agency of non-human actors.¹¹ I do not draw on this theoretical perspective because, methodologically, it would require creating a detailed descriptive account of an actant network, which does not seem a

⁷ I elaborate on this thought in Chapter 2.

⁸ For instance, the assumption that technology can be operationalized to exceed the limitations posed by the embodied reality of being human (Campbell 2006)—whereby death is a very pressing limitation.

⁹ The concept of *assemblage* (here, assemblage of data, code, computational power, and desires of those who enable the technology and create the actual deepfakes), would be extremely apt to conceive of whether the said ambiguity can be traced in the fantasies about the deepfake technology and to examine how the liberal humanist subject, criticized by N. Katherine Hayles (Hayles 1999: 288 and onwards), manifests itself in the desire to control death and the fantasies driving the development of technology enabling deepfake resurrection.

¹⁰ Several examples are: Beard 2001; Bassett 2015; Meese et al. 2015; Sherlock 2013.

¹¹ As done, for instance, by social anthropologist Alfred Gell in *Art and Agency: An Anthropological Theory* (Gell 1998).

realistic undertaking given that the subject of this thesis is the very possibility to resurrect people by deepfake, not a particular case of such resurrection.

Chapter 2. Methodology: *a blue sky and a full tank of postphenomenology*

In the upcoming section (2.1), I outline the methodological considerations of this thesis, situating discerning fantasies about deepfakes as a task informed by the postphenomenological approach based on analyzing accounts of deepfake mediation. I then specify how I shape the corpus of this research (2.2) and expound the research methods used to approach the corpus and answer the research question (2.3).

2.1. The meaning we make of things before they make us

The phenomenological approach provides that empirically accessing technology is inessential to perceiving it (Bucher 2017: 32), which often leads to abstracted, romanticized portrayals of technologies. In turn, the postphenomenological lens is empirically oriented, i.e., necessitating accounts of how the technology *mediates* human experiences and practices in the course of *intentional relations*. With its close attention to how technologies shape human experience, postphenomenological analysis allows for examining accounts of mediation in terms of epistemology, politics, and normative ideas (Rosenberger & Verbeek 2015: 31).

As language functions differently for groups of people depending on the ideological material that informs their judgments (Bernard-Donals 1994: 175), I turn to the publications upon deepfake and deepfake resurrection to distill the “meaning made of things” (Hasse 2018: 256). The postphenomenological optics helps me discover the ideological material in *how* such meaning is made without falling into “nothing-new-under-the-sun” explanations (Aagaard et al. 2018: xvii). My inquiry contributes to changing the ideological material—language—that constructs meaning (Bernard-Donals 1994: 183).

Excavating fantasies is, then, a largely metaphorical task— after all, we cannot measure or touch it, however precisely defined. The question arises: how to extract the fantasy, the imaginary, out of the technological thicket?

To facilitate her historical analysis of electric communication, Carolyn Marvin employs the concept of *textual communities* as theorized by Brian Stock to distinguish between insiders and outsiders as related to the technology. She delves into historical documents to analyze what media fantasies were prominent in shaping the discourse around electric communication called to maintain certain social stability in the course of messy appropriation of the new technology.

In turn, Lee Worth Bailey relies on Jungian epistemology in his exploration of technological enchantments that “reveal expressions of the collective unconscious” and “inhabit the machines we invent” (Bailey 2005: 6). While Bailey does not provide a direct methodological answer to extracting enchantments from discourse or practices around technology, his approach can be described as hermeneutics. In a circular motion, Bailey interprets the *expressions* and the *explications* of technological developments in a variety of sources (from films and advertisements to historical documents and existing research on technology).

Metaphorical excavation of fantasies is, thus, an interpretative endeavor. To establish a corpus for this research (i.e., determine what to interpret), I first set out to define who fantasizes about deepfake and its future. I then proceed to discuss the methods used to enable sound interpretations of this corpus.

2.2. Chasing authoritative texts beyond the “cool-or-creepy”

In terms of technical contributions, academics and engineers working in CGI, GAN development, and the film industry make deepfake resurrection possible as a curious by-product of technological development rather than its goal. Their commentary in technical publications is, thus, a major part of this research corpus. Technical papers likely disclose justifications and, accordingly, fantasies as to what new technology should do. I trace it where authors address possible applications of the technology and explain their benchmarks. To do so, I searched the ArXiv.org

e-print archive of scholarly articles and IEEE Xplore Digital library of engineering, computing, and technology-related publications, as these repositories feature peer-reviewed papers related to the fields of computer vision, CGI, and GAN development.

The criteria for the technical papers to be relevant are the following: a) discusses deepfake production from the technical viewpoint; b) addresses possible (mis)uses of technology; c) features (some) explanation of why/how the technology should develop; d) is published in English. In the course of an iterative search, supplemented by using the inbuilt search tools of the named platforms, I formulated the following queries (and their variations) for the technical articles: *deepfake*, *face AND swapping*, *generative adversarial network*, *face AND reenactment*.¹² These terms already prompt particular applications of the technology (e.g., reenactment). Seemingly functionally used in research titles, they imply how the technology is conceived of. For example, a search for technical publications by the term “deepfake” produces meager results. Apparently, the colloquial (vulgar?) *deepfake* is replaced by ostensibly more virtuous analogs such as *reanimation*, *facial expression transfer*, or *image-to-image translation*. These formulations could originate from the very genre of technical papers, requiring formal language use. The articles that do use the word *deepfake* are covering deepfake detection, not creation. On second thought, the functional epithets can be read as demonstrating how the technology is perceived by its makers: as a neutral tool for translating/transferring fairly abstract facial expressions; as a mathematical model to be improved, endlessly, without moral implications.

There are also lively discussions on deepfake in non-academic publications, including interviews with experts creating the software (e.g., Ajder 2019) and publications explaining the deepfake phenomenon to wider audiences (e.g., Roose 2019; Cole 2018). These instances constitute the second part of the research corpus. The criteria for such publications to be relevant are the following: a) specifically discusses applications of deepfake technology; b) is authored by developers

¹² The full selection is provided in Appendix 1.

(with(out) academic or industrial affiliations), “sellers,”¹³ journalists, policymakers, or artists reflecting *upon the uses* of the technology; c) is published in English. By performing the same procedure as when searching for technical articles, I list numerous terms to search for relevant publications. Some of them are: *deepfake AND resurrection OR reanimation; deepfake AND death; deepfake application; deepfake use; GAN AND death*. I manually select publications to make sure the content is not republished or rewritten.¹⁴

My search queries yield hundreds of relevant publications. Given the scope of this thesis and (possibly yet provisional) finiteness of human life, quantitative limits to the corpus are unavoidable. I somewhat arbitrarily limit the number of publications by around 25 instances per corpus part. I do not intend to subject the corpus to statistical analysis or draw generalizing conclusions as to what fantasies drive *all* the parties interested in developing the technology. Therefore, I select the publications arriving at the top of search queries with preference to those with high citation scores, supposing that high relevance to the search terms makes these texts authoritative in how the technology is imagined (Marvin 1990: 52).

As my preliminary research has shown, in non-technical publications, references to particular cultural texts recur.¹⁵ To contextualize my interpretations of fantasies, I draw upon these fictional texts as they can be very explicit about the limits of what is (in)conceivable and (un)acceptable. I, however, do not presume any of these texts being familiar to *all* those making deepfake possible.

These corpus considerations ascribe expertise to people mostly male, Western, and having enjoyed an academic education. In a way, my sampling illustrates the “artificial intelligence’s white guy problem” (Crawford 2016). In her opinion piece, Kate Crawford demonstrates how AI development is discussed from the position of the affluent, the privileged. A recitation from a random article on the dangers of deepfake confirms this statement: the main concerns address corporate and

¹³ This generic category designates business owners, marketers, and popularizers who monetize on technologies making deepfake resurrection possible.

¹⁴ The complete selection is provided in Appendix 2.

¹⁵ For instance, Mary Shelley’s *Frankenstein*, the episode “Be Right Back” of the *Black Mirror* series, Bram Stoker’s *Dracula*.

governmental security issues (cf. Westerlund 2019). The goal of my research is not to deepen this problem. Therefore, my corpus selection is not meant to be *universally* representative but to be representative of the desires of a group having the resources to develop the technologies enabling deepfake resurrection and whose desires, accordingly, drive this development. Just as the desires of Kris Kelvin in *Solaris* (1972) are not representative of the dreams of all people who have experienced loss, this research corpus is not representative of any “universal” desires.

2.3. Getting to the implied: method

2.3.a. Anything is a Foucault when you are brave enough

As argued in chapter 1, fantasies driving deepfake and deepfake resurrection can hardly be considered disregarding their political economy tinge. Given the corpus of this research, attention to the political economy aspect can be realized through performing Foucauldian discourse analysis (FDA). Not being a strictly defined combination of analytical practices, FDA is specifically attentive to the ways power relations and knowledge are explicated through language use. Within the FDA methodological framework, discourse is clearly differentiated from language (Sam 2019). The method is thus particularly suitable to inquire into ideological and symbolical characteristics of the research corpus without concentrating on its linguistic aspects, ensuring the feasibility of this research.

Operationalizing FDA means following a number of somewhat diffuse steps reflected upon in various disciplines (ibid.). In chapter 3 of this thesis, I employ the plan outlined by Terri Bourke and John Lidstone:

- crystalizing statements (units of discourse) in both parts of the corpus;
- setting the epistemological limits of my interpretation and hypothesizing how the statements are created with regard to the non-discursive context;
- hypothesizing what is “not-said” and why (Bourke & Lidstone 2015).

The research corpus could have been approached by conducting a semiotic analysis to consider how the text elements are likely to produce meaning (Fry 2009). This

would include blueprinting how denotations and connotations organize into a myth—seemingly “natural reality” (ibid.). Given the size of the research corpus, such analysis does not appear feasible. Furthermore, discovering the sense-impressions evoked by the texts (Berger 2016: 106) is beyond the scope of my inquiry.

2.3.b. Hovering between media theory and speculative work: digging

The method of Foucauldian discourse analysis allows me to answer the first sub-question of this research: what fantasies are voiced in the legitimations of deepfake technology development in technical papers and non-technical expert publications on deepfakes and deepfake resurrection. Informed by notions of media fantasies, enchantments of technology, and with specific interest towards the political economy of deepfakes, I use the results obtained by an FDA to hypothesize what interests and hopes are furthered in how deepfakes are conceived of and applied. However, answering the research question requires a deeper dig.

How can the fantasies extracted from the corpus be traced in political-economic configurations of the technology enabling deepfake resurrection? My interest in the fantasies *preceding* the very possibility of deepfake despite its readiness to be analyzed as a part of linear progress of deceiving media implies a media-archaeological approach. Practitioners of this approach share an interest in “alternate histories (...) that do not point teleologically to the present media-cultural condition as their “perfection” (Huhtamo & Parikka 2019: 3). Instead, media-archaeological optics seeks to discover how various and often forgotten pasts are reenacted in the present, to locate recurring tropes, uncanny feelings of déjà vu, and media-cultural ruptures (ibid., 14). While media-archaeology is a “traveling discipline” with few ready-made guidelines, it provides the ways of “reading” technologies by excavating the fantasies and aspirations realized in their design, use, and the ways they are discussed (Elsaesser 2016: 207). This helps conduct a “correction” (ibid., 183) of the teleological perspective, paying attention to how humans and machines coexist, how technologies shaped by humans, in turn, shape human experiences. Accordingly, the space for rethinking the future of technology might expand.

I, therefore, subject the statements obtained by an FDA to media-archaeological exploration (Sam 2019) in chapter 4 of this thesis to look at the seemingly evident origins of the technology. I do so through the media-archaeological method of *digging*. Digging works both metaphorically (“excavating” texts to trace paths and possibilities of configurations of technology) and as hands-on practice of “opening up historically constructed material reality” (Parikka 2018: 165–66). I here resort to the metaphoric component of digging, looking for the (dis)similarities and changes occurring in research and publications when the fantasies about deepfake resurrection are voiced, and drawing upon relevant academic discussions¹⁶ to situate similar fantasies in the field of historical media analysis. Because celebrity culture has a long history of death commercialization, I refer to its historical context to inform the digging.¹⁷

As to the hands-on component of digging, my initial plan was to operationalize the method of *critical making* and the theoretical premises behind it. The method emphasizes the shared act of making (Ratto 2011: 252) to increase engagement with theoretical and pragmatic issues. Critical making “frees” technologies from the state where they constitute the noisy background of social and cultural phenomena, allowing makers to “(re)evaluate the assumptions and values embedded into the designs” (Hertz 2017: 3). After thorough consideration, I refrain from it, despite the media-archaeological premises of my research forwarding research ‘close to the machine’ (Ernst 2011: 242). I assayed the critical making of a deepfake for another research project. It has provided me with exciting results in terms of the (im)possibilities of the technology and communicative component of this medium. Nevertheless, it has also shown that the method of critical making highlights the *wicked* (not yet even defined) *problems* in technology’s applications (Ratto, 2011), not necessarily answering questions about fantasies and aspirations in how the technology is imagined.

¹⁶ Such as those by Zoe Beloff (in Kluitenberg 2006: 214–40), Eric Kluitenberg (ibid., 157–84), and Edwin Carels (ibid., 186–214).

¹⁷ One such account is *Supernatural Entertainments: Victorian Spiritualism and the Rise of Modern Media Culture* by Communication and Media scholar Simone Natale.

Chapter 3. Who gazes, who controls the gaze?

The crucial question of discourse analysis informed by the work of Michel Foucault is how the objects of discourse might have been formed through what is made explicit and what is not (Graham 2005: 7). In this chapter, I conduct a Foucauldian discourse analysis (FDA) of technical research enabling deepfake resurrection (section 3.1) and expert publications (section 3.2), and provide some recapitulatory remarks (3.3). I do so in the search for statements that express the problems that the technology must solve (e.g., lack of speed and control in the production of synthetic/reanimated faces); the statements that set a *phantasmic scene*—a scene of fantasies through which the subject constitutes itself (Žižek 1997: 21)— and to discover for what kind of gaze this scene is set. I refer to Appendix 1 (designated with T. for technical research) and Appendix 2 (marked E. for expert publications), using index numbers to specify particular texts and letters in alphabetical order to indicate relevant passages.

3.1. Crossing the uncanny valley: harder, better, faster, ~~stronger~~ cheaper

As statements function in discourses to “enable objects to appear” (Graham 2005: 8), the question is then, what objects appear in the technical research allowing for deepfake production and thus deepfake resurrection. To start with (and quite contrary to my expectations), the applications of the technology are mentioned scarcely in the selected corpus of technical research. Only several publications are explicit about how the technology can be used, providing instances of video editing, film dubbing, social media content creation (T.21; T.22), creating virtual/physical avatars for virtual assistance, and other (not specified) face-to-face human-machine interactions (T.16). What is not mentioned are the more well-known-of applications of similar technologies such as revenge pornography or otherwise defamatory materials. Hypothesizing why these instances of image generation and GAN-training are not vocalized by technical researchers, I can locate an urge to clearly distinguish *neutral*, even *noble*, technology from its harmful or ethically-dubious appropriations. Through such formulations, authors position themselves at something that could be

read as “ethical height,” where the tools created are beyond doubt, having nothing to do with their messy real-life usage. Curiously enough, almost all models featured in the research I review are also reposted as open-source code on GitHub. On the one hand, open-sourced code is often viewed as inherently critical or thoughtful (Natalie Jeremijenko in conversation with Garnet Hertz in Hertz, 2015: 29), and making these models available for the broader public can be read as an open call to tinker with the technology. On the other hand, it renders the technology accessible for misuses the authors of the said models pronouncedly avoid pronouncing.

Another essential point here is that framing reasons for technology development as common-sensical and unquestionable (either by *not* mentioning use instances or by doing so lapidary, causally¹⁸) provides that the space shrinks where discussions could happen about the reasons for the very existence of this technology and what it has to mediate, how it has to structure our experiences. This collapsing space would otherwise allow for examining technologies’ meditative power in terms of epistemology and ideology.

I find the greatest number of recurring tropes in the parts of the research discussing the advantages of the technology as compared to other existing models (contribution sections and qualitative/quantitative evaluation sections). I thematically divide the frequent tropes into three clusters. The first one addresses the industry’s deficiencies and, thus, its goals. Authors define those by listing what their research contributes. Common here are the tropes of realism (e.g., “video-/photo-realistic animation”), naturalness, identity preservation (T.15),¹⁹ high-fidelity (T.14), convincingness, faithfulness of representation. The second cluster describes the deficiencies of the existing solutions in rather utilitarian, economic

¹⁸ It is worth hypothesizing that discussions about good and malicious usage of the deepfake technologies are happening elsewhere. How is the genre of technical research conceived by authors themselves, and what do they find the “right place” (if any) to discuss the ethics of the models they create? How might the delimitation of spheres of reflection work (mathematicians formulate exciting immaterial models, engineers tinker, humanities scholars reflect upon what is tinkered)? These apt questions could have been considered in the methodological part of this thesis, inevitably turning this analysis into something completely different.

¹⁹ Across all instances of the corpus, “identity” implies facial features in motion that look realistic to a human eye, e.g., T.4.e, T.8.a, T.12.b.

terms: needing a lot of calibration and human supervision (too much to stay usable on a large scale), slow, hard to use, lacking flexibility, and too expensive (T.11.a, d). The third cluster addressed the improvements and efficiency, building strongly upon how the problems in the industry are defined: “improving the state-of-art technology,” increased efficiency (e.g., T.3.d), producing better (even “superior”) results quicker (T.2.a), with less supervision, at reduced costs (T.4.h).

What is striking in these formulations is that people whose faces “need” to be “transferred”/“reanimated” are spoken of as objects, their “identities” reduced to coherent sets of facial features and expressions. Some of the research papers explicitly state that the unpredictable movements and individual mannerisms, overly detailed human hair and teeth are quite messy objects to work with (e.g., T.1.c). The implied solution to these inconveniences is eliminating unpredictability by training better, faster, cleverer models for “identity transfer.” This echoes yet another recurring trope I meet across all corpus instances: control. Several articles imply increased controllability of results by mentioning how their model facilitates “editing expressions at will” (e.g., T.22), while authors of others explicitly address the possibility to control facial expressions, artificial emotions of the “transferred identities,” and outputs in general (T.4; T.6; T.11). The seeming utility of the terms used to describe how the technology should be improved reminds of Slavoj Žižek’s assertion that ideology is at work where it “innocently” refers to the utility, while the claim of utility is reflective, used to signalize certain attitudes (Žižek 1997: 2). In the case of technology enabling deepfakes, I can hypothesize the attitude signalized here is what Lee Worth Bailey calls “rational, objective mastery” (Bailey 2005: 12), indicative of a utopian program yearning for ever-increasing autonomy for the rational subject and including that the surrounding object-world is a field of subject’s experiment and exploitation.

Curious are the references to the uncanniness of the results of image generation and reanimation, including mentions of the *uncanny valley effect* (e.g., T.1; T.4; T.24). Authors of technical research describe the uncanniness as “division between a synthetic-looking face and a real, animated, expressive person” (T.1). Ideally, escaping the uncanniness means producing a “real, emotionally compelling actor” (T.1). Partly,

this task is perplexed by the limitations of the existing models of deepfake production, but the uncanniness is also blamed upon how humans *learn* to distinguish between “generated” and “real” images (T.19.c). While learning quickly is an advantage in GANs, in humans, the very learning *feature* kills all fun and leads to the uncanniness of what is otherwise meant as *realistic*. While one research in my selection explicitly states the need to make sure humans can distinguish the generated images from the real ones (T.11.h), others implicitly present this skill as an obstacle to technology’s excellence (e.g., T.4.c, T.19.c).

Analyzing the technical research corpus, I notice an implicit opposition that can be addressed as “damned if we don’t.” If the technology is improved in terms of quality, realism, and cost-effectiveness, we harvest more efficient ways to exercise our rational mastery. If it does not, synthetic media will stay unimpressive; we can altogether abandon the dream of a realistic virtual assistance that looks like a breathing, feeling human being. What is worse, we then would not ascend to the utopian program of finding out who humans are in the mechanical universe (Bailey 2005: 158) by modeling the Pinnochios we could then (dis)assemble. The discourse sampled in my selection of technical research positions the technology allowing for deepfake resurrection as necessary for the sake of progress, objective and rational technological mastery, and informed control. Simultaneously, this discourse draws a clear line between expert and amateur applications of the technology, attributing “useful” applications to the technological elites (who are reaching for the progress) and the “low-culture” dubious applications to amateurs tinkering practices.

3.2. We’ve lost control (again, or maybe never had it in the first place)

One of the most prominent objects brought to a field of exteriority (Foucault 1972, cited in Graham 2005: 8) by the technical research enabling deepfake is the inevitable and necessary progress of synthetic media. The expert publications, however, en masse assume that this object might be closer than it appears in technical research. The progress of synthetic media in terms of speed and accessibility to non-professional audiences forms a starting point for the conversation. It then brings to light many other objects: the future of deepfake(able)

objects, digitally-resurrected people, and, surprisingly, the future of nation-states. In the analysis of the deepfake futures, celebrities (including performers, CEOs, politicians), the dead, and women play an exemplary role as people likely to “experience” the technology first.

Unlike technical research texts, non-technical expert publications are saturated with mentions of possible applications of deepfakes. At the same time, the realism or faithfulness of the generated images and “identity transfers” are scarcely talked about. Experts attribute the deluding capacity of the technology to the speed of its development rather than the realism of the end products. The fast-paced maturing of the tool is framed as a matter of fact²⁰ (Matt Ratto’s reading of Bruno Latour’s concept, Ratto 2011: 259), and authors actively voice the need to adapt to it, i.e., to “maximize the use (...) while being careful not to abuse others” (E.5).

The notion of speed plays an essential role in all the sampled publications. “We”—apparently, societies as a whole—need to hurry to anticipate where the technology will go next, prepare for it. The field of legislation is mentioned frequently. However, more curious are the reasons for such preparations that authors refer to, rooted in an image of a societal hierarchy given in the texts:

- “Simple users,” “regular folks” (respectively, E.7, E.12), unprepared to the deceptive powers of synthetic media. These people need to be educated on the subject.
- “Amateur users,” apparently, non-morals who make unacceptable deepfakes such as nonconsensual pornography. These users are “assholes” (E.19) jeopardizing the technology (E.14), “carnival barkers”²¹ (E.18). This last metaphor is particularly gaudy, as it refers very clearly to the imagined confrontation of the “low” and the “high” culture.
- Expert elites (trained academics and engineers) at the gates of using deepfakes exclusively for problem-solving. These people are described as a

²⁰ In the analyzed publications, experts speak about deepfake and its advancement either in impersonal forms or use “deepfake/technology” as a subject with “develops/advances” as a predicate.

²¹ Full quote: “The carnival barker gets an algorithmic boost on services like Facebook and YouTube, while the expert screams into the void.”

precious yet “outgunned” minority (E.16). Only one text (E.18) refers to this group in terms of relation to wealth instead of mastery and knowledge, even though its author still seems surprised that technical knowledge is not protected from “amateurs.”

The consequence of sketching such social structure is that while “technical elites” describe themselves as outgunned by persistent amateurs,²² broader elites (i.e., CEOs, politicians, celebrities) are astonished to find themselves among the first “preys” of the synthetic media revolution.²³ Beforehand secured from the perils of being replaced by algorithms, they are now the first ones to be “reanimated” due to the excessive availability of their images in open access. This way, for the expert publications, the cases of deepfake resurrection are not the cases worth pondering as such. They are a stepping stone to start a conversation about larger societal effects of the Fourth Industrial Revolution, and bring the discussion to what *really* causes a certain degree of moral panic. Where particular examples of deepfake resurrection (such as that of Salvador Dali) are generally framed as neutral (E.3.d, e) or positive (E.20.a, c, E.24.e) because of their economic effects, fantasizing about the consequences of deepfake reanimation or pornography applied to politicians and CEOs is consequently used as a warning example of technology’s disruptive power (E.8, E.10). Here, the (imaginary) dangers of deepfakes are transposed onto a larger societal fabric. One text (E.10.e) is particularly explicit about it: “distorting democratic discourse; manipulating elections; eroding trust in institutions; weakening journalism; exacerbating social divisions; undermining public safety; and inflicting hard-to-repair damage on the reputation of prominent individuals, including elected officials and candidates for office.” The underlying fantasy is that with progressing synthetic media, realities become uncurated, “anyone [is] able to (...) present any reality they want” (E.14.d), diminishing the “ability for a large community to agree on

²² At this point, I find myself in a struggle with imaginaries, too. I am inclined to refer to the image sketched of the amateurs with a Russian colloquial expression “*without a Tzar in one’s head*,” designating simultaneously a degree of *craziness* and a lack of respect to institutional norms. However, it has no direct equivalents in English, and I roughly and colloquially translate it here as “with no (moral) captain, no (normative) rudder.”

