

Małgorzata Kiełtyka / Jakub Gładkowski / Anna Bednarska / Mariusz Bujacz / Kamil Trzaskoś

Voice Cloning as a Global New Technology and its Challenges for EU and Polish Law

New technologies, such as Siri, Cortana or Google, process our personal data for various reasons. The paper focuses on voice, its protection and possible issues arising around voice cloning technologies. The authors present the voice as a personal right, the issue of personal data processing, the consequences related to penal liability in case of violations related to voice cloning, examples of voice cloning software and, voice biometrics.

Category: Articles

Region: Poland

Field of law: LegalTech

Citation: Małgorzata Kiełtyka / Jakub Gładkowski / Anna Bednarska / Mariusz Bujacz / Kamil Trzaskoś, Voice Cloning as a Global New Technology and its Challenges for EU and Polish Law, in: Jusletter IT 18 May 2017

Contents

1. Introduction
2. Background of Voice Cloning Technology
3. Voice as a Personal Right and its Protection in Polish Jurisdiction
4. Voice Cloning in the Light of Penal Liability (in the Polish Copyright Law)
5. Voice Cloning Software and its Risks
6. Data Protection in the Light of Voice Cloning
7. Voice Biometrics vs Voice Cloning
8. Conclusions

1. Introduction

[Rz 1] Siri, Cortana, Google and other applications use human voice to do a variety of things, e.g. searching for information, sending emails, calling somebody. Voice-based technologies are increasingly applied in legal environment and legal services, for example in legal advice rendered online and in legal translations. At the same time, new applications of innovative technologies caused the necessity to define the approach to privacy issues anew. The cases of Edward Snowden and Julian Assange showed us how meaningful privacy and its protection is and made us realize the excessive amount of personal data processed and stored daily. This is why privacy and its protection will soon become one of the most important personal rights. The issue of voice protection comes to the fore in this context. Voice is, obviously, a personal right. What is more, voice is becoming a tool used by most applications both for mundane activities as well as more complex ones, like ROSS AI operating on IBM's Watson, which can do legal research and is learning to understand law with every research conducted by it. What if it was possible for such applications as Watson to use the voice of a specific lawyer and, with the use of voice sample, produce speech of a different content, for example in the form of legal advice? Well, practically it is possible, since last November Adobe presented Adobe VoCo to the world, which (when having a voice sample) is able to read various content differing from the content sampled. The present article will try to shed some light to the issue of the risk involved with voice cloning technology in legal environment and will analyse whether law can adequately protect human voice as a personal right.

2. Background of Voice Cloning Technology

[Rz 2] Voice cloning technology is based on copying and reusing of recorded speech. In the future, such software will be able to record voice samples and, afterwards, produce an infinite number of combined syllables leading to an unlimited number of sentences, without the participation of the human being that provided the voice sample. In reference to the latest developments, we may be first to witness the creation of such software for commercial use. The first project worth mentioning is Google Deep Mind's WaveNet. It is a deep neural network for generating raw audio waveforms, including speech and music. WaveNet has outperformed other text-to-speech systems, but this product has not yet been declared to be available for consumers.¹ From this point of view, it is of importance to mention Adobe Project VoCo, presented during the Adobe MAX 2016 Sneak Peeks. It is a software which is able to create a voice model of the speaker

¹ AARON VAN DEN OORD / KAREN SIMONYAN / NAL KALCHBRENNER / SANDER DIELEMAN / ORIOL VINYALS / ANDREW SENIOR / HEIGA ZEN / ALEX GRAVES / KORAY KAVUKCUOGLU, WaveNet: A Generative Model for Raw Audio, 19 September 2016, www.arxiv.org/pdf/1609.03499.pdf (all internet addresses last accessed 18 April 2017).

from an earlier given voice sample of 20 minutes duration by said speaker.² VoCo can construct new words and sentences which did not occur in the provided recordings.³ Such potential of said software, with the plans to release VoCo to the consumer market, raises considerable concerns, also legal ones, in respect of data and privacy protection.

