UNIVERSITY OF TURKU
Faculty of Law

The Unauthorised Commercial Use of Person's Voice in AI

OT00BG64 Critical International Law & Technology
Master's Thesis

Author:
Waltteri Messo

19.2.2026

UNIVERSITY OF TURKU
Faculty of Law

Abstract page

Master's thesis

Subject: OT00BG64 Critical International Law & Technology
Author: Waltteri Messo
Title: The Unauthorised Commercial Use of Person's Voice in AI
Supervisor: Outi Korhonen
Number of pages: 63 pages
Date: 19.2.2026

This thesis examines the unauthorised commercial use of the human voice in the context of generative artificial intelligence (AI) and addresses the research question: To what extent can copyright law and human rights frameworks effectively protect individuals from the unauthorised commercial use of their voice in AI-generated content? This thesis aims to answer this question by analysing the structural limitations of international copyright law and the potential of human rights instruments when a person's voice is cloned using emerging technologies and used commercially without authorisation. The study highlights the vulnerability of both public figures and ordinary citizens as technology outpaces current regulations.

Thesis aims to provide considerations for legislators regarding the necessity of stricter binding rules and the potential adoption of a *sui generis* right of publicity. The research uses legal dogmatics and critical legal studies as methods, approaching the topic from a human rights perspective. The primary sources of international law include treaties such as the Berne Convention, the TRIPS Agreement, and the European Convention on Human Rights. The main source of the case law is the European Court of Human Rights, particularly cases regarding the protection of privacy, private life, freedom of expression, and property, which form the basis of the protection framework presented in the thesis.

The findings indicate that while international copyright law is ineffective in protecting the human voice itself due to the fixation requirement, the human rights framework regarding the right to private life can be interpreted to protect the voice as personal data and an aspect of personality. However, current frameworks are insufficient, and the international community needs to consider establishing regulations, such as a specific right of publicity, to ensure adequate protection of personal identity and traits in the AI era.

Key words: artificial intelligence, voice cloning, copyright law, human rights, ECHR, personality rights, right of publicity, technology, fixation

UNIVERSITY OF TURKU
Faculty of Law

Abstract page

Tutkielma

Oppiaine: OT00BG64 Critical International Law & Technology
Tekijä: Waltteri Messo
Otsikko: The Unauthorised Commercial Use of Person's Voice in AI
Ohjaaja: Outi Korhonen
Sivumäärä: 63 sivua
Päivämäärä: 19.2.2026

Tässä tutkielmassa tarkastellaan ihmisäänen luvatonta kaupallista käyttöä generatiivisen tekoälyn (AI) kontekstissa ja vastataan tutkimuskysymykseen: Missä määrin tekijänoikeuslainsäädäntö ja ihmisoikeussäännökset voivat tehokkaasti suojella yksilöitä heidän äänensä luvattomalta kaupalliselta käytöltä tekoälyllä luodussa sisällössä? Tutkielman tavoitteena on vastata tähän kysymykseen analysoimalla kansainvälisen tekijänoikeuslainsäädännön rakenteellisia rajoituksia ja ihmisoikeussopimusten tarjoamia mahdollisuuksia tilanteissa, joissa henkilön ääni kloonataan uutta teknologiaa hyödyntäen ja sitä käytetään kaupallisesti ilman lupaa. Tutkimus tuo esiin sekä julkisuuden henkilöiden että tavallisten ihmisten haavoittuvuuden, kun teknologinen kehitys etenee nykyistä sääntelyä nopeammin.

Tutkielman tarkoituksena on tarjota lainsäätäjille näkökohtia tiukempien sitovien sääntöjen tarpeellisuudesta sekä mahdollisuudesta hyödyntää erillistä *sui generis* -tyyppistä *right of publicity* -sääntelyä. Tutkimusmenetelminä käytetään oikeusdogmatiikkaa ja kriittistä oikeusteoriaa, ja aihetta lähestytään ihmisoikeusnäkökulmasta. Kansainvälisen oikeuden ensisijaiset lähteet koostuvat sopimuksista, kuten Bernin yleisssopimus, TRIPS-sopimus ja Euroopan ihmisoikeussopimus. Keskeisimpänä oikeuskäytännön lähteenä käytetään Euroopan ihmisoikeustuomioistuimen oikeuskäytäntöä, ja erityisesti tuomioita koskien yksityisyyden, yksityiselämän, sananvapauden ja omaisuuden suojaa, mitkä muodostavat tutkielmassa esitettävän suojakehyksen perustan.

Tulokset osoittavat, että vaikka kansainvälinen tekijänoikeuslainsäädäntö on tehoton suojaamaan ihmisääntä itsessään aineellisen muodon vaatimuksen vuoksi, ihmisoikeussäännöksiä ja oikeutta yksityiselämään voidaan tulkita siten, että ne suojaavat ääntä henkilötietona ja osana persoonallisuutta. Nykyiset sääntelykehykset ovat kuitenkin riittämättömiä, ja kansainvälisen yhteisön on harkittava esimerkiksi erityisen *right of publicity* -sääntelyn luomista varmistaakseen henkilökohtaisen identiteetin ja piirteiden riittävän suojan tekoälyn aikakaudella.

Avainsanat: tekoäly, äänen kloonaus, tekijänoikeus, ihmisoikeudet, EIS, persoonallisuusoikeudet, right of publicity, teknologia, aineellisen muodon vaatimus

Table of contents

References

Bibliography

Baris, A. (2024). AI Covers: Legal Notes on Audio Mining and Voice Cloning. Journal of Intellectual Property Law & Practice, 19(7), 571.

Beckett, J. (2000). Critical international legal theory. Oxford Bibliographies Online: International Law. Retrieved from Academia.edu.

Borsos, Z., Marinier, R., Vincent, D., Kharitonov, E., Pietquin, O., Sharifi, M., Roblek, D., Teboul, O., Grangier, D., Tagliasacchi, M., & Zeghidour, N. (2023). AudioLM: A Language Modeling Approach to Audio Generation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31, 2521–2533. https://doi.org/10.1109/TASLP.2023.3288409.

Burley, S. (2023). Introduction to the International Intellectual Property Legal Framework (1st ed.). Geneva: World Intellectual Property Organization.

Cass, R. A., & Hylton, K. N. (2013). Laws of Creation: Property Rights in the World of Ideas. Cambridge, MA: Harvard University Press.

Chester-Taxin, S. (1992). Will the Real Bette Midler Please Stand Up? The Future of Celebrity Sound-Alike Recordings. University of Miami Entertainment & Sports Law Review, 9(1), 165–181.

Chisolm, T. E. (2018). In Lieu of Moral Rights for IP-Wronged Music Vocalists: Personhood Theory, Moral Rights, and the WPPT Revisited. St. John's Law Review, 92(3), 395–450.

Chopra, P., Sony, R. A., & Chopra, S. (2025). Generative AI, Copyright and Personality Rights: A Comparative Legal Perspective. Legal Issues in the Digital Age, 6(3), 23-51. https://doi.org/10.17323/2713-2749.2025.3.23.51.