²³ In one of the texts from my selection (E.8), an analogy is presented to the industrial revolution’s replacement of certain types of workers by machines. The synthetic media revolution would replace other types of people by synthesized images and data assemblages in the context of labor and leisure.

what is true” (E.10.h). The concern about the governability of large communities is brought to a field of exteriority very prominently in such statements.

In the majority of texts I analyze, experts seem uncertain about the positive applications of deepfakes. Higher personalization and a boost of creativity are often mentioned (e.g., E.20), along with claims about “better engagement” and “accessibility of solutions” (E.21).²⁴ Essentially, deepfakes are expected to “lessen barriers in communication” (E.24), echoing the dream of seamless communication in a mythological golden age before “the Lord scattered them abroad from thence” (Genesis 11:8 KJV). The expectation that, despite all odds, “deepfakes can give people a voice, purpose, and ability to impact scale and speed” (E.21.h) also reminds of what Lee Worth Bailey calls the enchantment of Sublime Speed, originating from an “illusion of rationalism that it can control the passions driving its speeding technologies” (Bailey 2005: 87).

However, not only the utopia of streamlined communication justifies the development of deepfakes. Authors of the publications rather frame the situation as not having any options to stop the technology without crossing ideological lines of democracy (E.14). While deepfakes are framed as a force of distortion of democracies and nation-states, the impossibility to unplug them is emphasized repeatedly. The potential modifications of the technology’s design are not mentioned at all, and the accountability for what the technology does “in the wild” is generally placed upon anonymous “assholes.” Deepfake might endanger the elites and whole social structures—but the carnival barkers, not the experts, are responsible.

3.3. Reconciling the found and the envisioned

Quite contrary to my expectations, I do not discover any preoccupation with the horrors and delights of otherwise ungraspable technological immortality in how the experts write about deepfake reanimation. It is the fantasies about governability of

²⁴ It is further not specified in the cited text engagement with or solutions to what exactly.

populations, specifically, in Western countries²⁵ and on terms of liberal democracies that fascinate them. I formed the corpus of expert publications to trace discussions concerning death and deepfake reanimation, using the search terms that would address these issues: *deepfake AND resurrection OR reanimation; deepfake AND death; GAN AND death*. Because I searched for authoritative texts reflecting upon the use of deepfakes, I have then settled for the most quoted, most mentioned, and best-search-engine-optimized texts.

Despite digging for death and reanimation with my queries, I thus have harvested texts captivated not with the death and immortality of individuals but rather the death and mortality of social groups. The big cliffhanger, then, is how to reconcile this finding with my research question to proceed? First of all, I must partly attribute the encounter with the layer of discourse concerning the *Untergang* of Western democracies to the very character of the texts I gathered. Being mostly published in large news outlets with specific interests in the economy (Forbes), social, political, and economic current affairs (BBC, The Guardian), and technological developments (Vice, Wired), the instances of my corpus can hardly be expected to analyze the deepfake phenomenon on the micro-level. Instead, larger societal processes constitute their main subject. On the one hand, this might be why the experts discuss the future of democracies and nation-states instead of the possibilities to employ deepfake-resurrected people for movie performance or grief processing. On the other hand, this explanation falls short given that the texts in the selection are authoritative, thus, tone-setting in how the technology is and must be comprehended. The authoritative texts analyzed by Carolyn Marvin conveyed how electricity (provided that used wisely and under “expert” control) would render cultural differences meaningless, making politics (understood as a struggle for scarce resources) abundant (Marvin 1990: 206—207). The authoritative texts I snatched use the case of resurrecting dead people as a starting point for a

²⁵ This derives at least partly from the fact that I analyze English-language publications only. Notwithstanding, if I were to play a drinking game, I would get very drunk very quickly by taking a sip every time China or Russia are mentioned in the publications in the context of jeopardizing Western democracies. Another way to get very drunk would be drinking every time institutionalized racism in Western societies is talked about as if it were completely hypothetical; as if the only way to get footage of officials making racist remarks were deepfakes.

conversation that is consistently framed as more important: that about the politics that are anything but abundant. Carolyn Marvin asserts that the texts marking the appropriation of electricity gently removed the contradiction between electric communication's potential to threaten social boundaries and to provide a future of cognitive imperialism where differences are finally obsolete. The texts marking the appropriation of deepfakes constitute the following "sensible" trade-off: deepfake resurrection of individuals might be unsettling, just as deepfake pornography can be devastating, but we must concentrate on bigger things—our market demands (after all, greater personalization is an undoubted value) and our political life. Not said is that deepfake highlights the contradiction between these bigger things. In terms of presenting a particular view of the technology and preparing the ground for the practices deriving from it (Graham 2005: 10), I conclude that expert texts in my research corpus privilege the view that individual deepfake immortality and its forms and consequences are out of the field of what might matter. Relating to my research question, this is highly anticlimactic. However, I do not take it as a sign of lacking methodological rigor, rather, as an excavation exercise yielding an encounter with a different treasure than expected.

The scenes in both technical and expert publications are set for the eyes of the "elites," not the "regular folks." The authors of technical papers fantasize how their formulas will acquire the expert community a better grip on the "identities," making synthetic media faster, cheaper, less uncanny. The aspect of limitless mastery powered by reason and progress is apparent in these texts. Following Lee Worth Bailey's program for discovering the enchantments of technology, I hypothesize sacrifices required to reach these utopian ideals through deepfake technology (Bailey 2005: 6). There might be occasional casualties: victims of deepfake pornography, performers potentially losing jobs due to deepfakes performing better, or oddities of grieving facilitated by deepfake-resurrected people—seemingly, all nominal if the enchantment-driven stakes are as "high" as finding out who humans are in the mechanical universe (ibid.: 158).

The non-technical publications reflecting upon the applicability of the formulas anticipate something different: a future where governability is in jeopardy due to the

nonchalant tinkering of anonymous rebels of the Internet. The frequently emphasized ethical ambiguity of this tinkering serves as a backdrop to ponder how the status quo can be restored. The utopian program of deepfakes “for a better world” is thus obfuscated by the interference of carnival barkers.

What, then, happens to the utilitarian, pragmatic language of technical articles²⁶ that accentuate the control the deepfakes must bring? How does it turn into lamenting the loss of control over universal rational truth, deceptive artifacts “tearing the fabric of democratic societies” (E.19.c), “wreaking havoc” (E.23.a)? Makers reflect on their formulas; later, other experts reflect on the possible consequences of how these formulas become technologies. Somewhere between, an act of translation occurs—and dry utilitarianism of formulas transforms into a warm yeast dough of speculation. Apparently, different types of control are negotiated in these fantasies: the technical elites speak enthusiastically about control in terms of production, while the experts zoom onto control in terms of governability. The same applies to the trope of realism. In technical research, the realism of produced images is the necessary condition for deepfakes to thrive in the market. For non-technical experts, the realism of the synthesized is the greatest dread.²⁷ This is where the scenes diverge, despite being set for virtually the same viewers and being a product of very similar discursive practices.

The political economy of these scenes—the lack of reflectivity towards the very coordinates of democracies leaning upon neoliberalism—is not as outrageous as the fact that it is served on a silver platter of elitist arrogance. Going back to the Žižekian interpretation of fantasy that not only “conceals the horror of the Real” but (re)creates what it aims to conceal (Žižek 1997: 6—7) and extending the metaphor of a strawberry cake, how do experts know they desire better governability in Western societies and more control over the realities presented in the media? This is what the

²⁶ The very genre of a technical paper provides, as it were, no space for fantasizing, though I do, in turn, fantasize about how the authors *might have* fantasized and then cut the phantasmagoric out of their texts for those to fit the genre.

²⁷ As many experts imply in their publications (e.g., E.2.a, E.8.d, E.9.a), this dread hardly fits neat mathematical models and is measured in more ephemeral units: believability and attractiveness of synthesized images, or already existing political sympathies of the viewers (e.g., E.19.a).

deepfake fantasies seem to tell. In these fantasies, the elites are unequivocally assigned the subject's position, their desires being a command for everyone else—which refers to the utopian program of rational mastery once again.

Chapter 4. Choose your unanswerable questions

In this chapter, I trace—fragmentary and bitty as they come—(dis)continuities that become apparent when the phantasmic scenes discovered in chapter 3 are held against the backdrop of scenes that emerged around the media in the past. By operationalizing the media-archaeological method of digging, I first conduct metaphorical excavation of how the fantasies voiced in the corpus of this thesis are reminiscent of fantasies already discussed in the field of historical media analysis (4.1). I further explore ideological entanglements of deepfakes (4.2) and hypothesize how the fantasies can be traced in the stabilities, marginalized and dominant, of deepfake.

4.1. Where deepfakes encounter death of social groups

I resort to a metaphorical excavation of the discourse around deepfake technology, intending to locate discursive constructions that might guide the existing figuration and reception (Boddy 1994: 106–7) of deepfakes. There are similarities between how deepfakes are spoken about in terms of disrupting trust and wreaking havoc on political reality (E.23.a) and how other media—now almost historic—were discussed upon emergence. Digging up these similarities evokes a sense of *déjà vu*, shows how certain fantasies and desires cyclically reappear as “imprinted” on various historical media technologies (Huhtamo 1997). Critical examination of such recurrence allows eschewing the technical-determinist idea of constant progress (*ibid.*), which might not fully disenchant technologies, but remove at least some tension brought about by the notions of novelty and infinite advancement.

The notion of multistability, holding that technologies can be put to multiple purposes, being variously meaningful to different users (Rosenberger & Verbeek 2015:

31), is key to my analysis of the deepfake phenomenon. Unlike a hammer, deepfake hardly has a dominant stability so far. From the FDA conducted, I can conclude that to technical researchers, the dominant stability of deepfake is (re)animating images of people, existent or otherwise, for the reasons of increasing mastery (e.g., T.1.a, b, T.4.b, T.6.c, e). According to the non-technical experts, the dominant stability might be distorting the public trust in primarily political institutions (e.g., E.5.g, E.6.a, E.10.d). All other possible applications of the technology (i.a., non-consensual pornography production, grief processing assistance, disrupting performers' careers) are marginalized by both groups of experts as negligible collateral damage on the way to progress. Applying the postphenomenological lens to the deepfake, how these marginalized stabilities are reflected upon (or not) in the corpus can be telling of how, through imagining deepfake's (im)possibilities, particular interests, desires, and beliefs might be furthered.

The utopian program deepfakes offer is almost a caricature of how Lee Worth Bailey describes the enchantments of technology: limitless mastery of a rational subject, sublime speed of progress, a possibility of seamless communication, and an opportunity to finally discover what being human means in a mechanical universe. The statements pointing towards this program I acquaint when conducting an FDA of technical and expert publications resemble those heard when electricity (Marvin 1990, Kluitenberg 2006, Carels in Kluitenberg 2006: 186–214), photography and motion picture (Beloff in Kluitenberg 2006: 214–40), radio, television, VR (Boddy 1994), or mobile telephony (De Vries 2012) were the new big (communication) thing. Deepfakes will possibly let us achieve some perfected state of communication, just as telegraph, radio, and VR were supposed to.

Why, then, would anyone expect deepfakes (or any new technology) to do what other technologies were expected but did not manage to deliver before? Applying the theoretical lens provided by Slavoj Žižek in *The Plague of Fantasies* (Žižek 1997: 39), in an ideological space, the fantasies about deepfake are radically ambiguous, closing the span of choices available (if we as a society desire this—mastery, rationality—we cannot choose to “undo” deepfakes) and maintaining the idea that

these excluded choices (“undoing” deepfakes) might have happened. This could be the reason why the apparentness of subject/object ontology fueling the deepfakes’ utopia does not light the way out of the mirror maze of fantasies about absolute mastery. Fantasizing this utopia already allows dealing with some ontological lack (Jørgensen & Tafdrup 2017: 99) by letting groups, in this case, identify as rational enough to see the innovative potential of technologies through having technological fantasies as such. In turn, this identification invites more technological fantasies, preventing social groups with particular imaginaries from shutting down deepfakes—or virtually any other technology—even when dangers seem to outweigh benefits.

The ambivalent nature of fantasies works on both micro- and macro-levels. At the level of individual usage, despite experts discussing how uncanny and potentially dangerous deepfake resurrection is, services like Deep Nostalgia™ by MyHeritage gain popularity because their claim of innovation²⁸ realizes and recreates technological desire. At the macro-level, despite threatening governability of social groups, deepfakes promise to innovate and streamline at least certain types of communication, production, and labor, which is attractive because it allows for both identification (“we are not some Luddites who would waive a technology offering this much”) and producing more technological fantasies (“to make this technology work for the good, we need other technologies to tame it”).

Paradoxically, no utopian program comes without a dystopian counterpart (De Vries 2012: 52). “Some fucked-up dystopia”²⁹ driven by deepfakes includes a complete rupture of realistic representations, which will lead to the death of trustworthiness, allow “anyone (...) to (...) present any reality they want” (E.14.d), and disrupt the governability of communities (E.19.d). That very idea of realistic representations being fundamental for social groups to find common understanding echoes the XIX

²⁸ An opportunity of interacting with the dead (an “old” need that had to be solved with many “old” technologies) in a new, better, surprising way: “Animate the faces in your family photos with amazing technology. Experience your family history like never before!” (“MyHeritage Deep Nostalgia™, deep learning technology to animate the faces in still family photos,” MyHeritage. Accessed on April 15, 2021, through <https://www.myheritage.com/deep-nostalgia>).

²⁹ As the deepfake of Barack Obama puts it in the BuzzFeed video “You Won’t Believe What Obama Says In This Video! 😊” uploaded on YouTube on April 17, 2018.

century's turn to document-based historiography. Wolfgang Ernst contextualizes the shift from describing to showing as a result of the new optical media development, for instance, photography (Ernst 2005: 592). The dystopian view on deepfakes seems drenched with the same yearning to dispose of unreliable narratives and tell stories of what groups are/must be by means of evidence and authenticity provided by *reliable* machines. The utopia of deepfakes being the next big communication thing shatters against the same problem all other big things did—the yearning for diversity and unrestricted access invites lack of consensus and “righteousness” (De Vries 2012: 51). Hence the expert publications in the corpus of this research overemphasizing how the consensus and righteousness need to be restored/kept by disciplining the deluded object: technological means of making deepfakes more controllable, legislative efforts, and digital literacy initiatives that would anew teach “the masses” what is real and what is not (T.11.i; E.11.d; E.14.j). Hence the protrusive (re)construction of boundaries between the elites, the amateurs, and the “regular” recipients of the synthesized media in the corpus of this thesis. As Carolyn Marvin demonstrates with her analysis of electricity, the spaces where new media are made are the spaces where new elites form and ward their position (Marvin 1990: 253), which I trace in concerns (voiced by non-technical experts primarily) about undesirable changes in social relationships. These concerns can be read as re-actualizing “old” ideas about what the *proper* relationships should be even in the future transformed by innovative technologies.

In one of the corpus' publications, deepfakes are compared to the printing press (E.14.h) as able to enhance creativity and human expression. Deepfakes are expected to “give people a voice, purpose, and ability to impact scale and speed” (E.21.i). Yet, considering the applications of deepfakes technical and non-technical experts recount, the purpose, impact, and creativity amount to better, faster, more attractive consumption. Beyond consumption, “the masses” are portrayed as passive recipients of whatever (image) comes their way (e.g., E.10.d). Even though this view holds that “regular folks” are spectators (Green 2010) in the first place, discursively, they are still stripped of all sovereignty as well as the ability to meaningfully distinguish between realistic versions of events and non-realistic ones.

What interests are forwarded in this protrusive framing the masses as passive and the technology as dangerous yet irresistible? Despite the promises of seamless communication and unlike, for example, broadcasting media (Boddy: 109), deepfakes are not envisioned as a tool to integrate communities. Even if there were a linear progression of deceptive technologies, deepfakes' hyper-personalization would differentiate them drastically from broadcasting media or photography (provided that those are seen as a way to monologize the "truth" in some way). Repetition of the possible dystopian consequences of deepfakes can be read not as a warning furthering interests of larger communities staying governable but rather as a fantasy concealing what Slavoj Žižek calls *the horror of the Real*. Even a frightening dystopian narrative rearranges new phenomena in some bearable temporal sequence (Žižek 1997: 11), in this case, of the inevitable progress of synthetic media. It might not be a calming narrative per se, but its familiarity is calming. Claim that the reality will change so much that "regular people" could not distinguish between real and synthesized anymore—does it not let readers get used to the very idea a bit? Does it not reconcile readers with the future, however ungraspable?

Because "undoing" deepfakes is impossible—it would contradict the needs of the market and ideology—the need to restrict its use is immediately denounced, conforming to the coordinates set by the fantasies about the technology and preparing space for more fantasies.³⁰ This ambiguity shows the bare minimum required by the utopian program. Deepfakes should, at least, produce us better, more personalized content at lower prices, despite the often pronounced ethical doubtfulness of the tool. Emphasizing the doubtfulness might obfuscate how the technology cannot be reshaped precisely because it is shaped to comply with the market demands. Following Žižek's line of reasoning (Žižek 1997: 27), by pointing out the ethical pitfalls, experts create a space for the ideology of demanding self-regulating markets to "seize" the technology by explicating the awareness that beneath the ideological mask, there are unquestionable values of diversity, solidarity, empowerment, democracy. In the name of these values, non-technical experts

³⁰ Similarly, when interviewed for journalistic publications, technical experts lament the technology being available to virtually anyone (e.g., E.3.c, E.5.c, d) but make sure dissemination of the tool (through, for example, open-sourcing the code) is unobstructed in their technical papers (e.g., T.22.a).

announce the need to restore monologizing control with an eye on unpredictable, hostile forces, including amateurs (whose applications of deepfakes are deemed both anti-elitist and immoral) and Chinese/Russian/generic³¹ propaganda machines (E.19.c, E.12.c, e). While the only way to deal with deepfakes is to discipline ourselves to live with them, the struggle around trust and reliability of information is telling of how experts envision the meta-group whose interests are paramount: societies of primarily Western countries, especially from the point of view of their political contracts. Once again, just as with telegraph, radio, or VR, this group is endangered. But let me leave this ideological stance alone for a moment.

4.2. One cannot abuse quantified identities—deepfake machine goes brrrr

Looking for the “alternate histories” (Huhtamo & Parikka 2019: 3) of the media marketed as innovative, media-archaeological approaches allow to discern how forgotten utopian programs of the past are being reenacted in the present and thus “correct” (Elsaesser 2016: 207) overly-enthusiastic or purely dystopian teleological views of technologies. Excavating phantasmic origins of the deepfakes allows to detach them from their seemingly evident predecessors—other “deceptive” media—and thus examine the fantasies that might hide behind the linear myth of technologically-driven progressing disruption in the representations of reality.

An essential aspect of how deepfakes are made is that the working of generative adversarial networks—a form of artificial intelligence—is based on quantifying (emotional) facial expressions and bodily movements, further operationalizing faces and bodies as numbers. In the case of facial reanimation, the notion of control does not only include “trapping the human body and holding it prisoner by having the software record its movements” (Mihailova 2013: 138) but also the possibility to

³¹ I call it “generic” here because expert publications call out China and Russia as possibly willing to disrupt the Western societies often enough. However, there are also more subtle references to the disruptive use of deepfakes for unstabilizing societies. For instance, a case of hypothesizing about “someone fabricat[ing] a deepfake-assisted fraudulent incident that is primed to touch a sensitive nerve such as racism or the intense dislike that members of the public have toward a prominent political figure” (E.12.e, also E.12.d). Speculating the way this particular expert sees the “someone” is not particularly useful. Therefore, I leave the portrait of forces that, according to the non-technical experts, might want to weaponize deepfakes somewhat blurry.

manipulate these quantified identities. Viewing identities as numerical (and thus computable, predictable) (T.4.h, T.8.a) renders them more usable for the subject's needs. The requirement of usable, pliable objects is less apparent in deepfake production than it is in, say, robotics, yet it raises a number of questions as to how deepfakes realize certain epistemological programs historically intrinsic to AI development. For instance, as early as 1996, Alison Adam argues that “history of AI is a history of building AI systems that model the Cartesian ideal of transcendent male reason” (Adam 1996: 52), resulting in what Francesca Ferrando recently addressed as “male legacy, social exclusivism, and biological essentialism” silently informing AI, carrying old and introducing new forms of discrimination (Ferrando 2014: 2:43). While I have extensively addressed the fantasy of technological control over quantified identities in chapter 3, here, I would like to focus on the types of objects quantified and manipulated in the technical research part of the corpus to inform the *digging*. I thus argue that the seemingly evident origins of deepfakes—being an extension of easily forgeable media used to deceive the public—are not that evident at all and that using this lens for analyzing the phenomenon obfuscates many essential aspects of how deepfakes are shaped and shape human experiences.

Viewing deepfakes as an instance of AI and analyzing its development not as compared to photography and video production but as another prophecy of “cybernetic supremacy” (Barbrook in Kluitenberg 2006: 267) allows highlighting different aspects than technological desires for streamlined, always reliable communication. In several technical publications, the importance of developing technologies allowing for deepfake creation is bound to the idea of developing appealing, almost-human virtual assistants (e.g., T.16.b, T.21.a, b). In this light, it is interesting to examine what objects are used to serve this goal. Are there types of faces and identities deemed unserviceable as potential virtual assistants? To answer this question, I visit the datasets utilized in my selection of technical research. Scarce are quantifications of black people; Hispanic and Asian faces are also infrequent, Slavs are represented by Vladimir Putin alone. Strikingly more female faces are quantified to learn AI about what a close-to-life virtual assistant might be; celebrities (including politicians) constitute a large number of faces in the datasets, too. Obtaining a dataset of celebrities' faces is relatively easy because of their publicity.

However, the datasets inevitably bear the biases of their makers (Alberts-de Gier 2019) as to who is worth quantification or might serve as a virtual assistant. For instance, CelebFaces Attributes Dataset (CelebA) is a dataset widely used for facial reanimation (e.g., T.24, variation of the dataset in T.20). CelebA is often criticized (e.g., in Xu et al. 2020) as “extremely biased to good-looking white people.”³² The list of annotations used in the dataset (ibid.) features terms such as “male” (in the absence of “female,” while 61.4% of the subjects are female (ibid., 12), “attractive,” and “young.” Some of the terms are practical to guide the researchers training GANs (e.g., “mustache”). It is, though, unclear how the process of quantification should improve by using the category “attractive.” My conclusion is that training GANs on datasets primarily featuring “attractive” caucasian female faces reinforces biases already discovered in AI technologies (Ferrando 2014). By training GANs to quantify, recognize, and reproduce “attractive” caucasian- and female-looking faces, the idea is bolstered of how a perfect virtual assistant or artificial being must look, express emotions, and communicate, and this idea is Western-centric and exclusive. If a GAN trained on CelebA dataset had to generate artificial performers for a movie, it would not produce a *Time of the Gypsies* or an *Andrei Rublev* (even in terms of variety of emotional expression shown, other artistic qualities aside). Just as Google’s Deep Dream is known for generating surprisingly many dogs in every picture because trained on an extensive dataset of images of canines (Alexeev 2018), a CelebA-trained GAN would not produce any varied set of artificial beings.