3. Voice as a Personal Right and its Protection in Polish Jurisdiction

[Rz 3] In order to specify if the European or the Polish law can adequately protect the use of voice technology-based applications and word-building software, we need to indicate the legal status of the human voice at first. From the legal point of view, the human voice should generally be classified as a personal right and, more specific, as a non-pecuniary property of every human being connected with his individual existence, which is effective against everyone, inalienable and not inherited. Polish provisions (art. 23–24 of the Polish Civil Code) include a sample and open catalogue of personal rights and their protection, irrespective of other regulations. Polish jurisprudence and the majority of law practitioners⁴ express approval for the most essential judgement in this regard, which has been delivered by the Polish Court of Appeals in Gdańsk on 21 June 1991 (case citation: I ACr 127/91, LEX), where the Court acknowledged that the voice shall be regarded as a personal right (as defined in art. 23 of the Polish Civil Code)⁵ and protected pursuant to art. 24 of the Polish Civil Code.⁶ Voice serves the same purpose as a human image, namely: identification. It is an element of appearance, given that it relates to individual voice alteration, pitch, sound and the ways someone speaks, i.e. intonation and characteristic words. The violation of this right could occur, e.g., by duplication of voice records or their modification and, what is more, by imitation of distinctive voices, if it could be demonstrated that the above mentioned use was intended to deceive listeners in regard to the identity of the person speaking.⁷ In case of acknowledging that recognizing a person within the sphere of a sound is possible just as well as through an external image, principles related to images apply analogically to voices, provided that the voice is protected as a separate personal right.⁸

[Rz 4] In accordance with the latter provision, the one whose personal right is threatened by the activity of third party may demand this activity to cease, unless it is legitimate (art. 24 par. 1 of the Polish Civil Code). Nevertheless, there are also legal experts who intend to recognize the voice not as a separate personal right but rather as a part of human image or as «audio-image» / «sound-image» that makes it possible to identify a person by sense of hearing. If this view is adopted, then the voice is protected not only on the basis of the Polish Civil Code but also within the framework of copyright law (art. 24 par. 3 of the Polish Civil Code).

² Official live presentation during the Adobe MAX 2016 Sneak Peeks, co-hosted by Jordan Peele, www.youtube.com/watch?v=I3l4XLZ59iw.

³ SEBASTIAN ANTHONY, Adobe demos «photoshop for audio,» lets you edit speech as easily as text, arsTECHNICA, 11 July 2016, www.arstechnica.com/information-technology/2016/11/adobe-voco-photoshop-for-audio-speech-editing.

⁴ JANUSZ BARTA / RYSZARD MARKIEWICZ / ANDRZEJ MATLAK, *Media Law*, LexisNexis, Warsaw 2005; JUSTYNA BALCARCZYK, The right to image and its commercialization, *Oficyna Wolter Kluwer Business*, Warsaw 2009, pp. 52–54; JUSTYNA BALCARCZYK, Voice right – outline of basis issues, *Zeszyty Naukowe Uniwersytetu Jagiellońskiego* 2010/2/115–126, LEX; MAKSYMILIAN PAZDAN, Commentary on Article 23 of the Civil Code, in: Krzysztof Pietrzkowski (ed.), *Civil Code. Commentary on Articles 1–449*[10], Volume 1, Legalis.

⁵ Art. 23 of the Polish Civil Code dated on 23 April 1964, *Journal of Laws* No 16.94 as amended.

⁶ Art. 24 of the Polish Civil Code dated on 23 April 1964, *Journal of Laws* No 16.94 as amended.

⁷ MAŁGORZATA PYZIAK-SZAFNICKA / PAWEŁ KSIĘŻAK, *Civil Code – Comment. General Part. Edition II*, LEX, 2014.

⁸ JUSTYNA BALCARCZYK, The right to image and its commercialization, *Oficyna Wolter Kluwer Business*, Warsaw 2009, pp. 52–54.

[Rz 5] The protection system of the personal rights should be deeply analysed in regard to new technologies based on the use of human voice, since new ways of using (i.e. for online legal advice) or modifying it (i.e. in order to circumvent voice recognition technologies used by banks while making payment orders) could not be protected adequately enough.

[Rz 6] Under the Polish Civil Code, the conditions for legal protection of the voice as a personal right are to be viewed as a breach or threat of a breach of personal rights and unlawfulness of such breach or threat. The person who provides his or her voice may therefore demand, amongst others, that the consequences of said breach are removed and that monetary compensation is paid under this title. In this context, the controversy arises, whether – given the situation that a person voluntarily and in consent provides a voice sample – the element of unlawfulness can be demonstrated when the specific software clones the voice in an unintended manner. Accordingly, the open question is whether the means mentioned above provide sufficient protection in this respect. It appears, that nowadays new technologies use subjects of personal rights (protected by given legal methods) in pioneer ways, so that the effects of those activities require new concepts, i.e. applications editing an attorney's voice (such as VoCo by Adobe) could be used for providing unfounded legal advice and therefore we not only deal with a breach of the personal right related to the voice but also related to the image, scientific activity, freedom of conscience or other implied legal consequences. On the other hand, VoicePass technology, constructed by the Polish University of Science and Technology in Cracow, which is able to identify our voice and allows to verify our identity i.e. in banks, insurance offices or authority bodies⁹ is not only a great invention and simplifying various official procedures but also a potential risk of violating our personal data.