Drexl, J., Hilty, R., Beneke, F., Desaunettes-Barbero, L., Finck, M., Globocnik, J., Gonzalez Otero, B., Hoffmann, J., Hollander, L., Kim, D., Richter, H., Scheuerer, S., Slowinski, P. R., & Thonemann, J. (2019). Technical aspects of artificial intelligence: An understanding from an intellectual property law perspective. Max Planck Institute for Innovation & Competition Research Paper No. 19-13. Available at SSRN: https://ssrn.com/abstract=3465577.

Esquirol, J. L. (2021). Making the critical moves: A top ten in progressive legal scholarship. University of Colorado Law Review, 92, 1079.

Fenwick, M., Jurcys, P., Liaudanskas, A., Earp, B., Porsdam Mann, S., & Kozuka, S. (2024). Voice Cloning in an Age of Generative AI: Mapping the Limits of the Law & Principles for a New Social Contract with Technology. Kyushu University Legal Research Bulletin. Available at SSRN: https://ssrn.com/abstract=4850866.

Gervais, D. J. (2003). The TRIPS Agreement: Drafting history and analysis (2nd ed.). London: Sweet & Maxwell.

Ginsburg, J. C. (2020). People Not Machines: Authorship and What It Means in International Copyright Law. In G. W. Austin, A. F. Christie, A. T. Kenyon, & M. Richardson (Eds.), Across Intellectual Property: Essays in Honour of Sam Ricketson (pp. 57–78). Cambridge: Cambridge University Press.

Goldstein, P., & Hugenholtz, B. (2010). International Copyright: Principles, Law, and Practice (2nd ed.). Oxford: Oxford University Press.

Griffiths, J. (2009). The 'Three-Step Test' in European Copyright Law - Problems and Solutions. Queen Mary School of Law Legal Studies Research Paper No. 31/2009. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1476968.

Hallo de Wolf, A., & Moerland, R. (2023). The UDHR as a living instrument at 75 and beyond. Netherlands Quarterly of Human Rights, 41(4), 182-189. https://doi.org/10.1177/09240519231214481.

Juang, B. H., & Rabiner, L. R. (2005). Automatic Speech Recognition – A Brief History of the Technology Development. Georgia Institute of Technology. Retrieved from https://web.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354_LALI-ASRHistory-final-10-8.pdf.

Kallioniemi, I. (2022). Tekoälyoikeus: varallisuusoikeuden ja riskienhallinnan kysymyksiä. Helsinki: Alma Talent.

Klabbers, J. (2023). International law (4th ed.). Cambridge: Cambridge University Press.

Klabbers, J. (2023). Soft authority in global governance. In M. Eliantonio, E. Korkea-aho, & U. Mörth (Eds.), Research handbook on soft law (pp. 162–176). Cheltenham: Edward Elgar. https://doi.org/10.4337/9781839101939.00020

Klatt, D. H. (1987). Review of Text-to-Speech Conversion for English. The Journal of the Acoustical Society of America, 82(3), 737–793.

Kong, J., Kim, J., & Bae, J. (2020). HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis. In Proceedings of the 33rd Advances in Neural Information Processing Systems (NeurIPS 2020), 33, 17022–17033.

Korhonen, O., & Leppäkorpi, M. (2024). Elusive Interdisciplinarity in International Law in the Nordics. Nordic Journal of International Law, 93, 114–131.

Kreiman, J., & Sidtis, D. (2011). Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception. Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444395068.

Kumar, K., Kumar, R., de Haas, T., Wang, W., Nouthel, H., & Courville, A. (2019). MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis. Advances in Neural Information Processing Systems, 32. arXiv:1910.06711.

Lapter, A. J. (1998). How the Other Half Lives (Revisited): Twenty Years Since Midler v. Ford – A Global Perspective on the Right of Publicity. ExpressO (bePress Legal Series). Retrieved from https://law.bepress.com/cgi/viewcontent.cgi?article=6630&context=expresso.

Lessig, L. (1999). Code and other laws of cyberspace. New York: Basic Books.

Letsas, G. (2012). The ECHR as a Living Instrument: Its Meaning and its Legitimacy. Current Legal Problems, 65, 106–141. Available at SSRN: https://ssrn.com/abstract=2021836.

Liu, J. (2023). Artificial Intelligence and International Law: The Impact of Emerging Technologies on the Global Legal System.

Makovec Petrik, T. (2012). Moral Rights of Composers: The Protection of Attribution and Integrity Available to Musicians in the European Union and the United States. Fordham Intellectual Property, Media & Entertainment Law Journal, 22(2), 359–412.

McCarthy, J. (2007). What is artificial intelligence? Stanford University. Retrieved from https://www-formal.stanford.edu/jmc/whatisai.pdf.

Milošević, D. (2022). Detention observed through the prism of human right. Glasnik Advokatske komore Vojvodine, 94(1), 177-223. https://doi.org/10.5937/gakv94-32818.

Qin, Z., Yu, W., Tan, X., & Qin, T. (2024). OpenVoice: Versatile Instant Voice Cloning. arXiv preprint arXiv:2312.01479v4. https://arxiv.org/abs/2312.01479

Ren, Y., Hu, C., Tan, X., Qin, T., Zhao, S., Zhao, Z., & Liu, T. (2020). FastSpeech 2: Fast and High-Quality End-to-End Text to Speech. arXiv preprint arXiv:2006.04558.

Ricketson, S., & Ginsburg, J. (2022). International Copyright and Neighbouring Rights: The Berne Convention and Beyond (3rd ed.). Oxford: Oxford University Press.

Russell, S. J., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Hoboken: Pearson.

Salo, I. (2024). Luova tekoäly työn supervoimana. Helsinki: Helsingin seudun kauppakamari.

Schabas, W. A. (2015). The European Convention on Human Rights: A Commentary. Oxford: Oxford University Press.

Schwarz, D. (2005). Current Research in Concatenative Sound Synthesis. In Proceedings of the International Computer Music Conference (ICMC) (pp. 1–1). Barcelona, Spain.

Schwarz, D. (2006). Concatenative Sound Synthesis: The Early Years. Journal of New Music Research, 35(1), 3–22.

Sturm, B. L. T., Iglesias, M., Ben-Tal, O., Miron, M., & Gómez, E. (2019). Artificial intelligence and music: Open questions of copyright law and engineering praxis. Arts, 8(3). https://doi.org/10.3390/arts8030115.

Tokuda, K., Zen, H., & Black, A. W. (2002). An HMM-Based Speech Synthesis System Applied to English. In IEEE Speech Synthesis Workshop. Santa Monica.

van der Sloot, B. (2015). Privacy as personality right: Why the ECtHR's focus on ulterior interests might prove indispensable in the age of "big data." Utrecht Journal of International and European Law, 31(80), 25. https://doi.org/10.5334/ujiel.cp.

van Rijn, A. (2006). Right to the Peaceful Enjoyment of One's Possessions (Article 1 of Protocol No. 1). In P. van Dijk, F. van Hoof, A. van Rijn, & L. Zwaak (Eds.), Theory and Practice of the European Convention on Human Rights (4th ed., pp. 863–894). Antwerp: Intersentia.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. In I. Guyon et al. (Eds.),

Advances in Neural Information Processing Systems 30 (NIPS 2017) (pp. 5998–6008). arXiv:1706.03762.