Another striking aspect of training generative networks on pictures of females and celebrities is rendering these faces serviceable. Celebrities, as, for instance, Sharon Marcus conveys in her book *The Drama of Celebrity* (2019) or as Simone Natale demonstrates in his research of Victorian spiritualism (Natale 2016), are not fully agents of how they perform. Their performance and public image are for a large part driven by the expectations and internal desires of their fans. Also essential here is that mediatizing (Western) celebrities and disseminating their images through historic media played a crucial role in the global expansion of Western culture (ibid., 83)—and can play it again through deepfake virtual assistants. Male politicians in the

³² Jolicoeur-Martineau, Alexia. Twitter Post. June 20, 2020. https://twitter.com/jm_alexia/status/1274447446760927232?s=20.

datasets might seem a curious exception in this case, yet they are probably not. Politicians—specifically, heads of states, often used in the corpus for testing the quantification algorithms—historically serve as a white screen for collective expectations and projections and are, thus, too, serviceable.³³

Bringing these observations about the datasets used for GANs training in the technical research part of this thesis' corpus back to the domain of deepfake resurrection, I must add that the bias should inevitably influence how the faces of the deceased are reanimated. Time-, gender-, social position- and even individually-specific mimics, mannerisms, and emotive facial expressions are disregarded by quantifying and reanimating a face as if it was the same as those of the people used to train a GAN.

The positive mentions of streamlined communication and connectivity (E.7.e, E.20.a, c, E.21.g) used by non-technical experts can be both an enchantment of technology speaking through the technocratic discourse and the technocratic discourse once again employing the imaginary of streamlined communication. It probably is both. But beyond that déjà vu of the enchantment of technology, it can also be read as an instance of discursively making deepfakes (and thus AI) emotionally compelling in order for the capital to instrumentalize and monetize emotional responses (Stark & Crawford 2015: 8). Artificial beings in general are often presented in emotional terms (Selinger 2021): this is why the monster of *Frankenstein*, a synthetically created emotional being, is often used as a metaphor for AI. First of all, I must note that this widely disseminated reading of *Frankenstein* might not be the most relevant for understanding artificial intelligence. For instance, through the lens of feminist critique, the monster in the novel can be read as a metaphor for women in Victorian society, bereft of a possibility to realize “self-assertion and social acceptance” (Long Hoeveler 2003: 48, 52). Through the Marxist optics, *Frankenstein* can be read as an

³³ This quite psycho-historical idea is developed by, i.a., Lloyd deMause (e.g., in DeMause & Ebel 1977), though, there is research on projections of particular values and expectations upon political leaders without the contested discipline of Psychohistory as well, such as *The Royal Touch: Sacred Monarchy and Scrofula in England and France* by Marc Bloch (Bloch 2015 (1924)).

allegory of worker's alienation of the products of her labor, whereby Victor Frankenstein represents a worker entangled in a production process, and the Creature externalizes, as it were, Victor's alienation (Michie 1990). Repetition of the reading of the Creature as a metaphor for an emotional, relatable AI-being closes the span for discussing how we might conceive of artificial intelligence beyond the idea of the subject's (Victor's) urge to create life and object's (the Creature's) subsequent frustration and rage. Secondly, employing the metaphor of an artificial being that is able to experience and evoke emotions serves the opposite to a warning end: it nudges the readers to think of AI-generated beings as emotional or at least necessarily evoking strong emotions. The Deep Nostalgia™ example is again curious here. By reanimating the dear deceased of its users, it once again adds (positive) emotional value to the technology that otherwise has none. Does this feel like abuse of actual people behind the images for profit generation? Maybe. Is it deemed important by the experts of technical and non-technical publications? No.

GANs are hardly imagined by their makers within a binary gendered framework. I do not find any instances of gendering generative networks in the corpus. On the contrary, the virtual assistants that GANs must produce mostly are gendered, as the tasks they perform reinforce assumptions about femininity and feminine labor (Costa & Ribas 2019: 177–78). In this light, GANs can be read as a technological extension of the *Bottomless subject* (researcher, maker), while the deepfakes produced—along with people's quantified identities employed for production—are the docile objects. In my optics, this is why when some real-life applications of deepfake (pornography or resurrection) are explicitly disapproved of by experts, a few sentences later, this disapproval is denounced, and the focus shifts to deepfake as a perfect tool for greater personalization. The very statement that greater personalization is unconditionally positive already subtly suggests that hyper-personalized pornography or performance (of dead or living actresses) is an attractive opportunity.

When abuse of individual quantified identities does not matter much when there are profits to make, discussions unfolding in the non-technical publications as to values of democracy, trust, and protection of societal stability seem quite far-fetched.

The deepfake machine goes brrrr not necessarily as Chinese propagandists or immoral carnival barkers make disruptive deepfakes of political leaders, but rather because this technology is designed to regard anything and anyone as a means to the ends of a *Bottomless subject*. As one of the non-technical experts (E.24.i) puts it, the solution to the deepfake dilemma might be to deal with it the same way as we “dealt with other industries that externalize large costs to society while privately claiming the profits (such as industrial pollution, big tobacco, and fast food)”—except for we did not.

Conclusion

Fascinated by the examples of applying deepfake technology for resurrection, I have formulated the research question of this thesis the following way: *What fantasies can be found in justifications of the development of the technology allowing for deepfake resurrection, and how can they be traced in the workings of the technology?* I expected to write about possibly immortal people of the future, someone imagined as a sum of a *Black Mirror* episode, the Creature from *Frankenstein*, a Dracula who does not die, and a perfect virtual assistant such as *Her*.

This study aimed to analyze the phenomenon of deepfake through accounts of its mediation, i.e., technical papers and non-technical expert publications reflecting upon how GANs and similar technologies can be and are put to work. I aspired to conduct an analysis avoiding the premises of technological determinism, the notion of linear progress, and viewing deepfakes as a technological *matter of fact*. Therefore, I operationalized concepts of *(socio-)technical imaginary*, *(media) fantasies*, and *enchantments of technology* to render visible what cultural conditions and social stakes are reflected, what epistemological and ontological positions are furthered through imagining deepfakes as it is done in the corpus of this thesis.

To discover the fantasies behind deepfakes, I subjected the corpus to Foucauldian discourse analysis, distilling statements that enable discursive objects to appear and

steer how these objects are/must be understood. The application of an FDA proved helpful for finding out that deepfake resurrection is used as an entry to a conversation about governability and survival of Western societies in the light of synthetic media development. The method allowed me to conclude that the texts in my research corpus disregard individual deepfake immortality, its mechanics, and consequences; and examine the ideological aspects of deepfake development.

Why is individual deepfake resurrection framed as unimportant? In a measure, this results from the character of the corpus of research I have formed. Publications in large news outlets (e.g., Forbes, The Guardian, Vice) analyze large societal processes more often than the micro-level of technological adaptation. Nonetheless, I chose these publications because they are authoritative: setting coordinates for how the technology is/must be comprehended. I took the fact that these texts do not speak extensively about deepfake resurrection as an instance of what is *not-said* in discourse and thus influences the objects brought to the field of exteriority. Another question raised in the course of my FDA is: have I not overlooked discursive sites where deepfake resurrection is discussed in more detail? Transhumanist publications, for example, marketing materials about and interfaces³⁴ of programs allowing deepfake resurrection, would be these locations. I did not research them due to my concern with the authoritative texts. I list this observation as a limitation to my research and a possible direction for further inquiries. Conducting a discourse analysis of software communities as socio-technical systems would be helpful to address the image of a societal hierarchy I discovered by the method of FDA and see how open-sourcing the code and affordances of, i.a., GitHub or Google Collab play a role in deepfakes becoming a collaborative technology.

Operationalizing the method of digging and resorting to only its metaphorical component was a daring endeavor. The method having few ready-made guidelines was both liberating and challenging, and at some points, tracing the genealogical and phantasmatic kinships of deepfakes felt speculative. For instance, analyzing the serviceability aspect of female and celebrity faces in the GANs training datasets required more in-depth knowledge of celebrity culture and gendered software than I

³⁴ Or lack thereof if running code from one of the many GitHub repositories.

possess. The resulting bit of analysis is flimsy, and I am only glad that this flimsiness does not prevent me from answering my research question. Moreover, it turned out that the teleological perspective—at least at the level of being inclined to find relations of determination—is not easy to eschew even when one is guided by methodological premises unwelcoming a determinist point of view. I had to repeatedly remind myself that inquiring into (socio-)technical imaginary is a way to *enter* the analysis of the social meaning (James 2019: 44) that cannot unequivocally point out relations of (non)determination in how technologies are made and that it instead helps to critically access fantasies, ideologies, ontological and epistemological premises of them. Hopefully, this leads to changes in ideological material that constructs the social meaning of the technology in question.

Through the media-archaeological method of digging, I traced how the fantasies around deepfakes resemble and diverge from the fantasies discussed in the field of historical media analysis and described some ideological entanglements of the technology. I posed and answered several questions about the technology. First of all, what kind of mastery, seamless communication, “ability to impact scale and speed” (E.21.i) is imagined as the future of deepfakes—that where access to faster, personalized consumption is a priority along with the need to be disciplined as the boundary between realistic and unrealistic representations becomes blurry. The second question concerns how all instances of the corpus only ponder the future of Western societies, using the unquestionable values of diversity, empowerment, and democracy as the reason to both develop the technology and restore elites’ monologizing control over it. The third question addresses the fantasy of technological control upon quantified identities and subsequent quantification of only *certain types of identities* in the technical research part of the thesis’ corpus. Using one of the datasets frequented by technical experts training GANs as an example, I show how Western-centric and exclusive ideas of who and for what reasons should be quantified and manipulated reinforces in the case of deepfakes the same biases already discovered in other AI technologies. By operationalizing the method of digging, I tried to revise the teleological view on the deepfakes being a part of a linear progression of deceptive media and highlight its kinship with the utopian programs of artificial intelligence.

Generally, I can conclude that the ambiguous working of fantasies about deepfakes in the ideological space closes the span of choices available to the Western societies (desiring limitless mastery, one cannot shut down a new technology), maintaining the idea that such a shutdown was possible, and paving the way for new technological desires. At the same time, the protrusive trope of the inevitable progress of synthetic media creates a sense of familiarity, concealing the horror of an uncontrollable technological future. Imprinted in the sense of familiarity are ethnocentric and sexist biases that I find in the datasets for GANs training and in how the potential of deepfakes is conceived of. The conflicts addressed in the existing academic debate (deepfakes distorting trust, devastating political relations vs. streamlining communication, bringing new ways to generate profits) cannot be resolved legislatively or by ethical reflection as long as the technology is grounded in fantasizing anything and anyone as a means to the ends of a *Bottomless subject*. Bringing the deepfake instance to the discussion about the fantasies and desires that drive technological development, this thesis contributes to discussing technologies as they are (pre)configured in terms of political economy, hopefully reopening a span where these configurations can be altered.

In the course of this research, in spite of searching explicitly for publications that deal with deepfake resurrection, I shifted my focus from the issues of individual death and immortality to that of social groups in order to examine the phantasmic scene the corpus of this research revealed. One of the directions for future research would be, then, to examine how particular cases of deepfake resurrection function and how they are discussed. Such inquiries would profit from (n)ethnographic methods, forms of discourse analysis, discursive interface analysis, and the walkthrough method for studying software allowing for deepfake resurrection.

Another direction for future inquiries might be analyzing larger amounts of evaluations of the technology by methods of distant reading of deepfake-related discussions on fora or social media to extract fantasies and expectations of the “regular folks.” Working with this data might yield a more complex understanding of

how accurate the framing of “the masses” as easily deceived and misled by deepfakes (as asserted by non-technical experts) is.

Yet another question raised in this thesis and requiring a more thorough examination would relate to imagining the future of deepfakes on male and Western terms of viewing technology. The corpus of this thesis was limited to English-language publications for reasons of feasibility. Further exploring the bias of the Western gaze in the reflection upon the future of deepfakes and ways the technology works would contribute to deconstructing sexist and neo-colonialist optics that are still guiding the development of artificial intelligence.

References

- Aagaard, Jesper, Jan-Kyrre Berg Olsen, Jessica Sorenson, Oliver Tafdrup, Cathrine Hasse, and Don Ihde, eds. 2018. *Postphenomenological Methodologies: New Ways in Mediating Techno-Human Relationships. Postphenomenology and the Philosophy of Technology*. Lanham, Maryland: Lexington Books.
- Adam, Alison. 1996. "Constructions of gender in the history of artificial intelligence." *IEEE Annals of the History of Computing* 18 (3): 47–53. Accessed on April 14, 2021, through <https://www.computer.org/csdl/magazine/an/1996/03/00511944/13rRUxlqxOI>.
- Ajder, Henry. June 16, 2019. "The ethics of deepfakes aren't always black and white." *The next web*. Accessed on December 10, 2020, through <https://thenextweb.com/podium/2019/06/16/the-ethics-of-deepfakes-arent-always-black-and-white/>.
- Alberts-de Gier, Berend. July 4, 2019. "Data are always already biased: The datafication framework." *Datafication of experience*. Accessed on April 14, 2021, through <https://medium.com/datafication-of-experience/data-are-always-already-biased-the-datafication-framework-b56ab12168af>.
- Alexeev, Vladimir. December 6, 2018. "Deep Dream comes true." *Vlad Alex (Merzmensch)*. Accessed on March 10, 2021, through <https://medium.com/merzazine/deep-dream-comes-true-eafb97df6cc5>.
- Bassett, Debra. 2015. "Who Wants to Live Forever? Living, Dying and Grieving in Our Digital Society." *Social Sciences* 4 (4): 1127–39. <https://doi.org/10.3390/socsci4041127>.
- Beard, Joseph J. 2001. "Clones, Bones And Twilight Zones: Protecting the Digital Persona of the Quick, the Dead and the Imaginary." *Berkeley Technology Law Journal* 16, no. 3 (2001): 1165–271. Accessed on December 20, 2020, through <http://www.jstor.org/stable/24116971>.

- Berger, Arthur Asa. 2016. *Media and Communication Research Methods: An Introduction to Qualitative and Quantitative Approaches*. Fourth Edition. Los Angeles: SAGE.
- Bernard-Donals, Michael. 1994. "Mikhail Bakhtin: Between Phenomenology and Marxism." *College English* 56 (2): 170–88. <https://doi.org/10.2307/378728>.
- Bloch, Marc. 2015. *The Royal Touch: Sacred Monarchy and Scrofula in England and France*.
- Boddy, William. 1994. "Archeologies of Electronic Vision and the Gendered Spectator." *Screen* 35 (2): 105–122.
<http://screen.oxfordjournals.org.proxy.library.uu.nl/content/35/2/105>.
- Bourke, Terri, and John Lidstone. 2015. "What Is Plan B? Using Foucault's Archaeology to Enhance Policy Analysis." *Discourse: Studies in the Cultural Politics of Education* 36 (6): 833–53.
<https://doi.org/10.1080/01596306.2014.903611>.
- Brown, Nina I. 2020. "Deepfakes and the Weaponization of Information." *Virginia Journal Of Law & Technology* 23 (1). Accessed on December 20, 2020, through
https://static1.squarespace.com/static/5e793709295d7b60295b2d29/t/5f0b7e4e0f4f095f7e70613f/1594588750542/v23i1_1-Brown.pdf.
- Brubaker, Jed R., and Gillian R. Hayes. 2011. "'We Will Never Forget You [Online]': An Empirical Investigation of Post-Mortem Myspace Comments." *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work—CSCW '11*, 123. Hangzhou, China: ACM Press.
<https://doi.org/10.1145/1958824.1958843>.
- Bucher, Taina. 2017. "The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms." *Information, Communication & Society* 20 (1): 30–44.
<https://doi.org/10.1080/1369118X.2016.1154086>.
- Campbell, Heidi A. 2006. "Postcyborg Ethics: A New Way to Speak of Technology." *Explorations in Media Ecology* 5 (4): 279–96.
https://doi.org/10.1386/eme.5.4.279_1.

- Celona, Luigi, Simone Bianco, and Raimondo Schettini. 2018. "Fine-Grained Face Annotation Using Deep Multi-Task CNN." *Sensors* 18 (8): 2666.
<https://doi.org/10.3390/s18082666>.
- Cole, Samantha. January 24, 2018. "We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now." *Vice*. Accessed on January 22, 2021, through <https://www.vice.com/en/article/biye8a/reddit-fake-porn-app-daisy-ri-dley>.
- Costa, Pedro, and Luísa Ribas. 2019. "AI Becomes Her: Discussing Gender and Artificial Intelligence." *Technoetic Arts* 17 (1): 171–93.
https://doi.org/10.1386/tear_00014_1.
- Crawford, Kate. June 25, 2016. "Artificial Intelligence's White Guy Problem." *The New York Times*. Accessed on December 15, 2020, through <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html>.
- Deepware. July 23, 2020. "Deepfakes Can Put Flesh on the Bones of Artificial Consciousness." Accessed on February 3, 2021, through <https://deepware.ai/deepfakes-can-put-flesh-on-the-bones-of-artificial-consciousness/>.
- DeMause, Lloyd, and Henry Ebel, eds. 1977. *Jimmy Carter and American Fantasy: Psychohistorical Explorations*. New York: Two Continents.
- Dimitropoulos, Stav. October 28, 2020. "A new film with a dead star? Thanks to CGI and deepfake technology, it's possible James Dean 2.0 may be coming to a theater near you." *Experience*. Accessed on December 20, 2020, through <https://expmag.com/2020/10/a-new-film-with-a-dead-star-thanks-to-cgi-and-deepfake-technology-its-possible/>.
- Elsaesser, Thomas. 2016. "Media Archaeology as Symptom." *New Review of Film and Television Studies* 14 (2): 181–215.
<https://doi.org/10.1080/17400309.2016.1146858>.
- Ernst, Wolfgang. 2005. "Let There Be Irony: Cultural History And Media Archaeology In Parallel Lines." *Art History* 28 (5): 582–603.
<https://doi.org/10.1111/j.1467-8365.2005.00479.x>.

- Ferrando, Francesca. 2014. "Is the Post-Human a Post-Woman? Cyborgs, Robots, Artificial Intelligence and the Futures of Gender: A Case Study." *European Journal of Futures Research* 2 (1): 43.
<https://doi.org/10.1007/s40309-014-0043-8>.
- Fried, Ohad, Ayush Tewari, Michael Zollhöfer, Adam Finkelstein, Eli Shechtman, Dan B. Goldman, Kyle Genova, Zeyu Jin, Christian Theobalt, and Maneesh Agrawala. 2019. "Text-Based Editing of Talking-Head Video." ArXiv:1906.01524 [Cs], June. <http://arxiv.org/abs/1906.01524>.
- Fry, Paul. 2009. "Semiotics and Structuralism." *YaleCourses*, September 1, 2009. YouTube video, 51:27. <https://youtu.be/VsMfalOsT3M/>.
- Gell, Alfred. 1998. *Art and Agency: An Anthropological Theory*. Oxford; New York: Clarendon Press.
- Goddard, Michael. 2015. "Opening up the Black Boxes: Media Archaeology, 'Anarchaeology' and Media Materiality." *New Media & Society* 17 (11): 1761–76.
<https://doi.org/10.1177/1461444814532193>.
- Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. "Generative Adversarial Networks." ArXiv:1406.2661 [Cs, Stat], June.
<http://arxiv.org/abs/1406.2661>.
- Graham, Linda J. 2011. "The Product of Text and 'Other' Statements: Discourse Analysis and the Critical Use of Foucault." *Educational Philosophy and Theory* 43 (6): 663–74. <https://doi.org/10.1111/j.1469-5812.2010.00698.x>.
- Green, Jeffrey E. 2010. *The Eyes of the People: Democracy in an Age of Spectatorship*. Oxford: Oxford Univ. Press.
- Hayles, N. Katherine. 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, Ill: University of Chicago Press.
- Hertz, Garnet, ed. 2015. *Conversations in Critical Making*. Victoria, BC: CTheory Books.

- Hoskins, Andrew, ed. 2017. *Digital Memory Studies: Media Pasts in Transition*. 1st ed. New York : Routledge, 2017.: Routledge.
<https://doi.org/10.4324/9781315637235>.
- Huhtamo, Erkki. 1997. "From Kaleidoscomaniac to Cybernerd: Towards an Archeology of the Media." *Leonardo* 30(3): 221–24. Accessed on April 7, 2021, through
http://gebseng.com/media_archeology/reading_materials/Erkki_Huhtamo-Kaleidoscomaniac_to_Cybernerd.pdf.
- Huhtamo, Erkki, and Jussi Parikka, eds. 2019. "1. Introduction: An Archaeology of Media Archaeology." In *Media Archaeology*, 1–22. University of California Press. <https://doi.org/10.1525/9780520948518-003>.
- Jørgensen, Stina Marie Hasse & Tafdrup, Oliver Alexander. 2017. "Technological fantasies of Nao: remarks about alterity relations." *Transformations*. 2017; Vol. 29. 89–102. Accessed on January 8, 2021, through
http://www.transformationsjournal.org/wp-content/uploads/2017/02/Transformations29_Jorgensen-Tafdrup.pdf.
- Kluitenberg, Eric, ed. 2006. *The book of imaginary media: excavating the dream of the ultimate communication medium*. Rotterdam: De Balie.
- Kneese, Tamara. January 27, 2021. "How a Dead Professor Is Teaching a University Art History Class." *Slate*. Accessed on January 28, 2021, through
<https://slate.com/technology/2021/01/dead-professor-teaching-online-class.html>.
- Kwok, Andrei O. J., and Sharon G. M. Koh. 2020. "Deepfake: A Social Construction of Technology Perspective." *Current Issues in Tourism*, March. 1–5.
<https://doi.org/10.1080/13683500.2020.1738357>.
- Leaver, Tama. 2019. "Posthumous Performance and Digital Resurrection: From Science Fiction to Startups." Preprint. SocArXiv.
<https://doi.org/10.31235/osf.io/2gkw9>.
- Lemke, Jay L. 1990. "Technical Discourse and Technocratic Ideology." In *Learning, Keeping and Using Language*, M.A.K. Halliday, John Gibbons, and Howard

Nicholas, eds., 435. Amsterdam: John Benjamins Publishing Company.
<https://doi.org/10.1075/z.lkul2.31lem>.

Linder, Courtney. January 13, 2021. "After You Die, Microsoft Wants to Resurrect You as a Chatbot." *Popular mechanics*. Accessed on January 18, 2021, through
<https://www.popularmechanics.com/technology/robots/a35165370/microsoft-resurrects-the-dead-chatbots/>.

Long Hoeveler, Diane. 2003. "Frankenstein, Feminism, and Literary Theory." In *The Cambridge Companion to Mary Shelley*, edited by Esther Schor, 1st ed., 45–62. Cambridge University Press.
<https://doi.org/10.1017/CCOL0521809843.004>.

Marvin, Carolyn. 1990. *When Old Technologies Were New: Thinking about Electric Communication in the Late Nineteenth Century*. Nachdr. III.

Marcus, Sharon. 2019. *The Drama of Celebrity*. Princeton: Princeton University Press.

Meese, James, Bjorn Nansen, Tamara Kohn, Michael Arnold, and Martin Gibbs. 2015. "Posthumous Personhood and the Affordances of Digital Media." *Mortality* 20 (4): 408–20. <https://doi.org/10.1080/13576275.2015.1083724>.

Meese, James, Bjorn Nansen, Tamara Kohn, Michael Arnold, and Martin Gibbs. 2015. "Posthumous Personhood and the Affordances of Digital Media." *Mortality* 20 (4): 408–20. <https://doi.org/10.1080/13576275.2015.1083724>.

Michie, Elsie B. 1990. "Frankenstein and Marx's Theories of Alienated Labor." In *Approaches to Teaching Mary Shelley's "Frankenstein,"* edited by Stephen C. Behrendt, 93–8. New York: MLA, 1990. Accessed on April 5, 2021, through
<http://knarf.english.upenn.edu/Articles/michie1.html>.

Mihailova, Mihaela. 2013. "The Mastery Machine: Digital Animation and Fantasies of Control." *Animation* 8 (2): 131–48. <https://doi.org/10.1177/1746847713485833>.

"MyHeritage Deep Nostalgia™, deep learning technology to animate the faces in still family photos," *MyHeritage*. Accessed on April 15, 2021, through
<https://www.myheritage.com/deep-nostalgia>.