4. Voice Cloning in the Light of Penal Liability (in the Polish Copyright Law)

[Rz 7] It has to be considered that manufacturers of computer programs which allow voice cloning will provide adequate protection in the form of tags, digital watermarks or any other forms, so that it can be showed that somebody's voice being used in bad faith has been created by the program. But what if somebody circumvents effective technical devices applied to protect the software in order to remove digital watermarks and to use somebody's voice unlawfully? This has to be viewed as cracking and the accountable person can be treated as cracker or hacker. Polish Copyright Law indicates penalties in its art. 118 para. 1, stating that «anyone who produces devices or components of devices for the purpose of unauthorised removal or circumvention of effective technical devices applied to protect a work or the subject matter of related rights against replaying, copying or reproduction or trades in such devices or components of such devices, or advertises their sale or rental, is liable to a fine, restriction of personal liberty or imprisonment for up to 3 years». In turn, para. 2 of said article sets forth that «anyone who owns, stores or uses devices or components of devices as referred to in paragraph 1, is liable to a fine, restriction of personal liberty or imprisonment for up to a year».

[Rz 8] First of all, the term «effective technical devices» must be clarified: it means that the introduced technical security is objectively capable of fulfilling its function and – without it being removed or bypassed – replaying, copying or reproducing are impossible.¹⁰ The problem arising from the wording of the quoted

⁹ Polish Press Agency, You know your neighbour by his voice, 31 March 2014, <http://naukawpolsce.pap.pl/aktualnosci/news,399802,poznasz-blizniego-po-glosie-jego.html>.

¹⁰ ZBIGNIEW CWIĄKALSKI, Commentary on Article 118(1) of the Copyright Law, in: Barta Janusz / Markiewicz Ryszard (eds.), Copyright Law. Commentary, Volume 5, LEX no. 8545, 2011.

legal provision concerns computer programs and whether this provision also applies to computer programs striving for illegal neutralization of security. It is worth pointing out that computer programs are not devices, since, according to the Polish Languages Dictionary, a device is a mechanism or a set of mechanisms performing specific actions,¹¹ meaning that devices must be material and, apart from that, computer programs constitute intangible rights. In the literature, it is proposed that computer programs may be, at most, treated as components of devices.¹² This is a significant issue, since the removal or circumvention of the effective technical devices applied (in this case, a voice cloning computer program) is usually performed by special computer programs and the appropriate interpretation will decide whether art. 118 para. 1 applies in this regard.

5. Voice Cloning Software and its Risks

[Rz 9] It seems that the main anxiety in this area is connected with the use of audio recordings as an evidence in court. Obviously, audio recording could be used as a valid evidence in the course of litigation under the Polish jurisdiction.¹³ Some restrictions apply to recordings acquired illegally but a general rule states that such recordings are also admitted in court if they support reaching a fair ruling.¹⁴ The software enabling the creation of statements which sound, for example, like the defendant, can cause a considerable risk for the fairness of the trial. Accordingly, one of the ideas to provide protection against fake statements generated by means of voice cloning technology is adding audio watermarks to every output of such software. Digital watermarking is the process of imperceptibly embedding watermarks into digital media as a permanent sign to assure its authenticity.¹⁵

6. Data Protection in the Light of Voice Cloning

[Rz 10] Voice cloning technology requires the obtained data to be saved, therefore it is necessary to also look at this new technology from a personal data protection point of view. In the European Union, this issue is regulated by a number of directives, i.e. the Data Protection Directive (95/46/EC), the Telecommunications Act of 16 July 2004 (unified text of 2016, item 1489 as amended) or the Electronic Communications Data Protection Directive (2002/58/EC). In spite of sealing personal data protection (not only internationally but also on a national level), multiple problems occur in practice, e.g., when the data controller entrusts data to countries with insufficient data protection standards. Sufficient data protection standards shall be assessed in the light of all circumstances surrounding a data transfer operation, in particular, the nature of data, its purpose and the duration of the proposed processing operation. According to the Polish Data Protection Act of 29 August 1998 (unified text Journal of Laws of 2016, item 922), the above stated doubts arise whenever personal data is transferred to a country not belonging to the European Economic Area. However, the issue of voice filing as personal data requires a more extensive description and exceeds the scope of this paper.