Watt, D., Harrison, P. S., & Cabot-King, L. (2019). Who owns your voice? Linguistic and legal perspectives on the relationship between vocal distinctiveness and the rights of the individual speaker. International Journal of Speech, Language & the Law, 26(2), 137–180. https://doi.org/10.1558/ijsll.40571.

Wells-Edwards, B. (2024). What's in a voice? The legal implications of voice cloning. Arizona Law Review, 64, 1213–1239.

WIPO. (2003). Guide to the copyright and related rights treaties administered by WIPO and glossary of copyright and related rights terms (WIPO Publication No. 891). Geneva: World Intellectual Property Organization.

Zen, H., Dang, V., Clark, R., Zhang, Y., Weiss, R. J., Jia, Y., Chen, Z., & Wu, Y. (2019). LibriTTS: A Corpus Derived from LibriSpeech for Text-to-Speech. arXiv preprint arXiv:1904.02882.

Zhang, H., & Lin, Y. (2022). Improve Few-Shot Voice Cloning Using Multi-Modal Learning. arXiv preprint arXiv:2203.09708.

Conventions

Beijing Treaty on Audiovisual Performances. (2012).

Copyright, Designs and Patents Act 1988 (United Kingdom). (1988). Available at: https://www.legislation.gov.uk/ukpga/1988/48/section/3.

European Convention on Human Rights (ECHR). (1950).

International Covenant on Civil and Political Rights. (1966).

UNESCO Recommendation on the Ethics of Artificial Intelligence (SHS/BIO/PI/2021/1). (2021).

Universal Declaration of Human Rights (UDHR). (1948).

WIPO Copyright Treaty (WCT). (1996).

WIPO Copyright Treaty (WCT) – Explanatory Notes (WIPO Publication No. 226). (1996).

World Intellectual Property Organization (WIPO). (1961). TRT/ROME/001: International Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organizations.

World Intellectual Property Organization (WIPO). (1978). Guide to the Berne Convention for the Protection of Literary and Artistic Works (Paris Act, 1971) (WIPO Publication No. 615(E)).

World Intellectual Property Organization (WIPO). (1979). TRT/BERNE/001: Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979).

World Intellectual Property Organization (WIPO). (1996). TRT/WPPT/001: WIPO Performances and Phonograms Treaty (WPPT).

World Trade Organization (WTO). (1994). Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS).

Online references

Al-Sibai, N. (2023, November 18). David Attenborough disgusted by AI clone of his voice: "I am profoundly disturbed." Futurism. Retrieved from https://futurism.com/the-byte/david-attenborough-ai-voice-clone.

Cambridge Dictionary. Voice. Retrieved from https://dictionary.cambridge.org/dictionary/english/voice.

Cooban, A. (2024, September 18). This bank says ‘millions’ of people could be targeted by AI voice-cloning scams. CNN. Retrieved from https://edition.cnn.com/2024/09/18/tech/ai-voice-cloning-scam-warning/index.html.

Coscarelli, J. (2023, April 19). An A.I. hit of fake ‘Drake’ and ‘The Weeknd’ rattles the music world. The New York Times. Retrieved from https://www.nytimes.com/2023/04/19/arts/music/ai-drake-the-weeknd-fake.html.

DeepMind. (2016). WaveNet: A Generative Model for Raw Audio. Retrieved from https://deepmind.google/discover/blog/wavenet-a-generative-model-for-raw-audio/.

DeepMind. WaveNet. Retrieved from https://deepmind.google/technologies/wavenet/.

Dictionary Cambridge. Critique. Retrieved from https://dictionary.cambridge.org/dictionary/english/critique.

Digital Equipment Corporation. (2018). DECtalk DTC-01. Internet Archive. Retrieved from https://archive.org/details/dectalk.

EditMojo. Ethical considerations in preserving author's voice. Retrieved from https://editmojo.com/ethical-considerations-in-preserving-authors-voice/.

ElevenLabs. About. Retrieved from https://elevenlabs.io/about.

Eromäki, V. (2023, April 27). Tekoäly tekee nyt myös ääntä, ja hyvin tekeekin – "Voi olla, että kohta kaikkea on oletettava huijaukseksi", miettii tutkija. Yle. Retrieved from https://yle.fi/a/74-20027151.

Free Speech Center. (2009, updated 2024, July 9). Misappropriation. Retrieved from https://firstamendment.mtsu.edu/article/misappropriation/.

Grundhauser, E. (2017, January 16). The Voder, the First Machine to Create Human Speech: It Spoke Like A Demon. Atlas Obscura. Retrieved from https://www.atlasobscura.com/articles/the-voder-the-first-machine-to-create-human-speech.

Houser, K. (2019, November 6). Alarming AI clones both a person's voice and their speech patterns: Hear its best Bill Gates impersonation. Futurism. Retrieved from https://futurism.com/the-byte/ai-clones-voice-speech-patterns.

IBM. What is Artificial Intelligence (AI)? Retrieved from https://www.ibm.com/think/topics/artificial-intelligence.

International Trademark Association (INTA). (2019). Right of Publicity. Retrieved from https://www.inta.org/topics/right-of-publicity/.

Jia, Y., & Ramanovich, M. T. (2021). High-Quality, Robust and Responsible Direct Speech-to-Speech Translation. Google Research. Retrieved from https://research.google/blog/high-quality-robust-and-responsible-direct-speech-to-speech-translation/.

Lammers, T. (2024, September 9). James Earl Jones signed over rights for AI to recreate Darth Vader's voice. Forbes. Retrieved from https://www.forbes.com/sites/timlammers/2024/09/09/james-earl-jones-signed-over-rights-for-ai-to-recreate-darth-vaders-voice/.

Meta AI. (2023, June). Introducing Voicebox: The First Generative AI Model for Speech to Generalize Across Tasks with State-of-the-Art Performance. Retrieved from https://ai.meta.com/blog/voicebox-generative-ai-model-speech/.

Midler v. Ford Motor Co.Casebriefs. Retrieved from https://www.casebriefs.com/blog/law/intellectual-property-law/intellectual-property-keyed-to-merges/state-intellectual-property-law-and-federal-preemption/midler-v-ford-motor-co/.

Milmo, D. (2024, May 27). Scarlett Johansson's OpenAI clash is just the start of legal wrangles over artificial intelligence. The Guardian. Retrieved from https://www.theguardian.com/technology/article/2024/may/27/scarlett-johansson-openai-legal-artificial-intelligence-chatgpt.

Modern Diplomacy. (2025, November 16). Who owns the voice of the future – intellectual property in the age of AI media. Retrieved from https://moderndiplomacy.eu/2025/11/16/who-owns-the-voice-of-the-future-intellectual-property-in-the-age-of-ai-media/.