- Natale, Simone. 2017. *Supernatural Entertainments: Victorian Spiritualism and the Rise of Modern Media Culture*. Place of publication not identified: PENN State UNIV Press.
- Öhman, Carl, and Luciano Floridi. 2017. "The Political Economy of Death in the Age of Information: A Critical Approach to the Digital Afterlife Industry." *Minds and Machines* 27 (4): 639–62. <https://doi.org/10.1007/s11023-017-9445-2>.
- Packer, Jeremy. 2010. "What Is an Archive?: An Apparatus Model for Communications and Media History." *The Communication Review* 13 (1): 88–104. <https://doi.org/10.1080/10714420903558720>.
- Parikka, Jussi. 2018. "Digging." In *Routledge Handbook of Interdisciplinary Research Methods*, edited by Celia Lury, 164–168. London ; New York: Routledge, Taylor & Francis Group, 2018.
- Poulopoulos, Dimitris. April 2, 2020. "How to Produce a DeepFake Video in 5 Minutes." *Towards data science*. Accessed on January 4, 2021, through <https://towardsdatascience.com/how-to-produce-a-deepfake-video-in-5-minutes-513984fd24b6>.
- Ratto, Matt. 2011. "Critical Making: Conceptual and Material Studies in Technology and Social Life." *The Information Society* 27 (4): 252–60. <https://doi.org/10.1080/01972243.2011.583819>.
- Roose, Kevin. March 4, 2018. "Here Come the Fake Videos, Too." *The New York Times*. ISSN 0362-4331. Accessed on January 22, 2021, through <https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html>.
- Rosenberger, Robert, and Peter-Paul Verbeek. 2015. "A Field Guide to Postphenomenology." In *Postphenomenological Investigations: Essays on Human-Technology Relations*, edited by Robert Rosenberger and Peter-Paul Verbeek, 9–41. London: Lexington Books, 2015.
- Sam, Cecile H. 2019. "Shaping Discourse Through Social Media: Using Foucauldian Discourse Analysis to Explore the Narratives That Influence Educational Policy." *American Behavioral Scientist* 63 (3): 333–50. <https://doi.org/10.1177/0002764218820565>.

- Savin-Baden, Maggi, and Victoria Mason-Robbie, eds. 2020. *Digital Afterlife: Death Matters in a Digital Age*. 1st ed. Chapman and Hall/CRC.
<https://doi.org/10.1201/9780429322198>.
- Selinger, Evan. April 6, 2021. "A.I. Can't Detect Our Emotions." *OneZero*. Accessed on April 10, 2021, through <https://onezero.medium.com/a-i-cant-detect-our-emotions-3c1f6fce2539>.
- Sherlock, Alexandra. 2013. "Larger Than Life: Digital Resurrection and the Re-Enchantment of Society." *The Information Society* 29 (3): 164–76.
<https://doi.org/10.1080/01972243.2013.777302>.
- Siarohin, Aliaksandr, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, and Nicu Sebe. 2020. "First Order Motion Model for Image Animation." ArXiv:2003.00196 [Cs], October. <http://arxiv.org/abs/2003.00196>.
- Singer, David, and Camila Connolly. September 7, 2019. "How Hollywood Can (and Can't) Fight Back Against Deepfake Videos (Guest Column)." *The Hollywood Reporter*. Accessed on December 20, 2020, through <https://www.hollywoodreporter.com/thr-esq/how-hollywood-can-can-t-fight-back-deepfake-videos-guest-column-1237685>.
- Stark, Luke, and Kate Crawford. 2015. "The Conservatism of Emoji: Work, Affect, and Communication." *Social Media + Society* 1 (2): 1–11.
<https://doi.org/10.1177/2056305115604853>.
- Taylor, Charles. 2004. *Modern Social Imaginaries*. Public Planet Books. Durham: Duke University Press.
- Thurlow, David A., and Ben D. Sawyer. "Who Wants an Automated Vehicle?" In *Critical Issues Impacting Science, Technology, Society (STS), and Our Future*, edited by Heather Christina Lum, 171–196. IGI Global, 2019.
<https://doi.org/10.4018/978-1-5225-7949-6>.
- Walter, Tony. 2015. "Communication Media and the Dead: From the Stone Age to Facebook." *Mortality* 20 (3): 215–32.
<https://doi.org/10.1080/13576275.2014.993598>.

- Westerlund, Mika. 2019. "The Emergence of Deepfake Technology: A Review." *Technology Innovation Management Review* 9 (11): 39–52. <https://doi.org/10.22215/timreview/1282>.
- Whittaker, Lucas, Tim C. Kietzmann, Jan Kietzmann, and Amir Dabirian. 2020. "‘All Around Me Are Synthetic Faces’: The Mad World of AI-Generated Media." *IT Professional* 22 (5): 90–99. <https://doi.org/10.1109/MITP.2020.2985492>.
- Whyte, Christopher. 2020. "Deepfake News: AI-Enabled Disinformation as a Multi-Level Public Policy Challenge." *Journal of Cyber Policy* 5 (2): 199–217. <https://doi.org/10.1080/23738871.2020.1797135>.
- Xu, Tian, Jennifer White, Sinan Kalkan, and Hatice Gunes. 2020. "Investigating Bias and Fairness in Facial Expression Recognition." *ArXiv:2007.10075 [Cs]*, August. <http://arxiv.org/abs/2007.10075>.
- Žižek, Slavoj. 1997. *The Plague of Fantasies*. London and New York: Verso.
- Žižek, Slavoj. 2008. *The Sublime Object of Ideology*. London: Verso.

Appendix 1. Corpus of technical research publications

For quick references, this appendix contains relevant extractions from the titles analyzed (authors' punctuation and spelling are preserved). The clusters are color-coded:

- Hypothetical applications of technology—pale pink
- Realism of the end product (naturalness, high-fidelity, convincingsness, etc.) as the end-goal of the authors—light green
- Deficiencies of the existing solutions—light red berry
- Improvements and enhancements proposed by authors, building upon the definitions of industry's deficiencies, including streamlining, improvement, development, making technology easier to use faster—light cornflower blue
- Controllability of results, excitement of mastery, knowledge accumulation, and pleasure of heuristics—light yellow
- Framing uncanniness of the produced images—orange
- Actual functioning of the technology—light magenta
- Contextual definition of identity in the context of image generation—cyan
- Ease of adaptation of a proposed solution, democratization of technology, including open-sourcing the code—grey.

Several remarkable metaphors that do not form a recurring cluster are highlighted in magenta.

1	Alexander, Oleg, Mike Rogers, William Lambeth, Jen-Yuan Chiang, Wan-Chun Ma, Chuan-Chang Wang, and Paul Debevec. 2010. "The Digital Emily Project: Achieving a Photorealistic Digital Actor." <i>IEEE Computer Graphics and Applications</i> 30 (4): 20–31. https://doi.org/10.1109/MCG.2010.65 .
	<p>(a) The project aimed to cross the "uncanny valley" that divides a synthetic-looking face from a real, animated, expressive person. The key technologies included a fast high-resolution digital face-scanning process using USC ICT's Light Stage capture system and Image Metrics' video- based facial-animation system.</p> <p>(b) From this result, it's clear that a tractable process can achieve the goal of an emotionally compelling, photorealistic digital actor. What remains to be done</p>

	<p>is to streamline the techniques and to further explore the uses of such technologies in applications such as film production, interactive entertainment, simulation, and education.</p> <p>(c) Although accurate and detailed, the scans of Emily's expressions required preprocessing before we could construct appropriate blendshapes. For instance, many meshes had irregular boundaries with poor triangulation and mesh artifacts around the teeth and eye regions. Also, many scans contained surface regions not represented in the master mesh, such as the surfaces of the eyes and teeth. With some effort, we manually cleaned up such problem areas in mesh-editing software. (...) In future work, we would like to streamline this process for creating photorealistic digital characters. We spent a great deal of our effort forming correspondences between the scans, and the semiautomatic techniques weren't quite accurate to the level of skin pores. Having algorithms to automatically determine consistent surface coordinate correspondences across the scans—accurate to skin-pore detail—would improve the process significantly.</p>
2	<p>Bińkowski, Mikołaj, Danica J. Sutherland, Michael Arbel, and Arthur Gretton. 2021. "Demystifying MMD GANs." <i>ArXiv:1801.01401 [Cs, Stat]</i>, January. http://arxiv.org/abs/1801.01401.</p>
	<p>(a) [key contribution by authors is] resulting in a simpler and faster-training algorithm with matching performance. We also propose an improved measure of GAN convergence, the Kernel Inception Distance, and show how to use it to dynamically adapt learning rates during GAN training.</p> <p>(b) Generative Adversarial Networks (GANs; Goodfellow et al., 2014) provide a powerful method for general-purpose generative modeling of datasets.</p>
3	<p>Brock, Andrew, Jeff Donahue, and Karen Simonyan. 2019. "Large Scale GAN Training for High Fidelity Natural Image Synthesis." <i>ArXiv:1809.11096 [Cs, Stat]</i>, February. http://arxiv.org/abs/1809.11096.</p>
	<p>(a) Despite recent progress in generative image modeling, successfully generating high-resolution, diverse samples from complex datasets such as ImageNet remains an elusive goal.</p> <p>(b) We find that applying orthogonal regularization to the generator renders it amenable to a simple "truncation trick," allowing fine control over the trade-off between sample fidelity and variety by reducing the variance of the Generator's input. Our modifications lead to models which set the new state of the art in class-conditional image synthesis.</p> <p>(c) The state of generative image modeling has advanced dramatically in recent years, with Generative Adversarial Networks (GANs, Goodfellow et al. (2014)) at</p>

	<p>the forefront of efforts to generate high-fidelity, diverse images with models learned directly from data.</p> <p>(d) We introduce two simple, general architectural changes that improve scalability, and modify a regularization scheme to improve conditioning, demonstrably boosting performance.</p> <p>(e) We discover instabilities specific to large scale GANs, and characterize them empirically. Leveraging insights from this analysis, we demonstrate that a combination of novel and existing techniques can reduce these instabilities, but complete training stability can only be achieved at a dramatic cost to performance.</p> <p>(f) Code and weights for our pretrained generators are publicly available.</p> <p>(g) As a result, our models set a new level of performance among ImageNet GAN models, improving on the state of the art by a large margin.</p>
4	<p>Doukas, Michail Christos, Mohammad Rami Koujan, Viktoriia Sharmanska, and Anastasios Roussos. 2020. "Head2Head++: Deep Facial Attributes Re-Targeting." <i>ArXiv:2006.10199 [Cs, Eess]</i>, June. http://arxiv.org/abs/2006.10199.</p>
	<p>(a) We leverage the 3D geometry of faces and Generative Adversarial Networks (GANs) to design a novel deep learning architecture for the task of facial and head reenactment.</p> <p>(b) We conduct a comprehensive set of quantitative and qualitative tests and demonstrate experimentally that our proposed method can successfully transfer facial expressions, head pose and eye gaze from a source video to a target subject, in a photo-realistic and faithful fashion, better than other state-of-the-art methods. Most importantly, our system performs end-to-end reenactment in nearly real-time speed (18 fps).</p> <p>(c) Video editing, film dubbing, social media content creation, teleconference and virtual assistance are some indicative examples. However, generating artificial human faces indistinguishable from real ones is a very challenging task, particularly when it comes to video data. Adding the extra dimension of time, might give rise to the so-called problem of temporal incoherence. As the uncanny valley effect suggests, people are extremely perceptible in unnatural facial and head movements, thus even small discontinuities can expose a synthetic video.</p> <p>(d) Given that face reenactment systems offer no control over the target's head pose and eye gaze, there might be cases where the person's expressions do not match with the overall head movement and therefore seem unnatural.</p> <p>(e)</p>

	<p>A more holistic approach on reenactment involves generating all pixels within frames, including the upper body, hair and background of the target identity. In contrast to video manipulation techniques, such as face reenactment, full head reenactment methods [4], [9] aim to transfer the entire head motion from a source identity to a target one, along with the eye blinking and gaze, providing complete control over the target subject.</p> <p>(f) [State-of-art] methods do not provide complete control over the generated video, as they manipulate only the interior of the face.</p> <p>(g) Our recorded results indicate a human fake detection accuracy of 43.1%, which demonstrates the strong photo-realism of samples created by our <i>Head2Head++</i> system.</p> <p>(h) Furthermore, qualitative and quantitative comparisons suggest that <i>Head2Head++</i> outperforms other state-of-the-art methods, in terms of photo-realism, expression and pose transferability, as well as target identity preservation. For future work, we aim to make the network trainable in a few-shot manner, reducing the training time and eliminating the need for learning a different model per person. Additionally, we plan to transfer the facial expressions from the source in a way that preserves the target's speaking style.</p>
5	<p>Ekmen, Beste, and Hazım Kemal Ekenel. 2019. "From 2D to 3D Real-Time Expression Transfer for Facial Animation." <i>Multimedia Tools and Applications</i> 78 (9): 12519–35. https://doi.org/10.1007/s11042-018-6785-8.</p>
	<p>(a) In this paper, we present a three-stage approach, which creates realistic facial animations by tracking expressions of a human face in 2D and transferring them to a human-like 3D model in real-time.</p> <p>(b) The tracking is performed using a single camera to enable several practical applications, for example, using tablets and mobile devices, and the expressions are transferred with a joint-based system to improve the quality and persuasiveness of animations.</p> <p>(c) Realistic animation of an expressive human face, which is an important research field in computer graphics, has been a significant challenge due to human face perception. It is a complex, costly, and time-consuming process, which requires a considerable detail in texture, modelling, skin look, and muscle animation [29].</p> <p>(d) This research aims to simplify the process of creating realistic facial animations by using the facial expression transfer.</p> <p>(e) In this study, we have presented a calibration-free, fast, easy to use, and an affordable approach providing facial animation of a 3D input model. Our</p>

	<p>method requires a single camera for real-time tracking of facial movements of an actor in 2D and transfer regular, exaggerated or restricted expressions on a 3D model. The system has been implemented as a platform-based solution for animators, which allows additional arrangements on automatically generated animation frames before producing the final animation.</p>
6	<p>Elgharib, Mohamed, Mallikarjun BR, Ayush Tewari, Hyeongwoo Kim, Wentao Liu, Hans-Peter Seidel, and Christian Theobalt. 2019. "EgoFace: Egocentric Face Performance Capture and Videorealistic Reenactment." <i>ArXiv:1905.10822 [Cs]</i>, May. http://arxiv.org/abs/1905.10822.</p>
	<p>(a) Our lightweight setup allows operations in uncontrolled environments, and lends itself to telepresence applications such as video-conferencing from dynamic environments. The input image is projected into a low dimensional latent space of the facial expression parameters. Through careful adversarial training of the parameter-space synthetic rendering, a videorealistic animation is produced. Our problem is challenging as the human visual system is sensitive to the smallest face irregularities that could occur in the final results. This sensitivity is even stronger for video results. Our solution is trained in a pre-processing stage, through a supervised manner without manual annotations.</p> <p>(b) EgoFace captures a wide variety of facial expressions, including mouth movements and asymmetrical expressions. It works under varying illuminations, background, movements, handles people from different ethnicities and can operate in real time.</p> <p>(c) [Face performance capture algorithms] are crucial in many application areas, for instance, for the creation of digital actors in movies and VFX, as well as for the animation of avatars in virtual augmented reality.</p> <p>(d) However, even recent lightweight face performance capture setups are obstructive and prevent many ubiquitous applications of face performance capture and performance-driven face animation that have great potential.</p> <p>(e) While the aforementioned approaches produce compelling results, they only control facial expressions. Thies et al. presented HeadOn [31], a technique allowing control of the 3D pose and gaze, as well as the expressions. (...) Kim et al. [15] presented Deep Video Portraits, the first approach for controlling the 3D pose, facial expressions and gaze from a single RGB camera. Their approach is based on neural networks where they show generative adversarial networks can produce visually pleasing facial animations, inspired by the success of GANs [14]. (...) While [10, 36] can only control the face expressions, [22] can reenact both pose and expressions of the examined person.</p> <p>(f)</p>

	<p>Furthermore, the parametric face model gives us full control over the human pose in the final animation.</p> <p>(g)</p> <p>Given a single, highly distorted one sided face view as input, we produce a videorealistic face animation of the user from the front view. Our approach handles people from different ethnicities, variable illuminations, backgrounds and a variety of expressions. It outperforms unpaired image translation techniques. Our setup is suitable for mobile video conferencing applications and can run in real time.</p>
7	<p>Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. "Generative Adversarial Networks." <i>ArXiv:1406.2661 [Cs, Stat]</i>, June. http://arxiv.org/abs/1406.2661.</p>
	<p>(a)</p> <p>There is no need for any Markov chains or unrolled approximate inference networks during either training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.</p> <p>(b)</p> <p>So far, the most striking successes in deep learning have involved discriminative models, usually those that map a high-dimensional, rich sensory input to a class label [14, 22]. These striking successes have primarily been based on the backpropagation and dropout algorithms, using piecewise linear units [19, 9, 10] which have a particularly well-behaved gradient. Deep generative models have had less of an impact, due to the difficulty of approximating many intractable probabilistic computations that arise in maximum likelihood estimation and related strategies, and due to difficulty of leveraging the benefits of piecewise linear units in the generative context. We propose a new generative model estimation procedure that sidesteps these difficulties.</p> <p>(c)</p> <p>In the proposed adversarial nets framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine articles.</p>
8	<p>Ha, Sungjoo, Martin Kersner, Beomsu Kim, Seokjun Seo, and Dongyoung Kim. 2019. "MarioNETte: Few-Shot Face Reenactment Preserving Identity of Unseen Targets." <i>ArXiv:1911.08139 [Cs]</i>, November. http://arxiv.org/abs/1911.08139.</p>

	<p>(a) The identity preservation problem, where the model loses the detailed information of the target leading to a defective output, is the most common failure mode. The problem has several potential sources such as the identity of the driver leaking due to the identity mismatch, or dealing with unseen large poses. To overcome such problems, we introduce components that address the mentioned problem: image attention block, target feature alignment, and landmark transformer.</p> <p>(b) Comprehensive experiments are performed to verify that the proposed framework can generate highly realistic faces, outperforming all other baselines, even under a significant mismatch of facial characteristics between the target and the driver.</p> <p>(c) Given a target face and a driver face, face reenactment aims to synthesize a reenacted face which is animated by the movement of a driver while preserving the identity of the target.</p> <p>(d) Many approaches make use of generative adversarial networks (GAN) which have demonstrated a great success in image generation tasks. Xu et al.; Wu et al. (2017; 2018) achieved high-fidelity face reenactment results by exploiting CycleGAN (Zhu et al. 2017). However, the CycleGAN-based approaches require at least a few minutes of training data for each target and can only reenact predefined identities, which is less attractive in-the-wild where a reenactment of unseen targets cannot be avoided.</p> <p>(e) We introduce a novel method of landmark transformation which cope with varying facial characteristics of different people.</p> <p>(f) We report the ratio of deception, which demonstrates the identity preservation and the photo-realism of each model.</p> <p>(g) Proposed method do not need additional fine-tuning phase for identity adaptation, which significantly increases the usefulness of the model when deployed in-the-wild. Our experiments including human evaluation suggest the excellence of the proposed method. One exciting avenue for future work is to improve the landmark transformer to better handle the landmark disentanglement to make the reenactment even more convincing.</p>
9	<p>Heusel, Martin, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. 2018. "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium." <i>ArXiv:1706.08500 [Cs, Stat]</i>, January. http://arxiv.org/abs/1706.08500.</p>
	<p>(a) Generative Adversarial Networks (GANs) excel at creating realistic images with complex models for which maximum likelihood is infeasible.</p> <p>(b)</p>

	<p>Finally, to evaluate GANs, we introduced the ‘Fréchet Inception Distance’ (FID) which captures the similarity of generated images to real ones better than the Inception Score. In experiments we have compared GANs trained with TTUR to conventional GAN training with a one time-scale update rule on CelebA, CIFAR-10, SVHN, LSUN Bedrooms, and the One Billion Word Benchmark. TTUR outperforms conventional GAN training consistently in all experiments.</p>
10	<p>Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. 2018. “Image-to-Image Translation with Conditional Adversarial Networks.” <i>ArXiv:1611.07004 [Cs]</i>, November. http://arxiv.org/abs/1611.07004.</p>
	<p>(a) This makes it possible to apply the same generic approach to problems that traditionally would require very different loss formulations. We demonstrate that this approach is effective at synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images, among other tasks. Indeed, since the release of the pix2pix software associated with this paper, a large number of internet users (many of them artists) have posted their own experiments with our system, further demonstrating its wide applicability and ease of adoption without the need for parameter tweaking.</p> <p>(b) But, just like King Midas, we must be careful what we wish for! If we take a naive approach and ask the CNN [convolutional neural nets] to minimize the Euclidean distance between predicted and ground truth pixels, it will tend to produce blurry results [43, 62]. This is because Euclidean distance is minimized by averaging all plausible outputs, which causes blurring. Coming up with loss functions that force the CNN to do what we really want – e.g., output sharp, realistic images – is an open problem and generally requires expert knowledge.</p> <p>(c) Our primary contribution is to demonstrate that on a wide variety of problems, conditional GANs produce reasonable results. Our second contribution is to present a simple framework sufficient to achieve good results, and to analyze the effects of several important architectural choices. Code is available at https://github.com/phillipi/pix2pix.</p> <p>(d) Blurry images will not be tolerated since they look obviously fake. Because GANs learn a loss that adapts to the data, they can be applied to a multitude of tasks that traditionally would require very different kinds of loss functions.</p> <p>(e) First, we run “real vs. fake” perceptual studies on Amazon Mechanical Turk (AMT). For graphics problems like colorization and photo generation, plausibility to a human observer is often the ultimate goal.</p> <p>(f) The results in this paper suggest that conditional adversarial networks are a promising approach for many image-to-image translation tasks, especially those involving highly structured graphical outputs. These networks learn a</p>

	<p>loss adapted to the task and data at hand, which makes them applicable in a wide variety of settings.</p>
11	<p>Kim, Hyeongwoo, Pablo Garrido, Ayush Tewari, Weipeng Xu, Justus Thies, Matthias Nießner, Patrick Pérez, Christian Richardt, Michael Zollhöfer, and Christian Theobalt. 2018. "Deep Video Portraits." <i>ArXiv:1805.11714 [Cs]</i>, May. http://arxiv.org/abs/1805.11714.</p>
	<p>(a) We present a novel approach that enables photo-realistic re-animation of portrait videos using only an input video. In contrast to existing approaches that are restricted to manipulations of facial expressions only, we are the first to transfer the full 3D head position, head rotation, face expression, eye gaze, and eye blinking from a source actor to a portrait video of a target actor. The core of our approach is a generative neural network with a novel space-time architecture. The network takes as input synthetic renderings of a parametric face model, based on which it predicts photo-realistic video frames for a given target actor. The realism in this rendering-to-video transfer is achieved by careful adversarial training, and as a result, we can create modified target videos that mimic the behavior of the synthetically-created input.</p> <p>(b) To the best of our knowledge, our approach is the first to synthesize full photo-realistic video portraits of a target person's upper body, including realistic clothing and hair, and consistent scene background, under full 3D control of the target's head. To summarize, we make the following technical contributions:</p> <ul style="list-style-type: none"> • A rendering-to-video translation network that transforms coarse face model renderings into full photo-realistic portrait video output. • A novel space-time encoding as conditional input for temporally coherent video synthesis that represents face geometry, reflectance, and motion as well as eye gaze and eye blinks. • A comprehensive evaluation on several applications to demonstrate the flexibility and effectiveness of our approach. <p>We demonstrate the potential and high quality of our method in many intriguing applications, ranging from face reenactment and visual dubbing for foreign language movies to user-guided interactive editing of portrait videos for movie postproduction. A comprehensive comparison to state-of-the-art methods and a user study confirm the high fidelity of our results.</p> <p>(c) In contrast, our approach learns to synthesize photo-realistic facial motion and actions from coarse renderings, thus enabling the synthesis of expressions and joint modification of the head pose, with consistent body and background.</p> <p>(d) Olszewski et al. [2017] synthesize a realistic inner face texture, but cannot generate a fully controllable output video, including person-specific hair. Lassner et al. [2017] propose a generative model to synthesize people in clothing, and Ma et al. [2017] generate new images of persons in arbitrary</p>