¹¹ Polish Language Dictionary, <http://sjp.pwn.pl/sjp/urzadzenie;2533403.html>.

¹² JANUSZ RAGLEWSKI, Commentary on Article 118(1) of the Copyright Law, in: Damian Flisak (ed.), Copyright Law. Commentary, LEX no. 9083, 2015.

¹³ Article 308 §1 of Polish Code of Civil Procedure of 17 November 1964, Journal of Laws 2016.1822 as amended.

¹⁴ Resolution of the Supreme Court of 22 April 2016, ref. no. II CSK 478/15.

¹⁵ YIQING LIN / WALEED H. ABDULLA, Audio Watermark: A Comprehensive Foundation Using MATLAB, Springer, 2014, ISBN: 9783319079745.

7. Voice Biometrics vs Voice Cloning

[Rz 11] Voice cloning technology differs from voice biometrics technology, since the latter is a technology used to identify people by their voices. Biometrics refers to metrics related to human characteristics. Nowadays, this technology is being used increasingly, especially in matters of security (e.g., at the airport, where one can choose the facial recognition system instead of the traditional way of checking in).

[Rz 12] The human voice is as unique as fingerprints. Moreover, everyone articulates sentences in an original way: one puts emphasis differently, the rate of speech and the intonation are varying. The system records and picks out all the differences while taking into consideration details such as the size and the shape of throat, mouth cavity, nasal cavity, length and tension of vocal cords. Every recorded voice print is stored as a mathematical model. To avoid mistakes during voice recording, the commands are dictated by a speech synthesizer. The verification process consists of comparing samples of recordings to previous recordings. The said systems are currently equipped with technologies removing ambient noise, and can therefore recognize voices in most cases.

[Rz 13] Companies using voiceprint checks to verify their customers are at risk of voice cloning technologies, too. It is said that biometric systems would not be tricked by this, as the inspected items differ from what humans look for when identifying people.¹⁶ The authors of the VoicePIN, a new startup from Poland, claim that their product based on voice authentication is resistant to spoofing and can detect whether the sample is original or re-played.¹⁷ If such assumption is correct, it seems reasonable that the biometric system could likewise be protected from voice cloning software. The final answer will be known only after specific tests and experiments.

8. Conclusions

[Rz 14] The above remarks lead to the conclusion that voice cloning technology may cause new types of legal liability, both civil and penal, of the entities applying this technology. Consequently, voice cloning, like any other new technology, will involve the need to amend and adjust the existing legal provision. Nevertheless, these should not restrict the application of this technology in areas like providing services, e.g., legal advice and legal translation. The said changes are particularly required in the area of administrative law when defining the authority and supervisory competence of entities protecting personal data. Moreover, an important postulate would be to precisely define human voice as a specific personal interest. Furthermore, an unauthorised modification of such voice by means of computer devices should be classified as a specific type of infringement of human voice as a personal interest. Nevertheless, despite the existence of potential risks, when properly safeguarded by legal provisions, voice cloning software can indeed influence the effectiveness and cost-efficiency of legal services positively.

MAŁGORZATA KIEŁTYKA, Attorney, Founding Partner of KG LEGAL Law Firm, specialises in cross border cases, corporate law, life science and new technologies.

¹⁶ British Broadcasting Corporation, Adobe VoCo «Photoshop-for-voice» causes concern, 7 November 2016, www.bbc.com/news/technology-37899902.

¹⁷ Information given by CEO on VoicePIN in the interview for Business Insider, 28 March 2017, www.businessinsider.com.pl/technologie/nowe-technologie/voicepin-zabezpieczenia-biometryczne-thing-big-upc/120w4f3.

K. JAKUB GŁADKOWSKI, Attorney, Founding Partner of KG LEGAL Law Firm, specialises in cross border cases with its focus on life science, biotech, medtech, IT, new technologies and investment processes in Poland.

ANNA BEDNARSKA, Associate in KG LEGAL Law Firm, Graduate of law at the Faculty and Administration University and German Law School, particularly interested in matters related to life science, pharmaceutical law and aspects of cross border contracts.

MARIUSZ BUJACZ, Associate in KG LEGAL Law Firm, student of law at the Faculty of Law and Administration at the Jagiellonian University, mainly interested in commercial law and civil law procedure, focused on economic aspects of Polish jurisdiction.

KAMIL TRZASKOŚ, Associate in KG LEGAL Law Firm, student of intellectual property law and new media at the Faculty of Law and Administration at the Jagiellonian University, deals with cases involving new technologies and IT sector, industrial property and patents.