Mozilla. Common Voice Datasets. Retrieved from https://commonvoice.mozilla.org/en.

NVIDIA. Tacotron 2. GitHub. Retrieved from https://github.com/NVIDIA/tacotron2.

OpenAI. (2024, March 29). Navigating the Challenges and Opportunities of Synthetic Voices. Retrieved from https://openai.com/index/navigating-the-challenges-and-opportunities-of-synthetic-voices/.

Post, R. C., & Rothman, J. E. (2020). The First Amendment and the Right(s) of Publicity. Yale Law Journal. Retrieved from https://yalelawjournal.org/article/the-first-amendment-and-the-rights-of-publicity.

Prolific North. (2025). AI vs the human voice – how can the creative industries protect the rights of voice artists? Retrieved from https://www.prolificnorth.co.uk/feature/ai-vs-the-human-voice-how-can-the-creative-industries-protect-the-rights-of-voice-artists/.

Rosner, H. (2021, July 15). A Haunting New Documentary About Anthony Bourdain. The New Yorker. Retrieved from https://www.newyorker.com/culture/culture-desk/roadrunner.

Sky News. (2023). Cher hits out at AI after hearing fake version of herself covering a Madonna track. Retrieved from https://news.sky.com/story/cher-hits-out-at-ai-after-hearing-fake-version-of-herself-covering-a-madonna-track-12988405.

Speechify. (2022). What Is the History of Text to Speech and Voice Synthesis? Retrieved from https://speechify.com/blog/history-of-text-to-speech-voice-synthesis/.

Udio. About Us. Retrieved from https://www.udio.com/about-us.

Ward & Smith. (2025). The voice behind the brand: Who owns what you hear? Retrieved from https://www.wardandsmith.com/articles/the-voice-behind-the-brand-who-owns-what-you-hear.

What Is the Voder? (2025). Retrieved from https://www.whatisthevoder.com.

World Intellectual Property Organization (WIPO). Artificial Intelligence: Call for Comments on Intellectual Property Policy. Retrieved from https://www.wipo.int/documents/d/frontier-technologies/docs-en-artificial-intelligence-call-for-comments-ind_modkova.pdf.

Zhang, M., & Karch, G. (2019). Generate Natural Sounding Speech from Text in Real-Time. NVIDIA Developer Blog. Retrieved from https://developer.nvidia.com/blog/generate-natural-sounding-speech-from-text-in-real-time/.

Cases

European Court of Human Rights:

Aksu v. Turkey [GC], nos. 4149/04 and 41029/04, ECHR 2012.

Anheuser-Busch Inc. v. Portugal, no. 73049/01, ECHR 2007.

Bohlen v. Germany, no. 53495/09, ECHR 2015.

Chauvy and Others v. France, no. 64915/01, ECHR 2004-VI.

Dudgeon v. United Kingdom, no. 7525/76, 24 February 1983.

Lingens v. Austria, 8 July 1986, Series A no. 103.

Marckx v. Belgium, no. 6833/74, 13 June 1979, Series A no. 31.

Nikowitz and Verlagsgruppe News GmbH v. Austria, no. 5266/03, ECHR 2007.

P.G. and J.H. v. the United Kingdom, no. 44787/98, ECHR 2001-IX.

Palomo Sánchez and Others v. Spain [GC], nos. 28955/06, 28957/06, 28959/06, and 28964/06, ECHR 2011.

Reklos and Davourlis v. Greece, no. 1234/05, 15 January 2009.

Sapan v. Turkey, no. 44102/04, 8 June 2010.

Sousa Goucha v. Portugal, no. 70434/12, ECHR 2016.

Tyrer v. United Kingdom, no. 5856/72, 25 April 1978, Series A no. 26.

Vereinigung Bildender Künstler v. Austria, no. 68354/01, ECHR 2007.

Von Hannover v. Germany (no. 1), no. 59320/00, ECHR 2004-VI.

Von Hannover v. Germany (no. 2), nos. 40660/08 and 60641/08, ECHR 2012.

United States:

Midler v. Ford Motor Co., 849 F.2d 460 (9th Cir. 1988).

Waits v. Frito-Lay, Inc., 978 F.2d 1093 (9th Cir. 1992).

List of Abbreviations

AGI	General Artificial Intelligence Systems
AI	Artificial Intelligence
ANI	Narrow Artificial Intelligence Systems
ASI	Artificial Super Intelligence
Berne Convention	Berne Convention for the Protection of Literary and Artistic Works
CNN	Convolutional Neural Network
CoE	The Council of Europe
ECHR	European Convention on Human Rights
ECtHR, the Court	European Court of Human Rights
EU	European Union
GANs	Generative Adversarial Networks
HMMs	Hidden Markov Models
INTA	International Trademark Association
TRIPS	Agreement on Trade-Related Aspects of Intellectual Property Rights
TTS	Text-to-Speech
UN	United Nations
UDHR	UN's Universal Declaration of Human Rights
UNESCO	United Nations Educational, Scientific and Cultural Organization
WCT	The WIPO Copyright Treaty
WIPO	World Intellectual Property Organization
WPPT	WIPO Performances and Phonograms Treaty
WTO	World Trade Organization

1 Introduction

1.1 Background and motivation

This master's thesis discusses the problem concerning "Unauthorised Commercial Use of the Person's Voice in AI". Over the past few years, artificial intelligence (AI) has been developing at a furious pace and is still accelerating. This research discusses the AI models that are capable of mimicking human voices. AI models are able to mimic individual speech patterns and cadence via exposure to recordings of human speech.[1] In the context of the thesis, AI is understood as a branch of computer science. It is described as computer-based systems developed to mimic human behaviour. AI is considered an umbrella term that covers machine learning, evolutionary algorithms, and other technologies, such as rule-based systems. However, the exact definition of AI is difficult and a subject of controversy among different researchers.[2]

The meaning of the word *voice* is the sounds that are made when people speak or sing[3] - an inherently human trait closely tied to personal identity and expression. The voice is not just a functional attribute but a deeply personal aspect of human creation, deserving legal recognition and protection. Voice AI models work by feeding the speaker's voice, which in this case is data, into voice AI models for cloning. The model, however, does not copy the audio file, but computes a probability calculation from it, as if analysing how someone's speech might sound in different texts. The synthetic speech of the models themselves is generated from an acoustic model of the synthesiser's digital audio signal, so there is no actual human voice in the final result. Voice cloning AI models are generative. This means that they can reproduce random variation from the data. Thus, AI voice models are not samplers: they do not copy, cut, trim and paste the sound that is fed to them.[4]

In the past, high-quality voice cloning took a lot of time and cost, but it is now easily accessible and instant. The research will include cases of voice cloning where a person's voice has been used commercially in AI without permission. This includes cases where, firstly, an unauthorised person's voice model is created by for example a developer or company for a use

[1] Futurism 2019.
[2] *Drexl et al.*, p. 3.
[3] Cambridge Dictionary, Voice.
[4] In Finnish: Yle 2023.

in an AI technology. Secondly, those cases where a user creates a voice model of another person using an AI technology and uses it for commercial purposes.