	<p>poses using image-to-image translation. In contrast, our approach enables the synthesis of temporally-coherent video portraits that follow the animation of a source actor in terms of head pose, facial expression and eye gaze.</p> <p>(e) Our deep video portraits approach provides full control of the head of a target actor by transferring the rigid head pose, facial expression, and eye motion of a source actor, while preserving the target's identity and appearance.</p> <p>(f) We further demonstrate the potential of our approach on several video rewrite applications, such as reenactment under full head and facial expression control, facial expression reenactment only, video dubbing, and live video portrait editing under user control.</p> <p>(g) We conducted two extensive web-based user studies to quantitatively evaluate the realism of our results. We prepared short 5-second video clips that we extracted from both real and synthesized videos (see Figure 18), to evaluate three applications of our approach: self-reenactment, same-person-reenactment and visual dubbing.</p> <p>(h) The results in Table 1 show that only 80% of participants rated real 256×256 videos as real, i.e. (strongly) agreeing to the video looking real; it seems that in anticipation of synthetic video clips, participants became overly critical. At the same time, 50% of participants consider our 256×256 results to be real, which increases slightly to 52% for 512×512. Our best result is the self-reenactment of Vladimir Putin at 256×256 resolution, which 65% of participants consider to be real, compared to 78% for the real video.</p> <p>(i) On a broader scale, and not being a limitation, democratization of advanced high-quality video editing possibilities, offered by our and other methods, calls for additional care in ensuring verifiable video authenticity, e.g., through invisible watermarking.</p> <p>(j) We have shown through experiments and a user study that our method outperforms prior work in quality and expands over their possibilities. It thus opens up a new level of capabilities in many applications, like video reenactment for virtual reality and telepresence, interactive video editing, and visual dubbing. We see our approach as a step towards highly realistic synthesis of full-frame video content under control of meaningful parameters. We hope that it will inspire future research in this very challenging field.</p>
12	<p>Kosarevych, Ivan, Marian Petruk, Markian Kostiv, Orest Kupyn, Mykola Maksymenko, and Volodymyr Budzan. 2020. "ActGAN: Flexible and Efficient One-Shot Face Reenactment." <i>ArXiv:2003.13840 [Cs]</i>, March. http://arxiv.org/abs/2003.13840.</p>

	<p>(a) Given two images, the goal is to transfer the facial expression of the source actor onto a target person in a photo-realistic fashion.</p> <p>(b) To this end, we employ the Feature Pyramid Network (FPN) as a core generator building block – the first application of FPN in face reenactment, producing finer results. We also introduce a solution to preserve a person's identity between synthesized and target person by adopting the state-of-the-art approach in deep face recognition domain.</p> <p>(c) Face reenactment aims at transferring of facial expression from source to target [30, 36]. This field of research has a myriad of real-world applications, such as cinematography, mass media, AR/VR, computer games, telepresence, to name a few.</p> <p>(d) However, existing methods such as [14, 36] are limited in scalability when a target person is not predefined. Such methods have to train large networks, where both generator and discriminator have a vast number of parameters for each target person. Meanwhile, solutions that overcome this limitation, such as Zakharov et al. [39], require additional retraining.</p> <p>(e) We compare our model with the state-of-the-art approaches, as well as run ablation study of different model architectures and training methods. We evaluate three components, which are significant for face reenactment: 1) image realism, 2) expression accuracy, 3) identity preservation.</p> <p>(f) As we are aiming at better identity preservation and higher realism, we choose separate encoders to siamese ones.</p> <p>(g) We plan to extend ActGAN for real-time face reenactment on videos, and better handling edge-case scenarios. We also hope that our research will help researchers to improve the accuracy of the models in the face-forensics domain.</p>
13	<p>Liu, Zhaoxiang, Huan Hu, Zipeng Wang, Kai Wang, Jinqiang Bai, and Shiguo Lian. 2019. "Video Synthesis of Human Upper Body with Realistic Face." In <i>2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)</i>, 200–202. Beijing, China: IEEE. https://doi.org/10.1109/ISMAR-Adjunct.2019.00-47.</p>
	<p>(a) Recently, with the dramatic advances of Generative Adversarial Network (GAN) [9, 16, 17, 20], GAN-based human body information transfer, such as body motion transfer [6] and face transfer [4, 18], which avoids expensive 3D motion capture, has gained much attention. However, these works mainly focus on either holistic transfer or local transfer. Consequently, facial details are not realistic and body motions will be lost. This limits their further</p>

	<p>applications. Therefore, transferring human motion while keeping realistic facial expressions becomes more and more essential and promising.</p> <p>(b) To address this issue, we propose to combine UBKP, facial action units and pose (FAUP) to realize body motion transfer, to utilize FAUP instead of facial landmarks to enhance the details of the synthesized face. We firstly map the source's FAUP into target's facial landmarks, and then translate the landmarks into target's face image. Compared to facial landmarks, FAUP have the advantages of being independent of identity and containing no spatial coordinate information. It thus can provide an abstract description of facial expressions and local movements without complicated normalization. The proposed pipeline is able to transfer both the human upper body motion and face information at the same time.</p> <p>(c) We conduct an experiment to synthesize two videos of one individual which are driven by two different individuals. Experimental results are shown in Fig. 4. We can see that the synthesized target sequences are photorealistic and consistent with the source sequence in body motions, face expressions and poses, which is also a strong demonstration of our method's performance.</p> <p>(d) This paper presents a novel pipeline to perform human upper body video synthesis. Given an upper body video of a source person, we use upper body keypoints, facial action units and face poses as compact media to generate a target video consistent with the motion, expression, and pose of the source video. Experimental results show the effectiveness of our method. However, our method suffers from the problem of efficiency. Our future work will focus on exploring more lightweight networks to speed up the video synthesis.</p>
14	<p>Lucic, Mario, Michael Tschannen, Marvin Ritter, Xiaohua Zhai, Olivier Bachem, and Sylvain Gelly. 2019. "High-Fidelity Image Generation With Fewer Labels." <i>ArXiv:1903.02271 [Cs, Stat]</i>, May. http://arxiv.org/abs/1903.02271.</p>
	<p>(a) Recent work on conditional generative adversarial networks has shown that learning complex, high-dimensional distributions over natural images is within reach. While the latest models are able to generate high-fidelity, diverse natural images at high resolution, they rely on a vast quantity of labeled data. In this work we demonstrate how one can benefit from recent work on self- and semi-supervised learning to outperform the state of the art on both unsupervised ImageNet synthesis, as well as in the conditional setting. In particular, the proposed approach is able to match the sample quality (as measured by FID) of the current state-of-the-art conditional model BigGAN on ImageNet using only 10% of the labels and outperform it using 20% of the labels.</p> <p>(b) In this work, we take a significant step towards closing the gap between conditional and unsupervised generation of high-fidelity images using</p>

	<p>generative adversarial networks (GANs). We leverage two simple yet powerful concepts.</p> <p>(c) We believe that this is a great first step towards the ultimate goal of few-shot high-fidelity image synthesis. There are several important directions for future work: (i) investigating the applicability of these techniques for even larger and more diverse data sets, and (ii) investigating the impact of other self- and semi-supervised approaches on the model quality. (iii) investigating the impact of self-supervision in other deep generative models.</p>
15	<p>Ma, Tianxiang, Bo Peng, Wei Wang, and Jing Dong. 2019. "Any-to-One Face Reenactment Based on Conditional Generative Adversarial Network." In <i>2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)</i>, 1657–64. Lanzhou, China: IEEE. https://doi.org/10.1109/APSIPAASC47483.2019.9023328.</p>
	<p>(a) Face reenactment refers to the process of transferring the expressions and postures of a given face to the target face. We present a novel Any-to-one Face Reenactment Model based on Conditional Generative Adversarial Network, which has a simple dual converter structure: Any-to-one Face Landmarks Map Converter(AFLC) and Landmark-to-face Converter based on Conditional Generative Adversarial Network(LFC).</p> <p>(b) The latter has a generator that transfers the landmarks map of the target face into the realistic and identity-preserving target facial image. The whole model is purely learning-based without any 3D model, and can generate high quality transferred face comparable to the state-of-the-art. What's more the model is highly robust to wild faces, including various faces of different complexions, ages, and genders. We performed an ablation study on our proposed AFLC to verify its importance for face reenactment of any object. AFLC helps the overall model to achieve an effective facial reenactment.</p> <p>(c) Face reenactment has a wide range of applications, such as CGI, virtual reality, video games, video conference and so on.</p> <p>(d) In this transfer process, it is necessary to keep the identity attribute of the target face unchanged, and the target face generated at the same time also needs to have facial attributes such as expressions and postures of source face.</p> <p>(e) Its generator can transfer the landmarks map of the target face into the realistic and identity-preserving target face. Here the face landmarks map is used as an intermediate medium to connect the source face and the target face, which can solve the image pairing problem.</p> <p>(f) Our model achieves purely learning-based facial expressions and postures transformation, which is able to implement face reenactment for various wild</p>

	<p>faces, including faces of different complexions, ages, and genders. The whole model can effectively achieve any-to-one face reenactment.</p> <p>(g)</p> <p>The overall model structure is simple and lightweight, which has a dual converter structure: Any-to-one Face Landmarks Map converter(AFLC) and Landmark-to-face Converter based on Conditional Generative Adversarial Network(LFC). LFC we proposed can generate realistic and identity-preserving facial image from landmarks map. We create AFLC that guarantees to generate robust faces, and transfers facial expressions and postures while preserving the shape and textures of the target face. We verify that our proposed model enables pure learning-based any-to-one face reenactment. The state-of-the-art papers' source code is not available for comparison, hence we compare our results with those demos and results showed in their papers and perform a user study. From these comparison results we confirm the effect of our proposed model is comparable to or better than state-of-the-art.</p>
16	<p>Pham, Hai X., Samuel Cheung, and Vladimir Pavlovic. 2017. "Speech-Driven 3D Facial Animation with Implicit Emotional Awareness: A Deep Learning Approach." In <i>2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)</i>, 2328–36. Honolulu, HI, USA: IEEE. https://doi.org/10.1109/CVPRW.2017.287.</p>
	<p>(a)</p> <p>Human-machine interaction has been one active research area for decades, with the ultimate goal to make human-machine interaction transparent.</p> <p>(b)</p> <p>One such application is the development of talking agent, either in the form of virtual or physical (i.e. robotic) avatars, for face-to-face human-machine interaction, as in computer-assisted voice agent. In this scenario, the recorded speech can easily be manipulated, by changing the speed or pitch, to reflect the artificial emotion of the digital assistant. These changes can be automatically reflected visually on the avatar, and make the interaction more engaging. On the other hand, it can also make inter-person telecommunication more enjoyable by expressing speech via personalized avatars, especially in interactive role-playing games, where the gamers communicate with other characters in the virtual world.</p> <p>(c)</p> <p>Moreover, speech-emanating facial movements involve different activations of correlated regions on the geometric surface, thus it is difficult to achieve realistic looking, emotion-aware facial deformation from speech sequence.</p> <p>(d)</p> <p>To tackle the difficulty of avatar generation, we utilize the blendshape model in [7], which is purposely designed with enough constraints to ensure that, the final model would always look realistic given a specific set of control parameters. In addition, it can represent various emotional states, e.g. sadness, happiness, etc., without explicitly specifying them. In order to directly map the input features to face shape parameters, we use deep</p>

	<p>recurrent neural network with LSTM cells [15] to model the long range context of the sequence.</p> <p>(e) These approaches usually generate photo-realistic video output, as they compose the result from real images with natural textures. However, their performance and quality are limited by the amount of samples in the database, thus it is difficult to generalize to a large corpus of speeches, which would require a tremendous amount of image samples to cover all possible facial appearances. In contrast, although lacking in photo-realism, model-based approaches enjoy the flexibility of a deformable model, which is controlled by only a set of parameters, and more straightforward modeling.</p> <p>(f) Long short-term memory recurrent neural networks. Recurrent neural networks (RNNs) [27] have demonstrated highly desirable performance in sequence modeling with the ability to integrate temporal contextual information. Hochreiter et al. [15] introduced the Long short-term memory (LSTM) cell in RNN to overcome the vanishing gradient problem [14] in modeling long-term relation. In this work, we aim to estimate the facial transformation trajectory in real-time, hence we utilize unidirectional (forward) LSTM-RNN that only memorizes the past data.</p> <p>(g) This paper presents a deep recurrent learning approach for speech-driven 3D facial animation. Our regression framework, based on deep long short-term memory recurrent neural network, directly maps various acoustic features of an input speech sequence to head rotation and facial deformation parameters of a 3D blendshape model for realistic animation in real-time. Experimental results on a real audio-visual corpus consisting of speeches under various emotions demonstrate the effectiveness of our approach in recreating the affective state and facial deformation of the speaker. We believe our work is a reasonably good baseline for further research in speech-driven facial animation. In the future, we will explore the ability to learn features directly from the raw waveform data, and incorporate deep generative model in our framework to improve its facial parameter generation quality.</p>
17	Richardson, Elad, Matan Sela, Roy Or-El, and Ron Kimmel. 2017. "Learning Detailed Face Reconstruction from a Single Image." In <i>2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)</i> , 5553–62. Honolulu, HI: IEEE. https://doi.org/10.1109/CVPR.2017.589 .
	<p>(a) Reconstructing the detailed geometric structure of a face from a given image is a key to many computer vision and graphics applications, such as motion capture and reenactment. The reconstruction task is challenging as human faces vary extensively when considering expressions, poses, textures, and intrinsic geometries. While many approaches tackle this complexity by using additional data to reconstruct the face of a single subject, extracting facial surface from a single image remains a difficult problem. (...) In contrast, we propose to leverage the power of convolutional neural networks to</p>

produce a highly detailed face reconstruction from a single image. (...) The accuracy and robustness of the proposed model is demonstrated by both qualitative and quantitative evaluation tests.

(b)

Faces, with all their complexities and vast number of degrees of freedom, allow us to communicate and express ourselves through expressions, mimics, and gestures. Facial muscles enable us to express our emotions and feelings, while facial geometric features determine one's identity. However, the flexibility of these qualities makes the recovery of facial geometry from a flat image a challenge.

(c)

The usage of an end-to-end network here is exciting as it connects the problem of face reconstruction to the rapidly expanding applications solved by CNNs, potentially allowing us to further improve our results following new advances in CNN architectures. Moreover, it allows fast reconstructions without the need for external initialization or post-processing algorithms. The potential of using a CNN for reconstructing face geometries was recently demonstrated in [34]. However, their network can only produce the coarse geometry, and must be given an aligned template model as initialization. These limitations force their solution to depend on external algorithms for pose alignment and detail refinement.

(d)

Both qualitative and quantitative evaluations are used to demonstrate the strength of the proposed solution. Our method is compared to the template based method of [20], to the 3DMM based method introduced as part of [48] and to the data driven method of [34]. Note that unlike our method, all of the above require alignment information. We use the state-of-the-art alignment method of [19] to provide input for these algorithms.

(e)

Finally, as noted in Section 4.2 the fully convolutional FineNet can receive inputs with varying sizes. This size invariance is a vital property for our detail extraction network, as it allows the network to extract more details when a high quality input image is available. Figure 11 shows that although our network was trained only on 200×200 images it gracefully scales up for 400×400 inputs.

(f)

For example, we found that our network might fail when tested upon unique facial features that were not part of the training data, such as beards, makeup, and glasses, as can be seen in the supplementary material. The second phase of the training is the unsupervised end-to-end training scheme. While we found that this step successfully trains FineNet, it only slightly tunes CoarseNet. We believe that is because the loss function of FineNet is more sensitive to high frequencies, while the 3DMM model captures mainly coarse facial geometries. Still, it would be interesting to see whether one can push the idea of end-to-end training further, to significantly affect CoarseNet and possibly even to remove its dependency on synthetic data.

18	<p>Saatchi, Yunus, and Andrew Gordon Wilson. 2017. "Bayesian GAN." <i>ArXiv:1705.09558 [Cs, Stat]</i>, November. http://arxiv.org/abs/1705.09558.</p>
	<p>(a) The resulting approach is straightforward and obtains good performance without any standard interventions such as feature matching or mini-batch discrimination. By exploring an expressive posterior over the parameters of the generator, the Bayesian GAN avoids mode-collapse, produces interpretable and diverse candidate samples, and provides state-of-the-art quantitative results for semi-supervised learning on benchmarks including SVHN, CelebA, and CIFAR-10, outperforming DCGAN, Wasserstein GANs, and DCGAN ensembles.</p> <p>(b) Learning a good generative model for high-dimensional natural signals, such as images, video and audio has long been one of the key milestones of machine learning. Powered by the learning capabilities of deep neural networks, generative adversarial networks (GANs) [4] and variational autoencoders [6] have brought the field closer to attaining this goal.</p> <p>(c) We interpret data samples from the generator, showing exploration across several distinct modes in the generator weights. We also show data and iteration efficient learning of the true distribution. We also demonstrate state of the art semi-supervised learning performance on several benchmarks, including SVHN, MNIST, CIFAR-10, and CelebA. The simplicity of the proposed approach is one of its greatest strengths: inference is straightforward, interpretable, and stable. Indeed all of the experimental results were obtained without feature matching or any ad-hoc techniques.</p> <p>(d) Effective semi-supervised learning of natural high dimensional data is crucial for reducing the dependency of deep learning on large labelled datasets. Often labeling data is not an option, or it comes at a high cost – be it through human labour or through expensive instrumentation (such as LIDAR for autonomous driving). Moreover, semi-supervised learning provides a practical and quantifiable mechanism to benchmark the many recent advances in unsupervised learning.</p>
19	<p>Salimans, Tim, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. 2016. "Improved Techniques for Training GANs." <i>ArXiv:1606.03498 [Cs]</i>, June. http://arxiv.org/abs/1606.03498.</p>
	<p>(a) We present a variety of new architectural features and training procedures that we apply to the generative adversarial networks (GANs) framework. We focus on two applications of GANs: semi-supervised learning, and the generation of images that humans find visually realistic. Unlike most work on generative models, our primary goal is not to train a model that assigns high likelihood to test data, nor do we require the model to be able to learn well without using any labels. Using our new techniques, we achieve</p>

	<p>state-of-the-art results in semi-supervised classification on MNIST, CIFAR-10 and SVHN. The generated images are of high quality as confirmed by a visual Turing test: our model generates MNIST samples that humans cannot distinguish from real data, and CIFAR-10 samples that yield a human error rate of 21.3%. We also present ImageNet samples with unprecedented resolution and show that our methods enable the model to learn recognizable features of ImageNet classes.</p> <p>(b) One of the primary goals of this work is to improve the effectiveness of generative adversarial networks for semi-supervised learning (improving the performance of a supervised task, in this case, classification, by learning on additional unlabeled examples). Like many deep generative models, GANs have previously been applied to semi-supervised learning [13, 14], and our work can be seen as a continuation and refinement of this effort.</p> <p>(c) A downside of using human annotators is that the metric varies depending on the setup of the task and the motivation of the annotators. We also find that results change drastically when we give annotators feedback about their mistakes: By learning from such feedback, annotators are better able to point out the flaws in generated images, giving a more pessimistic quality assessment.</p> <p>(d) Generative adversarial networks are a promising class of generative models that has so far been held back by unstable training and by the lack of a proper evaluation metric. This work presents partial solutions to both of these problems. We propose several techniques to stabilize training that allow us to train models that were previously untrainable. Moreover, our proposed evaluation metric (the Inception score) gives us a basis for comparing the quality of these models. We apply our techniques to the problem of semi-supervised learning, achieving state-of-the-art results on a number of different data sets in computer vision. The contributions made in this work are of a practical nature; we hope to develop a more rigorous theoretical understanding in future work.</p>
20	<p>Siarohin, Aliaksandr, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, and Nicu Sebe. 2020. "First Order Motion Model for Image Animation." <i>ArXiv:2003.00196 [Cs]</i>, October. http://arxiv.org/abs/2003.00196.</p>
	<p>(a) Our framework addresses this problem without using any annotation or prior information about the specific object to animate.</p> <p>(b) A generator network models occlusions arising during target motions and combines the appearance extracted from the source image and the motion derived from the driving video. Our framework scores best on diverse benchmarks and on a variety of object categories. Our source code is publicly available.</p> <p>(c)</p>

Generating videos by animating objects in still images has countless applications across areas of interest including movie production, photography, and e-commerce. More precisely, image animation refers to the task of automatically synthesizing videos by combining the appearance extracted from a source image with motion patterns derived from a driving video. For instance, a face image of a certain person can be animated following the facial expressions of another individual (see Fig. 1). In the literature, most methods tackle this problem by assuming strong priors on the object representation (e.g. 3D model) [4] and resorting to computer graphics techniques [6, 34]. These approaches can be referred to as object-specific methods, as they assume knowledge about the model of the specific object to animate.

(d)

Recently, deep generative models have emerged as effective techniques for image animation and video retargeting [2, 42, 3, 43, 28, 29, 38, 41, 32, 22]. In particular, Generative Adversarial Networks (GANs) [14] and Variational Auto-Encoders (VAEs) [21] have been used to transfer facial expressions [38] or motion patterns [3] between human subjects in videos. Nevertheless, these approaches usually rely on pre-trained models in order to extract object-specific representations such as keypoint locations. Unfortunately, these pre-trained models are built using costly ground-truth data annotations [2, 28, 32] and are not available in general for an arbitrary object category.

(e)

Third, we extend the equivariance loss commonly used for keypoints detector training [18, 45], to improve the estimation of local affine transformations. Fourth, we experimentally show that our method significantly outperforms state-of-the-art image animation methods and can handle high-resolution datasets where other approaches generally fail. Finally, we release a new high resolution dataset, Thai-Chi-HD, which we believe could become a reference benchmark for evaluating frameworks for image animation and video generation.

(f)

In contrast to these works, we neither rely on labels, prior information about the animated objects, nor on specific training procedures for each object instance. Furthermore, our approach can be applied to any object within the same category (e.g., faces, human bodies, robot arms etc).

(g)

Conversely, our approach is able to generate significantly better looking videos in which each body part is independently animated. This qualitative evaluation illustrates the potential of our rich motion description. We complete our evaluation with a user study. We ask users to select the most realistic image animation. Each question consists of the source image, the driving video, and the corresponding results of our method and a competitive method. We require each question to be answered by 10 AMT worker[s]. This evaluation is repeated on 50 different input pairs. Results are reported in Tab. 2. We observe that our method is clearly preferred over the competitor methods. Interestingly, the largest difference with the state of the art is obtained on Tai-Chi-HD: the most challenging dataset in our evaluation due to its rich motions.