There have been numerous widely publicised cases of this kind in recent times. Recently, OpenAI was reported to have used a voice very similar to Scarlett Johansson's in an updated version of ChatGPT, which can listen to spoken prompts and respond to them verbally.[5] In April 2023, the music industry was stunned by the unauthorised use of the voices of famous artists The Weeknd and Drake in the AI-generated song "*Heart on My Sleeve*", which quickly became a hit on numerous streaming services.[6] Recently, there was an uproar over a case where the voice of nature documentarian and presenter David Attenborough was cloned to make comments on crises and politics.[7] In "*Roadrunner: A Film About Anthony Bourdain*", the late Bourdain's voice was replicated by an AI at three different scenes to read quotes he had written. The late Bourdain had not given his permission for replication, but nothing was put into Bourdain's mouth that he had not said himself, at least in writing. Is that wrong? There is probably no right answer. Undoubtedly, artificially replicating the voice of the subject of a film, without viewers necessarily even realising it, raises a whole new set of ethical questions.[8] For well-known voices, this has not always been the case, for example, James Earl Jones signed a contract before his death to use AI to create the voice of his character Darth Vader in future films.[9] However, alongside these high-profile cases involving public figures, it is essential to take into account all the ordinary people whose voices can be used without authorisation just as easily. Examples of this can be seen, for example, in AI voice-cloning scams.[10]

Why should we be concerned? Our voice lies at the core of who we are, reflecting our personality, values, and the messages we wish to convey. It is a unique expression of identity, expression of personality, values and message. Voice serves as a distinct representation of who we are and a primary means of connecting with others and influencing the world around us. As AI and other technologies continue to evolve, it's important that we own the personal parts of what we have. The rapid development of these technologies often outpaces legal protections, leaving individuals vulnerable to potential misuse and exploitation.

---

[5] The Guardian 2024.
[6] The New York Times 2023.
[7] Futurism 2023.
[8] The New Yorker 2021.
[9] Forbes 2024.
[10] CNN 2024.

Why is it even important to own anything or to own things on the internet? Ownership, whether of a physical object or an intangible feature such as a voice, provides control and autonomy to the individual. It allows access, control and decision-making over their property.[11] The Internet has created new challenges to the concept of ownership, especially in the case of digital content. In the physical world, it is easier to determine who owns which object. In the digital world, however, control is much more difficult because it is easy to copy and share content.[12]

The unauthorised use of a person's voice in AI is a complex issue with multiple aspects. Technologically speaking, recognisable features of voice, such as tone, pitch and speech pattern, can be stolen.[13] These features can be used in AI applications to create the impression that a person has said something they have not said, or AI can be used to create an output in which the original owner of the voice does not want to be involved.[14] Such manipulation of voice can lead, for example, to identity theft and damage to reputation. Indeed, the concept of theft is a challenging one in the topic area. When something is similar enough to be a clone? Determining the degree of similarity seems to be a subjective matter and there is no clear standard to follow.

From a copyright perspective, it is difficult to say whether the unauthorised use of voice constitutes copyright infringement. Copyright protects creative works, and the use of voice in AI may depend on how the AI is used and for what purpose the voice is used. In general, the problem arises when a copyrighted work is used to clone sound. In this case, even a few seconds of audio can be sufficient for copyright infringement.[15]

The unauthorised use of a person's voice in AI can be seen as an international law issue due to the global nature of the technology and the gaps in existing legal frameworks. AI voice-cloning technologies are developed and used across the borders, creating potential conflicts in applying diverse intellectual property, privacy, and publicity laws. International law could be a key to harmonising these differences and addressing transnational challenges effectively.

Current frameworks, such as the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), do not explicitly protect the human voice, leaving it vulnerable to

---

[11] *Lessig*, p. 108.
[12] Ibid., p. 128.
[13] *Wells-Edwards*, pp. 1214-1216.
[14] Ibid., p. 1230.
[15] *Sturm et al.*, p. 4.

misuse. Similarly, the principle of right of the publicity, which safeguards a person's commercial identity, lacks global recognition and enforcement mechanisms. Beyond legal gaps, voice cloning raises ethical and human rights concerns, including violations of privacy, autonomy, and dignity. The rapid development of AI technologies is outpacing the regulations, which leave individuals exposed to potential exploitation. Global initiatives like UNESCO's AI Ethics Recommendations and the EU AI Act (2024/1689) highlight the need for international cooperation to protect individual personality rights. The unauthorised use of voice in AI is a new issue and raises questions that may not be fully covered by current legislation. As AI and technologies develop, new legislation and regulatory frameworks will likely be needed to protect human rights in the AI era.[16]

Research question:

(1) To what extent can copyright law and human rights frameworks effectively protect individuals from the unauthorised commercial use of their voice in AI-generated content?

[16] Ibid, p. 12.

1.2 Relevant doctrines

The main subjects in the field of unauthorised commercial use of a person's voice in AI can be categorised into various groups. Firstly, individuals have an important role in the topic. Individuals are mostly the owners of the voice that is used or cloned without permission. Individuals are a large group that includes celebrities (for example musicians, actors and such) as well as ordinary people. Individuals can also be at the other end of the problem acting as the technology user of an AI model creating the unauthorised AI output.

Secondly, corporations are another important group in the topic. Corporations could also be the owners of the voice if the corporation has bought the rights of it. Corporations could also be unauthorised users of voice, for example, using AI-generated voices in marketing, films, or audio content without permission or creating the AI output. Corporations are examined mainly in this context in the thesis. Corporations can also be the AI developers that own and create the AI technologies behind voice cloning and unauthorised outputs. Corporations are usually also the ones that provide the platform for deploying AI-generated voices, such as social media and streaming services. The responsibility for the unauthorised content on the platforms is an issue outside the scope of this research.

Thirdly, states are playing a big role in the regulatory framework of the topic. States are responsible for enacting and enforcing laws concerning copyright as a part of intellectual property and also the human rights related to AI and unauthorised use of voice. Intergovernmental organisations are one of the most significant actors in the issue. Bodies are involved in setting international norms and treaties for AI governance and intellectual property. When approaching the issue from a copyright perspective, the important actors are the World Intellectual Property Organization (WIPO) and the World Trade Organization (WTO). When approaching the issue from a chosen human rights perspective, the important actors involved are the United Nations (UN) and the Council of Europe (CoE). The European Union is an important body in regulating AI with the new regulation, but it is set outside the scope of this research because the overall approach is international.

Fourthly, the national and international courts are important actors in the topic. Because the approach is international, the international courts and their case law are the most significant for the research. For example the European Court of Human Rights (ECtHR) has important case law related to personality rights that could be lead to consider unauthorised use of voice in technology. In addition, there is transnational associations, for example the United Nations

Educational, Scientific and Cultural Organization (UNESCO), that are offering soft law norms, for example UNESCO's Recommendation on the Ethics of Artificial Intelligence.[17]

From these actors, however, we can roughly identify the three main groups that are most relevant to the research: Voice owners, unauthorised voice users and the actors that bring regulation and control to the playing field. In general, voice is owned and used without permission by private individuals and corporations. Regulation in international law is the responsibility of various bodies. Because there is not a direct regulatory framework, the main actors in the field are the international intellectual property regulators and the international human rights regulators. Regulation takes the form of international conventions between states, which govern the action or are implemented in national legislations. The enforcement of infringements is generally controlled at the lowest level by national courts and, if they do not provide legal protection or if there is no appeal at national level, legal protection is provided by international courts.

When it comes to the binding nature of international conventions and the effective legal protection in a matter it has to be noted that there is a lot of potential for a soft law regulatory like the UNESCO's recommendation in the topic area. In fast moving issues such as the topic, where problems easily accumulate and require quick action, the benefits of soft law regulation step in. By not setting anything in stone, it allows for context and influences decision making. Soft law allows regulators to balance the different demands coming from different bodies and institutions that declare different objectives and propose various and possibly contradictory behaviours.[18] Compared to the lengthy processes of hard law, soft law can be regulated quickly and efficiently, for example in the form of recommendations and guidelines.[19] It is also important to note that soft law regulation could easily be used as a basis for future stricter regulation.

From the soft law to the hard law distinction, it is easy to move to the central doctrines of international law: *monism* and *dualism*. These doctrines are crucial for the legal validity of different regulations and conventions in different states. Monists accept that the domestic and

[17] UNESCO, Recommendation on the Ethics of AI.
[18] *Klabbers*, p. 169.
[19] Ibid, p. 167.

international legal systems form a whole. Domestic and international legal rules, adopted by a state, for example through a convention, determine whether an action is legal or illegal.[20]

Dualists emphasise the distinction between domestic and international law and call for the latter to be transferred to the former. Without this translation, international law does not exist as law. The monist system is therefore seen as favourable, for example, to the protection of human rights. A dualist system, on the other hand, is able to observe which doctrines are effective and useful and to selectively implement them into the system.[21]

According to *Jan Klabbers*, in the "new dualist" system, it is noticeable that international law standards are sometimes avoided. Often the reason for avoidance is denying international law or claiming that an issue that has been raised should not be dealt with because it involves international issues.[22] This would lead to a situation in the topic field where personality rights could be violated in a state that does not recognise their protection in any form. For example, the person's voice could be used without permission in the territory of a state that does not give human rights protection to the person's voice or has not implemented a new regulation in its domestic legal order. Klabbers has therefore pointed out that international law is not always about states and rules, but also about who gets to decide which regulation and how the rules are interpreted, applied and in which circumstances.[23] In my opinion this can be connected to the evolving issues involving technology and regulations. There is often no direct regulation in these cases, so the existing regulation has to be applied in different adopted ways. When there will be a new regulation in these areas, it will be a leading and significant one. Thus, there are major advantages and reputations to be gained from the drafting and implementation of such regulation.

In identifying the doctrines that are central to the research, we can make a distinction between the copyright related doctrines and human rights related ones. The idea behind the distinction is whether voice could be protected by copyright or human rights, or even both. Copyright protection is based on whether the voice is an independent work that receives copyright protection or whether the case is that the voice itself does not receive protection, but the work of which the voice would be a part does. The human rights approach would be based on the assumption that voice is part of a person's identity and personality, and whether voice could

---

[20] *Klabbers*, pp. 318-321.
[21] Ibid.
[22] Ibid., pp. 332-333.
[23] Ibid., p. 334.

be protected through this protection. The challenge with the current voice regulation is that there is no direct regulatory framework and the existing regulation is backdated regulation that was enacted well before the various technologies have developed to their current state. This results in a situation where the existing regulation has to deal with scenarios where regulation does not directly reflect all modern conditions.

1.2.1 Copyright doctrines

The fixation doctrine is a general principle of copyright law that creative works must be reduced in some material form. This doctrine is reflected in the Berne Convention for the Protection of Literary and Artistic Works (Berne Convention) Article 2(2), which allows member states to require fixation for copyrightable works.[24] However, the fixation requirement raises issues in the context of voice and AI-generated outputs. A person's voice itself is not in a material form if it is not a recording of it. Where recordings of songs or spoken performances are fixed and protected, the raw, unrecorded qualities of the voice remain outside the doctrine's scope.

Moral rights, in the context of protection under the Berne Convention Article 6bis, acknowledge the personal and reputational interests of authors in their works. These rights include the right to claim authorship, the right to be recognised as the creator and the right to object to certain modifications and other derogatory actions.[25] The issue there is that when the voices themselves, are not considered to be copyrightable works in more traditional frameworks, individuals may therefore lack standing to invoke moral rights over their voices. In addition, AI-generated voice outputs often become disconnected from the original work, with potential distortions or misrepresentations that moral rights protections were never designed to address. Nevertheless, the principles of moral rights could provide protection in a case of an infringement due to unauthorised use of the voice.

Performers' rights, established under the Rome Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organizations[26] and reinforced by the WIPO Performances and Phonograms Treaty[27], recognise the contributions of performers to creative works. These rights grant performers control over the fixation and reproduction of their

[24] Berne Convention.
[25] Ibid.
[26] Rome Convention.
[27] WPPT.

performances, as well as protection against unauthorised distribution or alteration. However, these rights typically apply to live or recorded performances, not to the voice itself as an element.

1.2.2 Human rights doctrines

The right to artistic expression is protected in Article 27 of the UN's Universal Declaration of Human Rights (UDHR).[28] This Article in the Universal Declaration of Human Rights is also strongly linked to the protection of copyright. The Article guarantees everyone the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author. The right could therefore be described as very similar to copyright, and under this right, works that include, for example, voice could be protected from being used in connection with AI.

The right to privacy is protected under Article 12 of the Universal Declaration of Human Rights and under Article 8 of the European Convention on Human Rights (ECHR) as part of the protection of private and family life.[29] As we can see in the case study later on, the ECtHR broadly interprets the ECHR and privacy protection has been extended to protect for example personal data, identity, and attributes. The question is whether the person's voice is protected under these Articles as part of their identity or personal data. The person's voice could be seen as a part of the personality or identity because it is unique to everyone, it can be used to identify you, it can be used to express things to those around you, and so on. The unauthorised use of an individual's voice can violate their control over personal data. With voice cloning, a person may be misrepresented, losing control of their own identity.

The right to freedom of expression is protected in Article 19 of the Universal Declaration of Human Rights and the ECHR Article 10.[30] The freedom of expression includes both the right to transmit information and the protection of creative expression. The right is raising dilemmas in the balancing of human rights. As we can see in the case study later on, the right to freedom of expression can clash with the right to privacy. Leading the issue into the topic of the research, the questions arising are that could the cloning of the voice have protection under these Articles as a creative or technological expression? If the unauthorised use of voice

---

[28] UDHR.
[29] The right to privacy can be found in the most of the leading human right's conventions, e. g. Article 17 of the UN's International Covenant on Civil and Political Rights (ICCPR).
[30] The right to freedom of expression is also protected in the UN's ICCPR Article 19.

is breaching the other rights of the person, for example, the right to privacy, could that kind of use of the voice have protection for parody purposes? These questions cannot be answered directly and need examining.