21	Thies, Justus, Mohamed Elgharib, Ayush Tewari, Christian Theobalt, and Matthias Nießner. 2020. "Neural Voice Puppetry: Audio-Driven Facial Reenactment." <i>ArXiv:1912.05566 [Cs]</i> , July. http://arxiv.org/abs/1912.05566 .
	<p>(a) We present Neural Voice Puppetry, a novel approach for audio-driven facial video synthesis. Given an audio sequence of a source person or digital assistant, we generate a photo-realistic output video of a target person that is in sync with the audio of the source input.</p> <p>(b) Through the underlying 3D representation, the model inherently learns temporal stability while we leverage neural rendering to generate photo-realistic output frames. Our approach generalizes across different people, allowing us to synthesize videos of a target actor with the voice of any unknown source actor or even synthetic voices that can be generated utilizing standard text-to-speech approaches. Neural Voice Puppetry has a variety of use-cases, including audio-driven video avatars, video dubbing, and text-driven video synthesis of a talking head. We demonstrate the capabilities of our method in a series of audio- and text-based puppetry examples, including comparisons to state-of-the-art techniques and a user study.</p> <p>(c) Digital voice assistants are now ubiquitous due to their integration into many commodity devices such as smartphone, tvs, cars, etc.; even companies use more and more machine learning techniques to drive service bots that interact with their customers. These virtual agents aim for a user-friendly man-machine interface while keeping maintenance costs low. However, a significant challenge is to appeal to humans by delivering information through a medium that is most comfortable to them. While speech-based interaction is already very successful, such as shown in virtual assistants like Siri, Alexa, Google, etc., the visual counterpart is largely missing. This comes to no surprise given that a user would also like to associate the visuals of a face with the generated audio, similar to the ideas behind video conferencing. In fact, the level of engagement for audio-visual interactions is higher than for purely audio ones [10,28].</p> <p>(d) The aim of this work is to provide the missing visual channel by introducing Neural Voice Puppetry, a photo-realistic facial animation method that can be used in the scenario of a visual digital assistant.</p> <p>(e) The key component of our method is to estimate lip motions that fit the input audio and to render the appearance of the target person in a convincing way.</p> <p>(f) Facial reenactment is the process of re-animating a target video in a photo-realistic manner with the expressions of a source actor [32,38]. It enables a variety of applications, ranging from consumer-level teleconferencing through photo-realistic virtual avatars [23,30,31] to movie production applications such as video dubbing [12,19]. Recently, several</p>

	<p>authors started to exploit the audio signal for facial reenactment [5,27,35]. This has the potential of avoiding failures of visual-based approaches, when the visual signal is not reliable, e.g., due to occluded face, noise, distorted views and so on. Many of these approaches, however, lack video-realism [5,35], since they work in a normalized space of facial imagery (cropped, frontal faces), to be agnostic to head movements.</p> <p>(g) A novel light-weight neural rendering network using neural textures is presented that allows us to generate photo-realistic video content reproducing the person-specific appearance. It surpasses the quality and speed of state-of-the-art neural rendering methods [11,33].</p> <p>(h) To enable photo-realistic facial reenactment based on audio signals, we employ a 3D face model as intermediate representation of facial motion. A key component of our pipeline is the audio-based expression estimation.</p>
22	<p>Tripathy, Soumya, Juho Kannala, and Esa Rahtu. 2020a. "ICface: Interpretable and Controllable Face Reenactment Using GANs." <i>ArXiv:1904.01909 [Cs]</i>, January. http://arxiv.org/abs/1904.01909.</p>
	<p>(a) The results indicate that ICface produces better visual quality, while being more versatile than most of the comparison methods. The introduced model could provide a lightweight and easy to use tool for multitude of advanced image and video editing tasks. The program code will be publicly available upon the acceptance of the paper.</p> <p>(b) A lightweight and easy-to-use tool for this type of manipulation task would have numerous applications in animation industry, movie post-production, virtual reality, photography technology, video editing and interactive system design, among others.</p> <p>(c) A commonly used procedure is to take a source face and a set of desired facial attributes (e.g. pose) as an input and produce a face image depicting the source identity with the desired attributes. The source face is usually specified by one or more example images depicting the selected person. The facial attributes could be presented by categorical variables, continuous parameters or by another face image (referred as a driving image) with desired pose and expression.</p> <p>(d) We make the following three contributions. i) We propose a data driven and GAN based face animation system that is applicable to a large number of source and driving identities. ii) The proposed system is driven by human interpretable control signals obtainable from multiple sources such as external driving videos and manual controls. iii) We demonstrate our system in multiple tasks including face reenactment, facial emotion synthesis, and multi-view image generation from single-view input. The presented results outperform several recent (possibly purpose-built) state-of-the-art works.</p> <p>(e)</p>

	<p>A recent work [25] proposed a method called GANimation that was capable of generating a wide range of continuous emotions and expressions for a given face image. They utilized the well known concept of action units (AUs) as a conditioning vector and obtained very appealing results. Similar results are achieved on portrait face images in [12, 1]. Unfortunately, unlike our method, these approaches are not suitable for full reenactment, where the head pose has to be modified. Moreover, they add the new attribute directly on to the existing expression of the source image which can be problematic to handle.</p> <p>(f) Furthermore, our method is based on interpretable and continuous attributes that are extracted from the driving video. The flexible interface allows user to easily mix the attributes from multiple driving videos and edit them manually at will. In the extreme case, the attributes could be defined without any driving video at all.</p> <p>(g) Moreover, even if one could record the proper driving frames, one may still wish to perform post-production editing on the result. This type of editing is hard to implement with previous methods like X2Face [32] and DAE [26] since the facial representation is learned implicitly and it lacks a clear interpretability. Instead, the head pose angles and AUs, utilised in our approach, provide human interpretable and easy-to-use interface for selective editing.</p> <p>(h) The proposed method produces good quality results and provides control over the animation process, unlike other methods.</p> <p>(i) In this paper, we proposed a generic face animator that is able to control the pose and expression of a given face image. The animation was controlled using human interpretable attributes consisting of head pose angles and action unit activations.</p>
23	<p>———. 2020b. "FACEGAN: Facial Attribute Controllable REenactment GAN." <i>ArXiv:2011.04439 [Cs]</i>, November. http://arxiv.org/abs/2011.04439.</p>
	<p>(a) These models perform best if the source and driving images depict the same person or if the facial structures are otherwise very similar. However, if the identity differs, the driving facial structures leak to the output distorting the reenactment result. We propose a novel Facial Attribute Controllable rEenactment GAN (FACEGAN), which transfers the facial motion from the driving face via the Action Unit (AU) representation. Unlike facial landmarks, the AUs are independent of the facial structure preventing the identity leak. Moreover, AUs provide a human interpretable way to control the reenactment. FACEGAN processes background and face regions separately for optimized output quality. The extensive quantitative and qualitative comparisons show a clear improvement over the state-of-the-art in a single source reenactment task. The results are best illustrated in the reenactment video provided in the supplementary material. The source code will be made available upon publication of the paper.</p>

	<p>(b) Face-reenactment is a process of animating a source face according to the motion (pose and expression) of a driving face. In general, the process involves three major steps: 1) creating a representation of the source face identity, 2) extracting and encoding the motion of the driving face, and 3) combining the identity and motion representations to produce a modified source face. Each part has a significant impact on the output quality. Number of algorithms, including traditional 3D face models [3, 4], data driven Neural-Networks [5, 6, 7, 1, 8, 9, 10, 11, 12], and their combinations [13] have been presented for creating photorealistic face animations.</p> <p>(c) Although this approach results in high quality outputs, they require substantial efforts in obtaining the faithful 3D representations of the faces. Therefore, such approaches are often limited to a few source identities.</p> <p>(d) We proposed a facial animator called FACEGAN that is capable of performing high quality reenactment from a single source image. Unlike many previous works, our model does not pose any restriction on the compatibility of the source and driving pairs. The model combines the best properties of the action unit and facial landmark motion representations for reducing the identity leakage problem and to optimise the reenactment quality. Furthermore, FACEGAN handles the face and the background separately which improves the output quality and gives additional control of choosing the desired background. We have compared our method with the state-of-the-art approaches and obtained superior results both quantitatively and qualitatively.</p>
24	<p>Yang, Chao, and Ser-Nam Lim. 2019. "Unconstrained Facial Expression Transfer Using Style-Based Generator." <i>ArXiv:1912.06253 [Cs]</i>, December. http://arxiv.org/abs/1912.06253.</p>
	<p>(a) Facial expression transfer and reenactment has been an important research problem given its applications in face editing, image manipulation, and fabricated videos generation. We present a novel method for image-based facial expression transfer, leveraging the recent style-based GAN shown to be very effective for creating realistic looking images.</p> <p>(b) We further introduce a linear combination scheme that fuses the style vectors of the two given images and generate a new face that combines the expression and appearance of the inputs. Our method can create high-quality synthesis with accurate facial reenactment. Unlike many existing methods, we do not rely on geometry annotations, and can be applied to unconstrained facial images of any identities without the need for retraining, making it feasible to generate large-scale expression-transferred results.</p> <p>(c) Editing in the latent manifold instead of the pixel space ensures that the image does not fall off the manifold and looks natural and realistic. Second,</p>

	<p>StyleGAN can learn hierarchical “style” vectors that are shown to explain attributes at different levels, from fine attributes such as hair color or eyes open/closed to high-level aspects such as pose, face shape, eyeglasses. At the core of our algorithm is an optimization scheme that infers and combines the style vectors to create a face that fuses the appearance and the expression of two images.</p> <p>(d) However, when a single image is provided, it needs either the neutral expression or manual initialization. Besides, the result often lacks realism and fine details due to the process of projecting the rendered texture onto a 3D model.</p> <p>(e) In the case of a single target image, the inner mouth interiors and fine details such as wrinkles are hallucinated by copying from the source, which often leads to uncanny results.</p> <p>(f) Failure cases We also observe some failure cases in our results. For example, when a large portion of the face is occluded, the model fails to recover the complete face and is unable to generate meaningful expressions (Fig. 16 left). The output may also look uncanny due to excessive shadows in the source identity (Fig. 16 right).</p> <p>(g) We propose a simple yet effective expression transfer method based on StyleGAN. Our method can easily apply to any pair of arbitrary face images and transfer the facial expression from one to another. Our approach not only generates compelling results but is also highly scalable and fully automatic. As future work, we are interested in extending our framework to incorporate head pose reenactment to generate more realistic video results. It would also lead to exciting applications if time efficiency could be improved such that expression transfer could run in real-time.</p>
25	<p>Yao, Guangming, Yi Yuan, Tianjia Shao, and Kun Zhou. 2020. “Mesh Guided One-Shot Face Reenactment Using Graph Convolutional Networks.” <i>Proceedings of the 28th ACM International Conference on Multimedia</i>, October, 1773–81. https://doi.org/10.1145/3394171.3413865.</p>
	<p>(a) Face reenactment aims to animate a source face image to a different pose and expression provided by a driving image. Existing approaches are either designed for a specific identity, or suffer from the identity preservation problem in the one-shot or few-shot scenarios. In this paper, we introduce a method for one-shot face reenactment, which uses the reconstructed 3D meshes (i.e., the source mesh and driving mesh) as guidance to learn the optical flow needed for the reenacted face synthesis. Technically, we explicitly exclude the driving face’s identity information in the reconstructed driving mesh. In this way, our network can focus on the motion estimation for the source face without the interference of driving face shape. We propose a motion net to learn the face motion, which is an asymmetric autoencoder. The encoder is a graph convolutional network (GCN) that learns a latent</p>

motion vector from the meshes, and the decoder serves to produce an optical flow image from the latent vector with CNNs. Compared to previous methods using sparse keypoints to guide the optical flow learning, our motion net learns the optical flow directly from 3D dense meshes, which provide the detailed shape and pose information for the optical flow, so it can achieve more accurate expression and pose on the reenacted face. Extensive experiments show that our method can generate high-quality results and outperforms state-of-the-art methods in both qualitative and quantitative comparisons.

(b)

The reenacted face preserves the same identity as the source person, while owning the pose and expression of the driving person.

Previous approaches have demonstrated great success in face reenactment for a specific identity using generative adversarial networks (GANs). For instance, the work of [47, 48] successfully performs face reenactment between a pair of specified identities using CycleGAN [52]. High-fidelity results are achieved with the help of 3D face reconstruction and GANs [22, 40]. Nevertheless, all these methods require a large number of images for a specific identity, which may be infeasible for many applications.

Consequently, a variety of one-shot or few-shot methods are proposed [36, 37, 45, 50].

(c)

We qualitatively and quantitatively evaluated our method on different datasets. Experimental results show that our one-shot method is able to better preserve the identity and yield a reenacted face with more accurate expression and pose, using only one source image and one driving image that are both unseen in the training data. Our method outperforms the state-of-the-art methods in both objective and subjective aspects.

(d)

More notably, the main problem in one-shot or few-shot face reenactment is the identity preservation problem, that is, the source face shape is difficult to be preserved during reenacting when there is a large shape difference between the source and driving faces. Although the latest few-shot work of [16] can alleviate this problem by introducing the Landmark Transformer mechanism, the result is still not satisfactory for the one-shot cases. In our approach, we explicitly exclude the driving face's identity information in the reconstructed driving face mesh. Thus our network can focus on learning the motion for the source face without the interference of driving face shape.

(e)

Furthermore, thanks to the full shape and pose information provided by dense meshes, our method can generate more natural-looking images with more details.

(f)

Also, our method has the best scores in PSNR and the second-best scores in SSIM, which implies that our method can generate images closest to real images. Finally, it is worth noting that, when reenacting a different person, our method outperforms other methods in all metrics at the Faceforensics++ [35] dataset, which is a very challenging dataset because it contains the most different identities.

(g)

	<p>Compared with other methods, our method can generate more realistic and natural-looking results. In the future, we plan to explore the temporal consistency in the network design to facilitate the face transfer in videos.</p>
26	<p>Zhang, Jiangning, Xianfang Zeng, Chao Xu, Jun Chen, Yong Liu, and Yunliang Jiang. 2020. "APB2FaceV2: Real-Time Audio-Guided Multi-Face Reenactment." <i>ArXiv:2010.13017 [Cs]</i>, October. http://arxiv.org/abs/2010.13017.</p>
	<p>(a) Audio-guided face reenactment aims to generate a photorealistic face that has matched facial expression with the input audio. However, current methods can only reenact a special person once the model is trained or need extra operations such as 3D rendering and image post-fusion on the premise of generating vivid faces. To solve the above challenge, we propose a novel Real-time Audio-guided Multi-face reenactment approach named APB2FaceV2, which can reenact different target faces among multiple persons with corresponding reference face and drive audio signal as inputs. Enabling the model to be trained end-to-end and have a faster speed, we design a novel module named Adaptive Convolution (AdaConv) to infuse audio information into the network, as well as adopt a lightweight network as our backbone so that the network can run in real time on CPU and GPU. Comparison experiments prove the superiority of our approach than existing state-of-the-art methods, and further experiments demonstrate that our method is efficient and flexible for practical applications.</p> <p>(b) Some qualitative experiments are conducted on AnnVI dataset to visually demonstrate the high quality of reenacted images and the flexibility of the proposed approach. Specifically, we randomly select one face of each identity (6 faces totally) as the reference face, and one drive frame of each identity for supplying audio, head pose, and eye blink signals. As indicated by the red rectangles in Figure 2, our proposed method can reenact photorealistic faces among multiple persons with one unified model that achieves multi-face reenactment task. Thanks to the decoupling design of our method, APB2FaceV2 can use input signals of other persons to reenact the target face that is consistent with the identity of the reference face. Experimental results show that our method has strong generalization ability, where the model can use non-self audio as input to reenact photorealistic faces.</p> <p>(c) Specifically, an Audio-aware Fuser is firstly used to predict a geometric representation from input signals, and then Multi-face Reenactor fuse it with the reference face that supplies appearance information to reenact photorealistic target face. Besides, a novel AdaConv module is proposed to inject geometric information in a more elegant and efficient way. Extensive experiments demonstrate the efficiency and flexibility of our approach. We will further combine Neural Architecture Search (NAS) with our approach to search for a more accurate and faster model for better practical applications, and we hope our work will help users to achieve better jobs.</p>

Appendix 2. Corpus of non-technical expert publications

This appendix contains relevant extractions from the non-technical expert publications. They are color-coded the following way:

- Specifically mentioning and framing deepfake resurrection—light blue;
- Hypothetical and actual applications and consequences of technology framed as positive or neutral—pale pink;
- Hypothetical and actual applications and consequences of technology framed as negative—light red berry;
- Inevitable independent development of the technology—dark yellow;
- Realism of the synthesized images (naturalness, persuasiveness, convincingness, etc.)—light green;
- References to the technology streamlining, improving, or otherwise facilitating communication, increasing personalization degree of entertainment—light cornflower blue;
- Excitement of mastery, knowledge accumulation, and pleasure of heuristics—light yellow;
- Increased productivity/profitability of labor relations due to introduction of digital beings—bright yellow;
- Framing uncanniness of the synthesized images—orange;
- Ease of use/adaptation of deepfake, democratization of technology, including open-sourcing the code—grey;
- References to hierarchical distribution of people producing/consuming deepfakes—light purple.

Remarkable metaphors that do not form a recurring cluster are highlighted in magenta.

1	Kneese, Tamara. November 2, 2020. "Robert Kardashian, Joaquin Oliver, and deepfakes of the dead." <i>Slate</i> . Accessed on March 1, 2021, through https://slate.com/technology/2020/11/robert-kardashian-joaquin-oliver-deep-fakes-death.html .
	(a) But in the case of a teenager who was more famous after death than during life, the technologists didn't have much material to work with.

Instead, they created a single image of his face using three different photographs. Still, the video is convincing enough. The uncanny valley effect comes after the fact, once you realize that the young man in the video who is speaking to you is not actually speaking.

(b)

Using deepfake technologies to render dead performers, from Tupac to Michael Jackson and Whitney Houston, is not a new practice. But, as Alyx Gorman recently noted in the Guardian, such deepfake resurrections have typically been used for money-making ventures, such as performances. In the case of Robert Kardashian, the deepfake hologram was intended as an unlikely gift to conjure an intimate connection and personal message. Reviving the dead through such technologies is connected to more mundane questions about inheritance, as well as what it means to grieve when the traces of the dead are intermingled with those of the living.

(c)

One fantasy associated with data is that it can approximate a human personality or possibly a human soul. (...) The examples of Oliver and Kardashian reveal fantasies and discomforts regarding the extension of human life through digital technologies. What does it mean to resurrect the dead, and not only bring them back to life to communicate with them, but to speak through them?

(d)

Living on through social media accounts is one thing, but actually using A.I. to replicate a person's personality in perpetuity is something different entirely, especially when it comes to deepfakes. Deepfakes pose problems for human content moderators and audiences as well as for platforms. Aside from issues sorting out real content from artificially produced content, there are fears regarding the loss of control over one's own image. Deepfakes are often framed as a potential problem in politics, where they are associated with disinformation campaigns intended to skew public opinion. Even unconvincing cheapfakes can go viral. But one of the earliest and most extensive applications of deepfake technology has been in pornography. TikTok creators and other influencers may find themselves the victims of deepfake pornographic images, just as more traditional celebrities have experienced. Not only do they emulate a person's appearance or communication patterns, they move, shake, and gesticulate as they would in life. Deepfake pornography—or, on the end of the spectrum, necromancy—is about the simulation of intimacy and simultaneously threatens the loss of control over one's likeness and one's person. Impersonating the dead can be as much of a violation.

(e)

This possible future echoes the contemporary cultural fantasy surrounding big data, which posits that if we gather enough information, this will be a sufficient proxy for the self as well as for the collective.

(f)

While many people would be horrified by the idea of an uncanny valley version of their dead loved ones or themselves, some technologists are enchanted precisely by the indecipherability of the living and the dead. The problems posed by deepfakes are actually these technologists' deepest wish: that blurring the boundaries between who is dead and who is alive will make it impossible to tell the difference between the two.

2	<p>Rees, Geriant. November 25, 2019. "Here's how deepfake technology can actually be a good thing." <i>World Economic Forum</i>. Accessed on March 1, 2021, through https://www.weforum.org/agenda/2019/11/advantages-of-artificial-intelligence/.</p>
	<p>(a) Recent advances in deepfake video technology have led to a rapid increase of such videos in the public domain in the past year. Face-swapping apps such as Zao, for example, allow users to swap their faces with a celebrity, creating a deepfake video in seconds. These advances are the result of deep generative modelling, a new technology which allows us to generate duplicates of real faces and create new, and impressively true-to-life images, of people who do not exist.</p> <p>(b) This new technology has quite rightly raised concerns about privacy and identity. If our faces can be created by an algorithm, would it be possible to replicate even more details of our personal digital identity or attributes like our voice – or even create a true body double?</p> <p>(c) While questions are rightly being asked about the consequences of deepfake technology, it is important that we do not lose sight of the fact that artificial intelligence (AI) can be used for good, as well as ill. World leaders are concerned with how to develop and apply technologies that genuinely benefit people and planet, and how to engage the whole of society in their development. Creating algorithms in isolation does not allow for the consideration of broader societal concerns to be incorporated into their practical applications.</p> <p>(d) These examples in healthcare highlight that AI is an enabling technology that is neither intrinsically good nor evil. Technology like this depends on the context in which we create and use it. Universities have a critical role to play here. In the UK, universities are leading the world in research and innovation and are focused on making an impact on real-world challenges. (...) Our academics are working with a broad range of experts and organizations to create new algorithms to support science, innovation and society. AI must complement and augment human endeavour, not replace it. We need to combine checks and balances that inhibit or prevent inappropriate use of technology while creating the right infrastructure and connections between different experts to ensure we develop technology that helps society thrive.</p>
3	<p>Ajder, Henry. June 16, 2019. "The ethics of deepfakes aren't always black and white." <i>The next web</i>. Accessed on March 10, 2021, through https://thenextweb.com/podium/2019/06/16/the-ethics-of-deepfakes-arent-always-black-and-white/.</p>

(a)

Since deepfakes emerged in December 2017, most media coverage has focused on their potentially catastrophic applications. These range from deepfake pornography, ransomfakes, smear campaigns against politicians, and a new age of fake news that could worsen the global 'post-truth' crisis.

(b)

While these malicious uses of deepfakes and synthetic media are rightly a cause for concern, there are also positive uses of the same generative AI technologies. For example, Lyrebird, a Canadian startup, has partnered with the ALS foundation on Project Revoice, an initiative that uses generative AI to create personalized synthetic voices for ALS sufferers who have lost the ability to speak. Similarly, DeepEmpathy, a project by MIT and Unicef, creates synthetic images that show what cities such as London and Tokyo would look like if they were bombed, with the aim of fostering empathy with those fleeing from war.

These examples show that, like most technologies, AI generated synthetic media has positive and negative applications. (...) Beyond these morally black and white examples, there are numerous 'grey' applications that don't fit neatly into this classification.

(c)

While governments and businesses are scrambling to counter the explicitly negative uses of synthetic media, and harness the positive ones, greyfakes are quietly developing under far less scrutiny.

(d)

Arguably the most promising commercial applications of AI generated synthetic media lie in the entertainment sector. Using CGI, experts have already realistically re-created the deceased actress Carrie Fisher as Princess Leia in Star Wars. With the rapid improvements in AI generated synthetic media, this practice will likely be automated and scaled with an even higher degree of realism.

(e)

[For] celebrities, the question focuses specifically on whether we can draw the line between respectful recreation and commercial exploitation.

(f)

It isn't just the dead who could be synthetically recreated with morally ambiguous results. When Vice first broke the story of deepfakes, the focus was on their nonconsensual 'deepfake pornography' involving celebrities. Since then, concerns have grown about how the commodified AI tools for creating synthetic media could be used to create revenge pornography that targets ordinary people. (...) Moreover, these services may also normalize the idea of synthetic pornography, which could further exacerbate existing concerns about the negative impact of pornography on psychological and sexual development.

(g)

The question is whether it all undermines real human interaction. (...) Moving forward, we need to ensure generative AI and synthetic media are developed responsibly, encouraging potentially positive applications, while ensuring they do not unwittingly cause harm.