The right to property is protected in the UN's Protocol to the Convention for the Protection of Human Rights and Fundamental Freedoms (Protocol 1), Article 1.[31] In the Article, every natural or legal person is entitled to the peaceful enjoyment of his possessions. In the research, the Article is raising questions about whether the voice could be that kind of *possession*. If we think for example a singer, for she or he the voice is an extremely important asset that has a value. If the voice itself could not be the possession, could the copyright of a song that could be used as a data in AI to create unauthorised output that includes person's voice be the possession? For example, in ECtHR case *Anheuser-Busch Inc. v. Portugal* an application for registration of a trademark, which is an intellectual property right, could be seen as a possession subject to ECHR property protection.[32]

1.2.3 ECtHR case study

When approaching the issue of "Unauthorised commercial use of the person's voice in AI" from the international law and human rights perspective, the important legal sources are the European Convention on Human Rights (ECHR) and the case law of the European Court of Human Rights (ECtHR). The person's voice could be protected from misuse under the ECHR's Article 8 of "Right to respect for private and family life". On the other hand, the AI output where the voice is used could have some protection under the ECHR's Article 10 of "Freedom of expression". In this research, we are going to examine cases related to the examples.

Article 8 - "Right to respect for private and family life"

The ECtHR has broadly interpreted Article 8 of the European Convention on Human Rights, which concerns the protection of private and family life. According to the Court, Article 8 covers not only the traditional aspects of privacy protection but also the development of a person's identity and personality.[33] Following this interpretation, Article 8 could then also apply to situations of unauthorised use of the voice if it violates a person's identity or prevents

---

[31] Ibid.
[32] *Case of Anheuser-Busch Inc. v. Portugal*, 78§.
[33] *van der Sloot*, pp. 25-26, p. 34.

the development of his or her personality. Whether the voice is part of a person's identity or personality is another question, which I will try to answer in the research.

The ECtHR developed the living instrument doctrine in *Tyler v. the United Kingdom* from 1978. Under this doctrine, the Court held that "the Convention is a living instrument which must be interpreted in the light of present-day conditions." Applying the doctrine to the modern day, the interpretation of the ECHR should take account of developments in society and technology.[34] Artificial intelligence and the use of voice in AI models are relatively recent developments, so the case law of the ECtHR does not provide direct answers to this problem.

Since there are no cases involving the recording of voice, the closest approach is with cases involving images, which discuss the same kind of issues. In the ECtHR case *von Hannover v. Germany (No. 2)* (nos. 40660/08 and 60641/08, 7 February 2012), Princess Caroline von Hannover and Prince Ernst August von Hannover challenged the German courts' refusal to prohibit the further publication of photographs taken of them during a holiday without their knowledge. These photos had been published in German magazines. The applicants argued that the German courts failed to sufficiently consider the European Court's earlier judgment in *von Hannover v. Germany (2004)*.[35]

The Court found no violation of Article 8 (right to respect for private life) of the European Convention on Human Rights. It noted that the German courts had appropriately balanced the publishing companies' right to freedom of expression against the applicants' right to private life. Key considerations included whether the photographs and accompanying Articles contributed to a debate of general interest and the context in which the photographs were taken.[36]

The Federal Court of Justice had revised its approach in light of the European Court's 2004 judgment, and the Federal Constitutional Court reinforced this, conducting a detailed analysis of the Court's case law. Given the national courts' margin of appreciation in balancing competing interests, the European Court concluded that Germany had fulfilled its positive obligations under Article 8 in this case.[37]

---

[34] Ibid p. 39.
[35] ECtHR, Factsheet – Right to the protection of one's image.
36 Ibid.
37 Ibid.

In the decision, the attention is focused on the following observations made by the Court. In 95§, the Court emphasises that the notion of private life broadly extends to the features of one's personal identity, including an individual's name, physical and moral integrity, and photographs. The primary purpose of the protection granted under Article 8 of the Convention is to allow individuals to freely develop their personalities and maintain relationships with others without external meddling. Consequently, even when an individual is in a public setting, their social interactions can still be safeguarded under the umbrella of private life. Therefore, releasing a photograph to the public can constitute a severe infringement on private life, regardless of whether the subject is a well-known public figure.[38]

In 96§, the Court highlights that an individual's visual representation is a fundamental trait of their personality. A person's image displays their distinct characteristics and separates them from their peers. As a result, safeguarding one's image is a vital element of personal growth and self-determination. Fundamentally, the protection grants the individual the right to control how their image is used, which inherently includes the power to deny its publication altogether.[39]

These observations made by the ECtHR are relevant to research. Indeed, the Court held that in this case the concept of respect for private and family life under Article 8 extends to aspects of personal identity, such as a person's name, photograph or physical and moral integrity. Moreover, a person's image is one of the most important aspects of their personality, as it reveals their unique characteristics and distinguishes them from their peers. I would argue that these points could quite reasonably be extended to the protection of the person's voice, which is very similar in its nature. In this case, however, the image did not receive protection under Article 8. The freedom of expression was a stronger interest in circumstances where there was a high public demand for images. If this were applied to the unauthorised commercial use of a person's voice in AI, one could well face the same issues and assess, for example, the level of publicity of the person, the way the AI output was produced and its unexpectedness.

Article 10 - "Freedom of expression"

Often the right to respect for private and family life under Article 8 conflicts with the right to freedom of expression under Article 10.

---

38 *Case of von Hannover v. Germany* (No. 2), 95§.
[39] Ibid 96§.

In the ECtHR case *Sapan v. Turkey* (8 June 2010), the applicant, who owned a publishing house, challenged the restriction ban of a book analysing the star phenomenon and its appearance in Turkey, using the example of a prominent singer. The restriction ban was ordered for nearly two years and eight months at the request of the singer, who argued that the book's title, which included his name, and the use of photographs featuring him, violated his right to protection of his image and personality rights.[40]

The European Court of Human Rights held that there had been a violation of Article 10 (freedom of expression) of the European Convention on Human Rights. It found that the restriction ban was not necessary in a democratic society because it lacked relevant and sufficient justification. The Court emphasised that the book analysed stardom and its emergence in Turkey through a scientific lens, distinguishing it from tabloid publications or gossip columns aimed at sensationalising celebrities' private lives.[41]

Additionally, the photographs in the book were previously published and posed for by the singer. The domestic courts failed to sufficiently assess the balance between the applicant's right to freedom of expression and the singer's right to privacy. Consequently, the restriction ban was deemed disproportionate and unjustified under Article 10.[42]

In this case, the commercial role of the case would be of special interest. The case held that the publication of the book served a social debate and was justified on scientific grounds. Similarly, AI-generated sounds may be used for entertainment, artistic or commercial purposes, but the question of whether such use violates individual rights should be assessed. In this case, the images published in the book had already been published earlier by the artist. For example, if an AI were to create a song using someone else's voice, these would also usually have been previously published. Similar legal issues would therefore arise when analysing the extent to which individuals have the right to control their voice and its commercial use in AI technologies, and to what extent such use could fall within the scope of freedom of expression.