4	<p>Stokes, Patrick. November 25, 2019. "Commentary: When is it okay to use the CGI of dead actors in new movies?" CNA. Accessed on March 1, 2021, through https://www.channelnewsasia.com/news/commentary/deepfakes-james-de-an-finding-jack-cgi-reanimation-ai-ethics-12110590.</p>
	<p>(a) "To be dead," wrote the 20th century French philosopher Jean-Paul Sartre, "is to be a prey for the living." Even Sartre, though, would have struggled to imagine casting James Dean in a movie 64 years after the actor's death.</p> <p>(b) Sadly, it's not an immortality we look forward to. When we fear death, one thing we particularly dread is the end of first-person experience. (...) The danger, as the philosopher Adam Buben put it, is that memorialisation could slip into replacement. (...) The line between remembrance and exploitation is surprisingly porous.</p> <p>(c) Worse, it suggests that James Dean can be replaced, just as algorithm-driven avatars might come to replace, rather than simply commemorate, the dead. The need to think about how we protect the dead from our digital predations isn't going away.</p>
5	<p>Panyatham, Paengsuda. March 10, 2020. "Deepfake Technology in the Entertainment industry: Potential Limitations and Protections." <i>Research center of Carnegie Mellon University's Master of Arts Management program</i>. Accessed on March 13, 2021, through https://amt-lab.org/blog/2020/3/deepfake-technology-in-the-entertainment-industry-potential-limitations-and-protections.</p>
	<p>(a) The use of this technology is undoubtedly expanding. Despite its common inappropriate use, Deepfake technology can surely benefit the entertainment industry, specifically film production. This comes with further consequences to be considered by various parties.</p> <p>(b) The story of Deepfake itself began in September 2017 when a user called 'Deepfakes' posted a computer-generated pornography video of a famous actress on Reddit, an online community. The post gained popularity, and eventually led to a dedicated subreddit r/deepfakes where users could post fake videos of celebrities' faces mapped on porn stars. A script for how to produce such videos was also published in a forum where users exchanged the latest know-hows and potential applications of the technology. Two months after its launch, the forum gained 25,000 members and sparked controversial debates about the use of such technologies. Once the technology raised concerns and threats in wider communities, Reddit decided to shut down the forum. But that was just the beginning.</p>

	<p>Deepfake has since gained wider recognition and been further developed by different users.</p> <p>(c) Though the anonymous developer shut DeepNude down within a day of its release, it was enough to show the public its capabilities. Deepfake technology is much more accessible today and users without any computer engineering background can create a fake video within seconds.</p> <p>(d) This has increased the fear of not only celebrities, but also normal people whose identity is not as ample as that of a celebrity. Though this method is still far from being realized, it demonstrates potential threats of this technology in the future.</p> <p>(e) Deepfake technology could threaten jobs for current actors while making it more difficult for immersing actors to gain success, as legendary actors can now be regenerated after their death.</p> <p>(f) Regarding the laws and tools we have nowadays, it is undeniably true that we are still a step behind the technology. However, when it comes to technology, it is either now or never. As said by the anonymous developer of the notorious DeepNude application “The technology is ready (within everyone’s reach). So if someone has bad intentions, having DeepNude doesn’t change much. ... If I don’t do it, someone else will do it in a year.” (Anonymous, 2019)</p> <p>(g) The only thing we can do is maximize the use of it while being careful not to abuse others at the same time. Therefore developing a legal framework is crucial at this time. Before doing that, it is important to understand the different impacts Deepfake can have in various industries. In the field of politics, Deepfake is considered a threat to reputation and can be used as a political tool to spread convincing fake news. However, in the entertainment industry, Deepfake can create both positive and negative impacts, and aid in the creative process. Lawmakers need to consider all these factors to realize this is not a zero-sum game, and a solution can exist that will benefit all parties.</p>
6	<p>Kemp, Luke. July 8, 2019. “In the age of deepfakes, could virtual actors put humans out of business?” <i>The Guardian</i>. Accessed on March 13, 2021, through https://www.theguardian.com/film/2019/jul/03/in-the-age-of-deepfakes-could-virtual-actors-put-humans-out-of-business.</p>
	<p>(a) Even that link with corporeal reality, though, is no longer absolute. You may have already seen examples of what’s possible: Peter Cushing (or his image) appearing in <i>Rogue One: A Star Wars Story</i> more than 20 years after his death, or Tupac Shakur performing from beyond the grave at Coachella in 2012. We’ve seen the terrifying potential of deepfakes – manipulated footage that could play a dangerous role in the fake news phenomenon.</p>

	<p>Jordan Peele’s remarkable fake Obama video is a key example. Could technology soon make professional actors redundant?</p> <p>(b) Like most of the examples above, the virtual Tupac is a digital human, and was produced by special effects company Digital Domain. Such technology is becoming more and more advanced. Darren Hendler, the company’s digital human group director, explains that it is in effect a “digital prosthetic” – like a suit that a real human has to wear.</p> <p>(c) “That’s [fully-AI-driven performance of digital humans] still pretty far away,” says Hendler. “Artificial intelligence is expanding so rapidly that it’s really hard for us to predict ... I would say that within the next five to ten years, [we’ll see] things that are able to construct semi-plausible versions of a full-facial performance.”</p> <p>(d) Nonetheless, Hendler believes that the advancement of digital and virtual human technology will become more visible to the public, coming much closer to home – literally – within the next 10 years or so: “You may have a big screen on your wall and have your own personal virtual assistant that you can talk to and interact with: a virtual human that’s got a face, and moves around the house with you. To your bed, your fridge, your stove, and to the games that you’re playing.” “A lot of this might sound pretty damn creepy, but elements of it have been coming for a while now.”</p>
7	<p>Caron, Vincent & Vincent Bergeron. June 26, 2019. “What are deepfakes impacts on the media and entertainment industry.” <i>Canada media fund</i>. Accessed on March 12, 2021, through https://cmf-fmc.ca/how-next/articles/what-are-deepfakes-impacts-on-the-media-and-entertainment-industry/.</p>
	<p>(a) Initially known by the general public in the past few years for its great potential for falsification and disinformation, the technology behind deepfake is generating increased business interest among content creators and promises to catalyze the development of hyperpersonalized entertainment products. What are the foreseeable impacts on the media and entertainment industry?</p> <p>(b) If there is one thing that deepfake teaches us, it’s that the expression “seeing is believing” deserves to be nuanced in light of how digital technologies are evolving. As we point out in our 2019 Trends Report, this technology manages to trick our senses by using deep learning, a form of artificial intelligence that is derived from machine learning and is used to model data with a high level of abstraction by training generative adversarial networks (or GANs). By applying this technology to image synthesis, deepfake makes it possible to superimpose images and videos created from pre-existing sources over</p>

other images and videos. It is therefore relatively easy to obtain credible results.

(c)

Rightfully criticized as a dangerous evolution of fake news, deepfake has prompted certain observers to highlight its potential dangers: diversion and violation of images, propagation of false information and 'dcredibilization' of audio and video evidence to name but a few. The rise of this technology also facilitated the development of an industry based on the detection of fake videos. However, although it is true that deepfake's development deserves to be cautiously explored with caution, the fact remains that it has major potential in terms of legitimate commercial exploitation—as have pointed out some of Hollywood's major special effects studios having expressed interest in this technology.

(d)

Although image synthesis has long been a part of the media and entertainment industry, the revolution announced by deepfake's apologists regards cost and access.

Until recently, switching faces with convincing results was reserved for the greatest of Hollywood productions. For example, following the "digital resurrection" of dead actor Peter Cushing allowing him to resume his role as Grand Moff Tarkin in the film *Rogue One: A Star Wars Story*, Industrial Light & Magic studio's creative director claimed that the effect had been extremely difficult and costly to develop. His team had had to resort to a mixed technique combining computer-assisted movement capture and image synthesis systems. By contrast, last year, a YouTuber known for using deepfake uploaded a short video in which he had replaced the face of actor Alden Ehrenreich with the face of a young Harrison Ford in several excerpts of the film *Solo: A Star Wars Story*. The YouTuber did it alone and practically at no cost! If there is a difference in the quality of both productions, creating such an effect on such a small scale would have been unthinkable a few years ago.

(e)

Given these technological developments, some observers emphasize that the future of the media and entertainment industry among several other industries will necessarily involve the hyperpersonalization of content and increasingly intense interactions between consumers and contents.

According to an Accenture report, 58% of consumers would change half or more of their expenses and choose to do business with vendors proposing personalized experiences without compromising their trust.

(f)

Having recognized the strategic importance of high-quality data for the future of image synthesis, some companies already offer actors and public figures high-priced "digital preservation" services that enable them to recreate their image in future productions and thus artificially preserve their youth.

(g)

Like the other technologies based on artificial intelligence, deepfake raises questions with respect to the rights and recourses that intellectual property rights holders have with respect to the training data used. On a more concrete level, if most of the deepfake models that are designed and deployed by simple users to copy the appearance of celebrities use

	<p>copyrighted works (including, for example, films or television series) as training data, the legality of this practice is far from certain.</p> <p>(h) Seeing as deepfake makes it possible to exploit the image of people who have died, its development risks feeding discussions on the right to one's image after death. Such discussions are already the object of public debate given the growth of the dead celebrity hologram industry. May it, however, be said that there are large grey areas and major legislative differences in that regard between different jurisdictions.</p> <p>(i) Deepfake's ramifications stem from political disinformation to cinema. However, regardless of what the years to come have in store for deepfake in political, legislative and commercial terms, chances are that it will become increasingly common on our screens.</p>
8	<p>Pomeroy, Robin. November 8, 2019. "James Dean comes back from the dead in Hollywood: what does that mean for us all?" <i>World Economic Forum</i>. Accessed on March 12, 2021, through https://www.weforum.org/agenda/2019/11/james-dean-cgi-deepfakes/.</p>
	<p>(a) Movie producers have announced that Dean will star in a Vietnam war film that is in pre-production, after they secured the rights to use his image.</p> <p>(b) But James Dean's return is thought to be the first time a dead actor will return to play an entirely new role.</p> <p>(c) As well as worrying about the 'authenticity' of CGI characters, actors may be concerned for their own careers. We all knew the rise of artificial intelligence in the Fourth Industrial Revolution would mean the end for many lines of work. But while bus drivers and manual workers might be at risk, surely computer code could never replace artists or actors?</p> <p>(d) The more alarming aspect of watching a dead actor play a new role is how convincing 'deepfake' technology is becoming. Pioneered as a way of superimposing movie stars' faces onto porn performers' bodies, deepfake technology can now be used to create a video of a person doing or saying something they have never done. Politicians can be deepfaked, and so can you. A single photograph can be enough to create a rudimentary deepfake video - although the makers of Finding Jack will have to do much more work to create a convincing James Dean on a big screen. The success or otherwise of that film is likely to be crucial when Hollywood moguls consider resurrecting other movie giants of the past.</p>
9	<p><i>The Economist</i>. July 5, 2018. "If AI made actors immortal—Performance anxiety The World If." Accessed on March 18, 2021, through</p>

<https://web.archive.org/web/20190203223825/https://www.economist.com/the-world-if/2018/07/05/performance-anxiety>.

(a)

Yet advances in special effects—and, increasingly, in artificial intelligence (AI)—are making it ever easier to manufacture convincing forgeries of human beings.

But the technology is improving fast, prompting a dozen AI researchers to place bets on whether a fake video will disrupt America's midterm elections later this year. (Tim Hwang, a Harvard academic, is overseeing the wager.)

(b)

Such digital actors would sit quietly in digital storage systems until their services were needed—there would be no need for luxury trailers, chefs and make-up people. Actors would be able to star in 100 different films in the same year, and carry on full and productive careers decades after their deaths. They would never age, and versions of them at different ages could appear in different films.

(c)

Close-ups are especially tough, says Mr Webber, because humans are hyper-sensitive to even small imperfections in how computer-generated faces look.

(d)

But this year's fake videos of Mr Obama and Mr Trump were made in a different way that is much cheaper and faster. They are so-called "deepfakes", created using a technique that, as the name suggests, uses a form of AI called deep learning to create fake videos.

(e)

The quality of this amateur face-swapping varies enormously; getting good results takes some fiddling, a powerful PC and a large selection of images of the face being superimposed. But none of those would be a problem for a Hollywood studio. As deepfake technology is incorporated into professional video-editing tools, face-swapping will surely become much faster, cheaper and more convincing than it is today.

Adobe, the maker of Photoshop, an image-manipulation tool, has developed a program called VoCo, which has been called "Photoshop for voice". It can simulate someone's voice saying almost anything, given 20 minutes of their speech to analyse. In 2017 Baidu, a Chinese technology firm, published details of a similar voice-cloning system called Deep Voice, which works with as little as three seconds' worth of training data. Other technology firms, including Google DeepMind, are doing similar work.

(f)

Hollywood, it seems, is opposed to today's unauthorised face-swapping, but wants to ensure it has a free hand to use such technology itself in the future.

(g)

Acting is already a winner-takes-all game, in which a small number of participants take the lion's share of the rewards. Being able to clone themselves, in effect, might allow big stars to take an even bigger share of the pie.

(h)

	<p>Those fed up with watching their hundredth Hugh Grant romantic comedy might have other options, however, if face-swapping technology can be made to work on the fly. Why not let viewers pick the actors they would like to see swapped into particular roles—or even let viewers splice themselves and their friends into their favourite films?</p> <p>(i)</p> <p>But the rise of immortal digital actors is the logical outcome as today's effects-heavy film-making techniques embrace the versatility of artificial intelligence. A trick that is currently resorted to only rarely could easily become a standard cinematic tool, like matte shots, green screens and CGI before it. Digital actors open up new possibilities in storytelling. But they also raise many new questions—and they will be able to answer them using any face, or voice, you like.</p>
10	<p>Toews, Rob. May 25, 2020. "Deepfakes Are Going To Wreak Havoc On Society. We Are Not Prepared." <i>Forbes</i>. Accessed on March 18, 2021, through https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/.</p>
	<p>(a)</p> <p>The State Farm ad was a benign example of an important and dangerous new phenomenon in AI: deepfakes. Deepfake technology enables anyone with a computer and an Internet connection to create realistic-looking photos and videos of people saying and doing things that they did not actually say or do.</p> <p>(b)</p> <p>While impressive, today's deepfake technology is still not quite to parity with authentic video footage—by looking closely, it is typically possible to tell that a video is a deepfake. But the technology is improving at a breathtaking pace. Experts predict that deepfakes will be indistinguishable from real images before long. "In January 2019, deep fakes were buggy and flickery," said Hany Farid, a UC Berkeley professor and deepfake expert. "Nine months later, I've never seen anything like how fast they're going. This is the tip of the iceberg."</p> <p>Today we stand at an inflection point. In the months and years ahead, deepfakes threaten to grow from an Internet oddity to a widely destructive political and social force. Society needs to act now to prepare itself.</p> <p>(c)</p> <p>The first use case to which deepfake technology has been widely applied—as is often the case with new technologies—is pornography. (...) Deepfake pornography is almost always non-consensual, involving the artificial synthesis of explicit videos that feature famous celebrities or personal contacts.</p> <p>From these dark corners of the web, the use of deepfakes has begun to spread to the political sphere, where the potential for mayhem is even greater.</p> <p>(d)</p> <p>It does not require much imagination to grasp the harm that could be done if entire populations can be shown fabricated videos that they believe are real. Imagine deepfake footage of a politician engaging in bribery or sexual</p>

	<p>assault right before an election; or of U.S. soldiers committing atrocities against civilians overseas; or of President Trump declaring the launch of nuclear weapons against North Korea. In a world where even some uncertainty exists as to whether such clips are authentic, the consequences could be catastrophic.</p> <p>(e) In a recent report, The Brookings Institution grimly summed up the range of political and social dangers that deepfakes pose: “distorting democratic discourse; manipulating elections; eroding trust in institutions; weakening journalism; exacerbating social divisions; undermining public safety; and inflicting hard-to-repair damage on the reputation of prominent individuals, including elected officials and candidates for office.”</p> <p>(f) In the words of Hani Farid, one of the world's leading experts on deepfakes: “If we can't believe the videos, the audios, the image, the information that is gleaned from around the world, that is a serious national security risk.”</p> <p>(g) But whether the video was real is almost beside the point. The larger lesson is that the emergence of deepfakes will make it increasingly difficult for the public to distinguish between what is real and what is fake, a situation that political actors will inevitably exploit—with potentially devastating consequences.</p> <p>(h) In a world in which seeing is no longer believing, the ability for a large community to agree on what is true—much less to engage in constructive dialogue about it—suddenly seems precarious.</p> <p>(i) One reason deepfakes have proliferated is the machine learning community's open-source ethos.</p> <p>(j) “We are outgunned,” said Farid. “The number of people working on the video-synthesis side, as opposed to the detector side, is 100 to 1.”</p> <p>(k) In the end, no single solution will suffice. An essential first step is simply to increase public awareness of the possibilities and dangers of deepfakes. An informed citizenry is a crucial defense against widespread misinformation.</p>
11	<p>Todd, Andrew. February 18, 2020. “Deepfakes: Visual Effects For The Post-Truth World.” <i>Birth Movies Death</i>. Accessed on March 18, 2021, through https://birthmoviesdeath.com/2020/02/18/deepfakes-visual-effects-for-the-post-truth-world.</p>
	<p>(a) Deepfakes differ from the 3D face replacement typically employed in Hollywood VFX studios. In fact, no 3D rendering is involved at all. Instead, a machine-learning algorithm analyses a target face's movement and lighting conditions, then replaces the face with another, sourced from whatever images the creator has at hand as reference. Feed it a wider variety of source images, and a more convincing result emerges. You can</p>

	<p>download the code and do it yourself; you only need time, a little technical nous, and the will to do it.</p> <p>(b) Like all video-adjacent technology, deepfakes first saw significant use in pornography.</p> <p>(c) Worse still are the potential implications for regular folks, whose potential appearance in deepfakes would likely be driven more by malice than comedy. As Carteris writes, "it is only a matter of time before a schoolteacher loses her job because of one of these deepfake videos." Angry exes, 8chan trolls, or anyone with an axe to grind can wield enormous power with this technology, especially if their targets live or work amongst the technologically unsavvy. The videos you'll find on YouTube right now may be funny, but to a more concerned mind, they feel like children naively playing with guns.</p> <p>(d) Deepfake technology is more accessible in 2020 than CGI was in 2008. It still takes work to produce convincing results, and even the best examples bear hallmarks of the process, but anyone can produce a deepfake. The legal and PR battles required to expose fraud or prove authenticity could become much more onerous, should deepfake tech grow a smidge more effective.</p> <p>(e) The ultimate destination of the Deepfake Express is more philosophically terrifying than any court case, identity theft, or political attack. If anything can be faked, everything could be faked - which, not coincidentally, perfectly reflects our current political situation. "Fake news" insults, coupled with outright lies, have achieved a broad goal of eroding public trust not just in reporting of truth, but in the truth itself. If everyone tells you conflicting stories, the instinctive mental response is to simply shut down. In this "fucked-up dystopia," as Peele puts it, what ends up mattering is what feels true. With no common understanding of fact and fiction, thought gets smaller, societies get more insular, and charismatic despots have the upper hand.</p> <p>(f) The technology isn't going away. We can't un-invent it. We can only learn to navigate this confusing world. As always, the best solution is to stay informed.</p>
12	<p>Stephan, Karl D. June 11, 2020. "That viral video of George Floyd's death shows why deepfakes are incredibly dangerous." <i>MercatorNet</i>. Accessed on March 18, 2021, through https://mercatornet.com/that-viral-video-of-george-floyds-death-shows-why-deepfakes-are-incredibly-dangerous/63626/.</p>
	<p>(a) A recent article on the Forbes website shows photos of four fairly normal-looking faces: two young women, a more mature woman, and a man. Thing is, none of these people exist. The faces are composites made</p>

	<p>with artificial-intelligence (AI) technology that produces what are called deepfakes: fictitious images that only sophisticated specialists armed with software can tell from the real thing, and sometimes they can't even tell either.</p> <p>(b) I've blogged about deepfakes before, but in our current volatile circumstances, the topic deserves some revisiting. Like almost anything having to do with AI these days, the quality of deepfakes is increasing as the computer horsepower needed to make them decreases. One way AI developers have found to improve the quality of deepfakes is to pit one program against another in what's called a "generational adversarial network," or GAN.</p> <p>(c) Without getting into the complicated issues of the relationship between videos and total reality, I will simply point out that the Floyd video created a hypersensitized public ready to be outraged by anything similar in nature, whether real or otherwise.</p> <p>(d) There are actors and institutions out there which would like nothing better than to sow discord in a city, a state, or a nation.</p> <p>(e) If someone fabricates a deepfake-assisted fraudulent incident that is primed to touch a sensitive nerve such as racism or the intense dislike that members of the public have toward a prominent political figure, it will not take much to set off reactions such as we have been seeing for the last couple of weeks since George Floyd's death.</p> <p>(f) As the technology of deepfakes advances, I expect that we will also learn how to deal with the fact that no matter how realistic a video looks, there is always the possibility that it was faked. And that thought is especially important to hang onto in distressed times such as these, and in critical times leading up to important elections.</p>
13	<p>Hao, Karen. October 10, 2019. "The biggest threat of deepfakes isn't the deepfakes themselves." <i>MIT review</i>. Accessed on March 18, 2021, through https://www.technologyreview.com/2019/10/10/132667/the-biggest-threat-of-deepfakes-isnt-the-deepfakes-themselves/.</p>
	<p>(a) The mere idea of AI-synthesized media is already making people stop believing that real things are real.</p> <p>(b) In the lead-up to the 2020 US presidential elections, increasingly convincing deepfake technology has led to fears about how such faked media could influence political opinion. But a new report from Deeptrace Labs, a cybersecurity company focused on detecting this deception, found no known instances in which deepfakes have actually been used in disinformation campaigns. What's had the more powerful effect is the knowledge that they could be used that way.</p> <p>(c)</p>

	<p>In the past two years, US tech companies and policymakers have focused almost exclusively on the first problem Ajder mentions: the ease with which the technology can make fake things look real. But it's the second that troubles experts more. While the barriers to creating deepfakes may be falling rapidly, calling the veracity of something into question requires no tech at all.</p> <p>(d) "It's an evolution of the claim that something is 'fake news,'" he says. "It gives another weapon to the powerful: to say 'It's a deepfake' about anything that people who are out of power are trying to use to show corruption, to show human rights abuses."</p> <p>(e) Finally, all the experts agree that the public needs greater media literacy. "There is a difference between proving that a real thing is real and actually having the general public believe that the real thing is real," [Aviv] Ovadya says. He says people need to be aware that falsifying content and casting doubt on the veracity of content are both tactics that can be used to intentionally sow confusion.</p> <p>(f) [Sam] Gregory cautions against placing too large a burden on news consumers, however. Researchers, platforms, and journalists should do as much of the work as possible to help make clear what is real and what is fake before news reaches the public.</p>
14	<p>Edwards, Benj. October 3, 2020. "How to stop deepfakes from tampering with history." <i>Fast company</i>. Accessed on March 19, 2021, through https://www.fastcompany.com/90549441/how-to-prevent-deepfakes.</p>
	<p>(a) Since deepfakes burst onto the scene a few years ago, many have worried that they represent a grave threat to our social fabric. Creators of deepfakes use artificial intelligence-based neural network algorithms to craft increasingly convincing forgeries of video, audio, and photography almost as if by magic. But this new technology doesn't just threaten our present discourse. Soon, AI-generated synthetic media may reach into the past and sow doubt into the authenticity of historical events, potentially destroying the credibility of records left behind in our present digital era.</p> <p>(b) In an age of very little institutional trust, without a firm historical context that future historians and the public can rely on to authenticate digital media events of the past, we may be looking at the dawn of a new era of civilization: post-history.</p> <p>(c) But in a world where information flows through social media faster than fact-checkers can process it, this disinformation sows enough doubt among those who don't understand how the technology works (and apathy among those who do) to destroy the shared cultural underpinnings of society—and trust in history itself. Even skeptics allow false information to slip through the cracks when it conveniently reinforces their worldview.</p> <p>(d)</p>

	<p>The technology is still in its early stages. And right now, detection is relatively easy, because many deepfakes feel “off.” But as techniques improve, it’s not a stretch to expect that amateur-produced AI-generated or -augmented content will soon be able to fool both human and machine detection in the realms of audio, video, photography, music, and even written text. At that point, anyone with a desktop PC and the right software will be able to create new media artifacts that present any reality they want, including clips that appear to have been generated in earlier eras.</p> <p>(e) In 2030, most people know that it’s possible to fake any video, any voice, or any statement from a person using AI-powered tools that are freely available. They generate many thousands of media fictions online every day, and that quantity is only going to balloon in the years to come.</p> <p>(f) And thanks to new AI-powered tools, the barriers to undetectably synthesizing every form of digital media are potentially about to disappear.</p> <p>(g) Let’s say that, in the future, there’s a core group of historians still holding the torch of enlightenment through these upcoming digital dark ages. They will need a new suite of tools and cultural policies that will allow them to put digital artifacts—real and synthesized alike—in context.</p> <p>(h) At some point, it’s likely that politicians in the U.S. and Europe will widely call to make deepfake tools illegal (as unmarked deepfakes already are in China). But a sweeping ban would be problematic for a free society. These same AI-powered tools will empower an explosion in human creative potential and artistry, and they should not be suppressed without careful thought. This would be the equivalent of outlawing the printing press because you don’t like how it can print books that disagree with your historical narrative.</p> <p>(i) With public trust in institutions diminished in the future, it’s possible that people (especially those not born in the 20th century who did not witness the media events firsthand) won’t believe officials who claim those physical artifacts are genuine if they don’t have the opportunity to study them themselves.</p> <p>(j) While we must not let disinformation destroy our understanding of the past, we also must not descend so far into fear that we stifle the creative tools that will power the next generation of art and entertainment. Together, we can build new tools and policies that will prevent digital barbarians from overwhelming the gates of history. And we can do it while still nourishing the civilization inside.</p>
15	<p><i>Tech Times</i>. July 2, 2020. “Disney is using deepfakes and facial recognition to bring back dead actors.” Accessed on March 19, 2021, through https://www.techtimes.com/articles/250776/20200702/disney-is-using-deep-fakes-and-facial-recognition-to-bring-back-dead-actors.htm.</p>

	<p>(a) Disney has unveiled its newest technology, deepfakes and facial recognition, to resurrect and revive dead actors and let fans see their stunts once more, The Verge and Screenrant reported. Deepfake is a term referring to the face-swapping algorithm that took the world by storm recently. The company believes these outputs will let fans see the stunts and acting prowess of amazing actors once more. With several actors the world lost in the past decade, this technology has become the bridge that will "immortalize" their onscreen performances. The Disney deepfakes are centered on realism and practicality.</p> <p>(b) What does it mean? With the facial details of this "dead" actor, the appearance is superimposed onto the chosen model. The method is similar to how deepfakes do it--though utilizing more sophisticated measures to maintain consistency and precision. This is not relatively new as there are companies that have ventured into these solutions for commercial purposes. However, it makes the project shine from above the rest, making it look more natural and at the same time, more compelling.</p> <p>(c) Digitally-created facial images are usually met with challenges when connecting them with the real world. However, this facial technology can address the problem of showcasing deceased actors in current works of art or in movies in the making. Among the adjustments are geometric alterations, contrast, and motion. The audio deepfake should also go along with this. There have already been instances of deceased actors reprising their roles with the use of this technology.</p>
16	<p>Christian, Jon. February 1, 2018. "Experts fear face swapping tech could start an international showdown." <i>The outline</i>. Accessed on March 19, 2021, through https://theoutline.com/post/3179/deepfake-videos-are-freaking-experts-out.</p>
	<p>(a) Video forensic specialists are worried Deepfakes could have national security repercussions.</p> <p>(b) And now bad actors have access to Deepfakes, which is becoming more advanced all the time. Shortly after Motherboard first reported on an early, difficult-to-use version of the software, an expert predicted that a more user-friendly version could appear within a year or two. In reality, it took just two months for a coder on Reddit to assemble an app that can produce face-swapped videos with no advanced machine learning knowledge.</p> <p>(c) And in the long term, [Hany] Farid has a structural concern about artificial intelligence research. Most work on machine learning, he said, has focused on what the technology can create — but if the internet of the future could turn out to be a wash of computer-generated hoaxes, maybe more resources should be dedicated to understanding how to defend against the</p>

	<p>material they produce. (...) “The reality is that the number of people working on the forensics side, like me, is relatively small compared to the number of people working on the other side,” Farid said. “We are greatly outnumbered and out-resourced. Google is not developing forensic techniques. Facebook is not developing forensic techniques. It’s a bunch of academics. We’re outgunned.”</p>
17	<p>Joshi, Shamani. August 21, 2020. “Face swap app Reface might make deepfake issue worse.” <i>Vice</i>. Accessed on March 19, 2021, through https://www.vice.com/en/article/wxqkbn/viral-reface-app-going-to-make-deepfake-problem-worse.</p>
	<p>(a) The app that allows you to superimpose your selfie onto videos of celebrities might be endlessly entertaining, but its quick, slick results can lead to sinister outcomes too.</p> <p>(b) In case you didn’t already know, Deepfake AI, the technology that replaces an existing image or video with someone else’s likeness using a machine learning technology called GAN (generative adversarial network), is categorised by criminology experts as the biggest technological threat humanity will face moving forward—even bigger than killer robots. That is because they’re perceived to be harder to defeat as opposed to other AI-enabled future crimes, and could lead to high harm or profit too. In their short span of existence, deepfakes have already been abused endlessly to manipulate everything, from election campaigns to non-consensual pornographic imagery, leading to a landmine of misinformation on the internet that is often difficult to detect and blurring the lines of consent.</p> <p>(c) And though the app [Reface] doesn’t allow for users to upload their own videos or full-body swaps, such technology already exists anyway and might be exploited by other apps looking to replicate Reface’s commercial success. That the very root of deepfakes lies in diligently cranking out fictional footage of women through a machine-learning meat grinder and posting them to public porn forums is important to remember as well.</p> <p>(d) Let’s face it: apps like Reface, which lure you in with their simple technology and capture attention by integrating your likeness into trendy video formats, may not set out with the intention to cause harm. But as we evolve into an increasingly digital existence, with everything from our workouts to our self-isolation outfits splashed across various social media for all the world to see, it’s important to remember that the lines between what is real and what can be made to seem real are just one app download away.</p>
18	<p>Roose, Kevin. March 4, 2018. “Here come the fake videos, too.” <i>The New York Times</i>. Accessed on March 19, 2021, through https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html.</p>

(a)

Artificial intelligence video tools make it relatively easy to put one person's face on another person's body with few traces of manipulation.

(b)

The video, which appeared on the online forum Reddit, was what's known as a "deepfake" — an ultrarealistic fake video made with artificial intelligence software. It was created using a program called FakeApp, which superimposed Mrs. Obama's face onto the body of a pornographic film actress. The hybrid was uncanny — if you didn't know better, you might have thought it was really her.

(c)

Until recently, realistic computer-generated video was a laborious pursuit available only to big-budget Hollywood productions or cutting-edge researchers.

(d)

Deepfakes are one of the newest forms of digital media manipulation, and one of the most obviously mischief-prone. It's not hard to imagine this technology's being used to smear politicians, create counterfeit revenge porn or frame people for crimes. Lawmakers have already begun to worry about how deepfakes could be used for political sabotage and propaganda. Even on morally lax sites like Reddit, deepfakes have raised eyebrows. Recently, FakeApp set off a panic after Motherboard, the technology site, reported that people were using it to create pornographic deepfakes of celebrities. Pornhub, Twitter and other sites quickly banned the videos, and Reddit closed a handful of deepfake groups, including one with nearly 100,000 members.

(e)

Some users on Reddit defended deepfakes and blamed the media for overhyping their potential for harm. Others moved their videos to alternative platforms, rightly anticipating that Reddit would crack down under its rules against nonconsensual pornography. And a few expressed moral qualms about putting the technology into the world. Then, they kept making more. The deepfake creator community is now in the internet's shadows. But while out in the open, it gave an unsettling peek into the future. "This is turning into an episode of Black Mirror," wrote one Reddit user. The post raised the ontological questions at the heart of the deepfake debate: Does a naked image of Person A become a naked image of Person B if Person B's face is superimposed in a seamless and untraceable way? In a broader sense, on the internet, what is the difference between representation and reality? The user then signed off with a shrug: "Godspeed rebels."

(f)

What I learned is that making a deepfake isn't simple. But it's not rocket science, either. (...) I also sent Derpfakes, my outsourced Reddit expert, several video options to choose from. (...) I reached out to the anonymous creator of FakeApp through an email address on its website. I wanted to know how it felt to create a cutting-edge A.I. tool, only to have it gleefully co-opted by ethically challenged pornographers. A man wrote back, identifying himself as a software developer in Maryland. Like Derpfakes, the man would not give me his full name, and instead went by his first initial, N.

	<p>He said he had created FakeApp as a creative experiment and was chagrined to see Reddit’s deepfake community use it for ill.</p> <p>(g) [N. says,] “I knew this was brilliant tech that should be accessible to anyone who wants to play around with it; I’ve decided I don’t think it’s right to condemn the technology itself — which can of course be used for many purposes, good and bad.” (...) “It’s precisely the things that make them so powerful and useful that make them so scary,” he [N.] said. “There’s really no limit to what you can apply it to with a little imagination.”</p> <p>(h) Online misinformation, no matter how sleekly produced, spreads through a familiar process once it enters our social distribution channels. The hoax gets 50,000 shares, and the debunking an hour later gets 200. The carnival barker gets an algorithmic boost on services like Facebook and YouTube, while the expert screams into the void.</p> <p>(i) So, O.K. Here I am, telling you this: An A.I. program powerful enough to turn Michelle Obama into a pornography star, or transform a schlubby newspaper columnist into Jake Gyllenhaal, is in our midst. Manipulated video will soon become far more commonplace.</p> <p>(j) And there’s probably nothing we can do except try to bat the fakes down as they happen, pressure social media companies to fight misinformation aggressively, and trust our eyes a little less every day. Godspeed, rebels.</p>
19	<p>Schwartz, Oscar. November 12, 2018. “You thought fake news was bad? Deep fakes are where truth goes to die.” <i>The Guardian</i>. Accessed on March 19, 2021, through https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth.</p>
	<p>(a) The party’s [sp.a, Belgium] communications team had clearly underestimated the power of their forgery, or perhaps overestimated the judiciousness of their audience. Either way, this small, left-leaning political party had, perhaps unwittingly, provided a deeply troubling example of the use of manipulated video online in an explicitly political context.</p> <p>(b) It was a small-scale demonstration of how this technology might be used to threaten our already vulnerable information ecosystem – and perhaps undermine the possibility of a reliable, shared reality.</p> <p>(c) When Danielle Citron, a professor of law at the University of Maryland, first became aware of the fake porn movies, she was initially struck by how viscerally they violated these women’s right to privacy. But once she started thinking about deep fakes, she realized that if they spread beyond the trolls on Reddit they could be even more dangerous. They could be weaponized in ways that weaken the fabric of democratic society itself.</p>

(d)

[Danielle] Citron, along with her colleague Bobby Chesney, began working on a report outlining the extent of the potential danger. (...) both scholars became increasingly concerned that the proliferation of deep fakes could catastrophically erode trust between different factions of society in an already polarized political climate. In particular, they could foresee deep fakes being exploited by purveyors of “fake news”. Anyone with access to this technology – from state-sanctioned propagandists to trolls – would be able to skew information, manipulate beliefs, and in so doing, push ideologically opposed online communities deeper into their own subjective realities.

(e)

At the same time, [Tim] Hwang acknowledges that as deep fakes become more realistic and easier to produce in the coming years, they could usher in an era of forgery qualitatively different from what we have seen before.

(f)

Christian Theobalt, a researcher involved in the study [Deep video portraits], told me via email that he imagines deep video portraits will be used most effectively for accurate dubbing in foreign films, advanced face editing techniques for post-production in film, and special effects. In a press release that accompanied the original paper, the researchers acknowledged potential misuse of their technology, but emphasized how their approach – capable of synthesizing faces that look “nearly indistinguishable from ground truth” – could make “a real difference to the visual entertainment industry”.

Hany Farid, professor of computer science at the University of California, Berkeley, believes that although the machine learning-powered breakthroughs in computer graphics are impressive, researchers should be more cognizant of the broader social and political ramifications of what they’re creating. “The special effects community will love these new technologies,” Farid told me. “But outside of this world, outside of Hollywood, it is not clear to me that the positive implications outweigh the negative.”

(g)

This, he [Hany Farid] explains, is due to the flexibility of machine learning. “All the programmer has to do is update the algorithm to look for, say, changes of color in the face that correspond with the heartbeat, and then suddenly, the fakes incorporate this once imperceptible sign.” (For this reason, Farid chose not to share some of his more recent forensic breakthroughs with me. “Once I spill on the research, all it takes is one asshole to add it to their system.”)

(h)

Although [Hany] Farid is locked in this technical cat-and-mouse game with deep fake creators, he is aware that the solution does not lie in new technology alone. “The problem isn’t just that deep fake technology is getting better,” he said. “It is that the social processes by which we collectively come to know things and hold them to be true or untrue are under threat.”

20	<p>Chandler, Simon. March 9, 2020. "Deepfakes are net positive for humanity." <i>Forbes</i>. Accessed on March 19, 2021, through https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/?sh=61cfad32f84f.</p>
	<p>(a) Deepfakes are your friend. Yes, deepfake technology has understandably become notorious in the wake of deepfake porn videos and the threat deepfakes seemingly pose to politics. However, the ability to generate realistic simulations using artificial intelligence will, on the whole, be only a positive for humanity. Whether recreating long-dead artists in museums or editing video without the need for a reshoot, deepfake technology will allow us to experience things that no longer exist, or that have never existed. And aside from having numerous applications in entertainment and education, it's being increasingly used in medicine and other areas.</p> <p>(b) Much more realistically, deepfake technology will play an increasingly constructive role in recreating the past and in envisioning future possibilities. (...) What was most novel about this is that, by using deepfake technology, you can automatically generate video reports personalised for each individual news viewer.</p> <p>(c) The examples above [cases of "resurrecting" Salvador Dali and John F. Kennedy for edutainment purposes] show how deepfakes can serve to help bring history and art 'alive' for a wider audience. And if this helps to get thousands or millions of people interested in art and history, then the world can only benefit.</p> <p>(d) For now, of course, deepfakes will continue to have a bad reputation, given that they're most well-known for (potentially) threatening democracy. Still, if we can educate ourselves to trust only videos delivered by reputable sources (in the same way we trust only certain sources of text-based news), we'll soon find that the good of deepfakes will more than outweigh the bad.</p>
21	<p>Jaiman, Ashish. August 14, 2020. "Positive use cases of deepfakes." <i>Towards data science</i>. Accessed on March 19, 2021, through https://towardsdatascience.com/positive-use-cases-of-deepfakes-49f510056387.</p>
	<p>(a) Technology is an excellent enabler and can be very empowering. New ideas and capabilities for empowerment have emerged because of the advancements in data science and Artificial intelligence. AI can create possibilities and opportunities for all people, regardless of who they are and how they listen, speak, or communicate.</p> <p>(b)</p>

	<p>I firmly believe that as technology evolves and societies adjudicate the deepfake phenomenon, most of the use cases of deepfakes will be to improve people's lives and empower communities and institutions.</p> <p>(c) Artificial intelligence can build tools to hear, see, and soon with Artificial General Intelligence (AGI), reason with increasing accuracy. AI-Generated Synthetic media can help make the accessibility tools smarter and, in some cases, even affordable and personalizable, which can help people augment their agency and gain independence. Technology, specifically AI-based tools, can bring accessible solutions to all.</p> <p>(d) Deepfakes can help an educator to deliver innovative lessons that are far more engaging than traditional visual and media formats.</p> <p>(e) Deepfakes can democratize the costly VFX technology as a powerful tool for independent storytellers at a fraction of the cost.</p> <p>(f) Synthetic media can help human rights activists and journalists to remain anonymous in dictatorial and oppressive regimes. Using technology to report out atrocities on traditional or social media can be very empowering for citizen journalists and activists. Deepfake can be used to anonymize voice and faces to protect their privacy.</p> <p>(g) In April last year, a health charity partnered with David Beckham to produce a video and voice campaign to help end malaria. In the "malaria no more" campaign, Beckham spoke nine languages seamlessly for public appeal. The social campaign was a great example of using deepfakes to broaden the reach of a public message. Synthesia, the company behind the Beckham video, and VOCALiD, a voice startup, create tools to natively localize video and audio content for learning tools, brand marketing, audience engagement, customer service, and public messaging to broaden the reach and amplification of the message.</p> <p>(h) Deepfakes can be used by influencers to broaden their reach and increase their audience. Using deepfakes, a brand can reach many customers with highly targeted and personalized messaging.</p> <p>(i) AI-Generated synthetic media can be very empowering and a great enabler. Deepfakes can give people a voice, purpose, and ability to impact scale and speed.</p> <p>(j) However, as access to synthetic media technology increases, so does the risk of exploitation. Deepfakes can be used to damage reputations, fabricate evidence, defraud the public, and undermine trust in democratic institutions.</p>
22	<p>Kishore, Aseem. May 23, 2019. "What is a Deepfake and How Are They Made?" <i>Online Tech Tips</i>. Accessed on March 19, 2021, through</p>

	<p>https://www.online-tech-tips.com/computer-tips/what-is-a-deepfake-and-how-are-they-made/.</p>
	<p>(a) Well, the answer isn't "witchcraft", but a technology that has been dubbed "deepfake" and it's causing quite a ruckus in various industries and communities.</p> <p>(b) There are a number of applications that allow people to make deepfakes. FakeApp was the first app we know of aimed at giving normal people a shot at making deepfake content. The website it now defunct and finding a copy is not easy at all.</p> <p>(c) Our goal is to help you understand the technology, since you are bound to encounter it more and more as time goes by. With that being said, these are the broad phases of creating a deepfake with DeepFaceLab.</p> <p>(d) Deepfakes can be used to create hoaxes, such as a video of a president saying crazy things. It can be used to insert people into pornographic films with the purpose of harassing or otherwise harming them.</p> <p>(e) Now that this technology exists and works as well as it does, it means we have to look at media like videos in a whole new light. If someone circulates a video on social media of a famous person or politician saying or doing something controversial, you'll first have to ask if the video is even real. Most poorly-done deepfakes are obviously fake, once you know what to look for and have seen a few. However, in some cases even a trained eye may have trouble telling that some sort of CG manipulation is happening and as the tech improves it will eventually become impossible.</p>
23	<p>Pangburn, Dj. September 21, 2019. "You've been warned: Full body deepfakes are the next step in AI-based human mimicry." <i>Fast company</i>. Accessed on March 19, 2021, through https://www.fastcompany.com/90407145/youve-been-warned-full-body-deepfakes-are-the-next-step-in-ai-based-human-mimicry.</p>
	<p>(a) This developing branch of synthetic media technology has commercial applications—but also has the potential to disrupt elections and spread disinformation. Within the field of deepfakes, or "synthetic media" as researchers call it, much of the attention has been focused on fake faces potentially wreaking havoc on political reality, as well as other deep learning algorithms that can, for instance, mimic a person's writing style and voice. But yet another branch of synthetic media technology is fast evolving: full body deepfakes.</p> <p>(b) In August 2018, University of California Berkeley researchers released a paper and video titled "Everybody Dance Now," demonstrating how deep learning algorithms can transfer a professional dancers' moves onto the</p>

bodies of amateurs. While primitive, it showed that machine learning researchers are tackling the more difficult task of creating full body deepfakes. Also in 2018, a team of researchers led by Dr. Björn Ommer of Heidelberg University in Germany published a paper on teaching machines to realistically render human movements. And in April of this year, the Japanese artificial intelligence company Data Grid developed an AI that can automatically generate whole body models of nonexistent persons, identifying practical applications in the fashion and apparel industries.

(c)

While it's clear that full body deepfakes have interesting commercial applications, like deepfake dancing apps or in fields like athletics and biomedical research, malicious use cases are an increasing concern amid today's polarized political climate riven by disinformation and fake news. For now, full body deepfakes aren't capable of completely fooling the eye, but like any deep learning technology, advances will be made. It's only a question of how soon full body deepfakes will become indistinguishable from the real.

(d)

"With deepfake images and videos, we've essentially democratized CGI technology," he [Hany Farid] says. "We've taken it out of the hands of Hollywood studios and put it in the hands of YouTube video creators."

(e)

[Björn Ommer:] "The research community itself has moved in a direction—and this is very much appreciated by much of the community that is responsible for a lot of this steady progress that we see—where the algorithms are easily available, like on Github and so on. So, you can just download the most recent code from some paper, and then, without much knowledge of what's under the hood, just apply it." (...) "Not every person will be able to create a 'blockbuster deepfake.' But, given more time, Ommer says money will no longer be an issue in terms of computational resources, and the applicability of software will also become much easier.

(f)

"It's not the big shots, the big politicians, and the big famous guys who are the most threatened," says [Tom] Van de Weghe. "It's the normal people—people like you, me, female journalists, and sort of marginalized groups that could become or are already becoming the victims of deepfakes." (...) "There is money to be earned with deepfakes," says Van de Weghe. "People will order it. So, a government doesn't have to create a deepfake—they just have to contact a person who is specialized in deepfakes to create one."

(g)

[Hany] Farid thinks a combination of social media and deepfake videos, whether of faces or full bodies, could easily wreak havoc. Social media companies are largely unable or unwilling to moderate their platforms and content, so deepfakes can spread like wildfire.

(h)

The technology is evolving fast, spurred by university research like "Everybody Dance Now" and private sector initiatives such as Zao to monetize deepfakes.

(i)

	<p>I think it can be helpful to consider how we've [legislatively] dealt with other industries that externalize large costs to society while privately claiming the profits (such as industrial pollution, big tobacco, and fast food/junk food)," [Rachel] Thomas adds.</p> <p>(j) A blanket law to stop deepfakes would be misguided, says [Giorgio] Patrini. Instead, he advocates government support for synthetic media applications that benefit society, while funding research into creating tools to detect deepfakes and encouraging startups and other companies to do the same.</p> <p>(k) [Björn Ommer:] "Government agencies and so on who have interviewed me obviously see their share in this responsibility. Companies say and probably—in the interest of their stockholders—have to say that they see their responsibility; but, we all know how they have handled this responsibility up until now."</p>
24	<p><i>Think automation.</i> "Yes, positive deepfake examples exist." Accessed on March 19, 2021, through https://www.thinkautomation.com/bots-and-ai/yes-positive-deepfake-examples-exist/.</p>
	<p>(a) Artificial intelligence technology, to many, is scary. It's packed full of ethical questions and uncertain future applications. Perhaps one of the more alarming functionalities enabled by AI advancement is the creation of deepfakes.</p> <p>(b) Deepfakes undermine our trust in the information we see. Plus, they present plenty of opportunities for mischief and malicious use. It's hard to see how we could hope to show such functionality in a positive light — or how anyone could ever come to trust it. But, as with any tool, there are beneficial uses for deepfakes too. Here, we explore some oft-forgotten positive deepfake examples. How can deepfake technology be a force for good?</p> <p>(c) Deepfakes are convincing content depicting artificially constructed events. (...) This most often affects celebrities and predominant public figures. Their likeness is used, but the content isn't real. They aren't truly saying those words, or physically enacting those movements.</p> <p>(d) Aggravating this alarm is the fact that plenty of stories about deepfakes discuss deepfake pornography, fake news and fraud. But where there's bad, there's good. Positive deepfake examples also exist.</p> <p>(e) Deepfake technology holds positive potential for education. It could revolutionise our history lessons with interactivity. It could preserve stories and help capture attention. How? With deepfake examples of historical figures. (...) As deepfake technology advances, this kind of virtual history could become achievable on a much wider scale.</p> <p>(f)</p>

	<p>Positive deepfake examples also show how the technology can make language barriers (and bad dubbing) a thing of the past.</p> <p>(g) So, with the help of positive deepfakes, we can better share thoughts, films and other creative works on a worldwide basis. Even those with lower budgets. This stands to improve the diversity of our entertainment and content consumption.</p> <p>(h) In fact, there are a lot of positive deepfake examples for use in the entertainment industry. Deepfakes can keep film characters consistent. (i) This also demonstrates another positive use of deepfake technology: ageing and de-ageing.</p> <p>(i) When it comes to deepfakes, there's plenty of potential for the technology, good and bad. Naturally, the desire is to see more positive deepfake examples, rather than fraud and negativity.</p>
25	<p><i>BBC Bitesize</i>. "Deepfakes: What are they and why would I make one?" Accessed on March 19, 2021, through https://www.bbc.co.uk/bitesize/articles/zfkwcqt.</p>
	<p>(a) The vast majority of deepfake videos available online are created by amateurs, and they are unlikely to fool anyone. The process that creates deepfakes often leaves very visible artifacts around a person's face. Blurring or flickering, for example, are very common, especially when the face changes angles very rapidly.</p> <p>(b) That being said, there are many examples of videos that are virtually perfect, or that might seem real to the untrained eye.</p> <p>(c) The vast majority of deepfakes circulating the Internet are featuring celebrities and politicians - something that has worried many. The process of creating a deepfake video requires many pictures of a person's face, which makes celebrities particularly vulnerable.</p> <p>(d) Deepfakes have often been used to create pornographic content without the direct consent of the people involved. It is very important to understand that that is not OK, even if the final result is not believable. Bullying does not need to be photo-realistic to be effective.</p> <p>(e) With these techniques becoming more accessible and reliable, it is not hard to imagine a future in which deepfakes will be just another tool in video editing programs.</p> <p>(f) There are so many positive ways in which deepfakes could be used to make our life better - and that is why this is a technology that needs to be understood, not just feared.</p> <p>(g)</p>

Others are worried that, without any proper regulations, many late actors and historical figures might be 'digitally resurrected' without the explicit consent of their relatives. (...) On the other hand, there is a fear that stricter rules might shine a bad light on deepfakes, which could slow down research in many fields related to AI. But with technology getting better and better, we are heading for a future in which videos cannot be so easily trusted.

(h)

Creating a deepfake video is surprisingly easy, although it is a very tedious process! The technology is easy to understand, and anyone with a decent laptop and some basic understanding of programming can make their own. Making a realistic, seamless deepfake videos, however, requires a much deeper knowledge and commitment.

If you are up for the challenge, there are two things that are absolutely necessary: a powerful GPU (graphics processing unit) and consent of the people in the footage and/or images you are manipulating.