40 ECtHR, Factsheet – Right to the protection of one's image, pp. 4-5.
41 Ibid.
42 Ibid.

1.3 Theory and methods

The theoretical approach of the research utilises critical legal studies and approaches the topic of the research from a human rights perspective. The research uses legal dogmatics and case studies as legal methods. Through these lenses, the research tries to answer the research question.

Critique is described in its usual context as a report of something such as a political situation or system, or a person's work or ideas, that examines it and provides a judgment, especially a negative one.[43] While discussing critique in a law context, the critique is something that can be shown towards the law in all its forms. The form of law in this context can be anything between the legal literature, legal system and judgment. The method to use when showing the critique in the legal context can be the critical theory. The critical theory can be distributed into the different critical moves. Those moves can be described as a means of expressing the critique.

In the legal context, the critique does not always need to be negative. One of the goals is to demystify the hold of traditional legal thought. In the demystification of the legal thought, the thought is analysed in parts and it shows how the argument is put together.[44] While demystifying the legal thought it aims to reveal that the law is not a neutral or objective system but that the law strives to promote the interests of those in power and it also reflects these interests, which are also reflected in the criticisms of the law.[45] When discussing scientific legal thoughts the open academic critique is essential for them to actually be scientific. Scientific legal thoughts also need to be repeatable or the research method needs to be conducted known. This can usually also be found out while exposing the thoughts to criticism.[46]

According to *Jorge L. Esquirol*, defining the critical theory and approach is difficult if even impossible. The concepts of criticism and the critiques take different meanings depending on their community of reference. While there are broad legal similarities in the world, the meanings of laws can vary considerably depending on how members of a given interpretive community make conceptual and relational associations that make sense to them.[47] Esquirol

---

[43] Dictionary Cambridge, Critique.
[44] *Esquirol*, p.1106.
[45] Ibid p. 1125.
[46] *Korhonen – Leppäkorpi*, p. 117.
[47] *Esquirol*, p. 1080.

argues that this happens because critiques are *metabolised* differently in different places depending on situational understandings and have relative effects. The same critique can lead to different outcomes or even reinforce opposing discourses depending on the context.[48] In this context, metabolising means that the critiques and the critical moves are adapted into the needed purpose of the communities. *Outi Korhonen* and *Mervi Leppäkorpi* argue the importance of community and context for scientific research. They argue that scientific objectivity should be pursued and it might be complicated to achieve because the researcher is always part of the community and context he or she is studying.[49] Esquirol thus points out that critical theory must take this community and contextual aspect into account. It cannot assume that legal concepts and critiques are universal or neutral, but it must always take into account how they are understood and how they affect in different communities.[50] A critic must therefore find out how they are interpreted in the context of the critique.

Esquirol's critical moves offer a useful critical analysis tool for analysing AI critically. The gaps, conflicts, and ambiguities move presents that the legal materials are riddled with gaps, conflicts, and ambiguities. The move's target is to critically find and reveal those problems. The problem with the gaps, conflicts and ambiguities is that those usually fall onto the shoulders of the society's most disadvantaged. In the context of this research, the vulnerable in society could be seen as people who have unexpectedly become unprotected pawns in the playground of AI. This is partly because regulation is lagging badly behind technological developments.

The flipping move as the name suggests, means turning an argument on its head. The intention is to use the same arguments to reach the opposite conclusion to the original claim.[51] This move is smoothly applied to the AI. Developers and supporters of AI often emphasise its positive effects, such as improved efficiency and new innovations. However, these same arguments can be flipped to show that AI can also have negative effects, such as job losses of the voice actors. In this topic area, it is also noticeable that advanced technology has created new problems that were not predicted, such as the unauthorised use of a person's voice.

Jason Beckett's article about the critical international legal theory (CILT) is closely related to Esquirol's critical theory and critical moves. CILT questions the traditional assumptions of

---

[48] Ibid p. 1081.
[49] *Korhonen – Leppäkorpi*, p. 116.
[50] *Esquirol*, p.1082.
[51] Ibid p. 1116.

international law and seeks to expose its shortcomings and contradictions. Beckett emphasises that CILT does not consider the shortcomings of international law to be marginal or exceptional, but rather structural and permanent. CILT brings the whole world outside the legal texts and processes into the critical review. This means that it is also essential to assess the concrete consequences and implementation of laws, as opposed to simply making the necessary laws, which may not be implemented in practice.[52]

Beckett describes CILT as taking responsibility for our actions. By this, he means that if a structure that is created, for example, creates discrimination, inequality or oppression, then progressive efforts should be made to make real and positive changes in the world.[53] This is where CILT differs from Esquirol's critical theory. Critical theory focuses on demystifying legal ideas, breaking them down, rather than providing solutions. In a sense, the role of CILT could be seen in legal research as a theory that could be used to demonstrate resolutions after critical theory has demonstrated flaws, which are then highlighted by critical moves. CILT could therefore provide possible resolutions for how to improve the position of threatened and nearly captured pawns on the AI playground.

In this research, a human rights approach could help to frame the unauthorised use of voice as a violation of rights like privacy, dignity, and autonomy. The approach would allow the critique towards existing protections and argue for a legal evolution to protect individuals' rights in the context of AI technologies.

When it comes to discussing the methodology of research, it must be pointed out that the use of the legal dogmatics is necessary up to a certain point. Legal dogmatism focuses on interpreting and systematically organising existing legal norms, doctrines, and regulations. In legal dogmatism the existing laws and doctrines related to copyright and publicity rights are examined and clarified. This method clarify the current legal landscape and identify specific legal gaps or ambiguities regarding AI voice replication.

This research use case studies to address the challenges of unauthorised AI-generated voice use. Case studies of notable incidents provide practical insights into how courts interpret existing laws. These examples help evaluate the adequacy of current legal responses and expose gaps needing attention.

---

[52] *Beckett*, p. 1-2.
[53] Ibid. p. 2.

1.4 Early brief conclusions, sources and tools used in the research

The unauthorised commercial use of a person's voice in AI raises critical legal, ethical, and technological questions in copyright law and human rights. Voice, as a unique aspect of identity, requires protection to protect individual autonomy, privacy, and creative expression while AI and other emerging technologies are evolving quickly.

Copyright law struggles to address the nature of voice when it's not in a material form. Human rights frameworks, while offering robust protections for privacy, identity, and property, face challenges in applying these principles to the complexities of AI-generated outputs and new technologies. Currently the human right framework can be lead to cover some of the new issues with the technologies but if the technologies are evolving on a current rate in the future, a lot of more problems are ahead.

A unified approach that integrates the strengths of copyright doctrines and human rights protections and could combine those to personality rights is essential in the future. Harmonising international regulations and fostering ethical AI practices can ensure the respectful and legal use of voices in AI technologies, protecting individuals in this evolving digital era.

This research is based on legal literature, mainly books and articles, internet sources, including news and blog posts, international legal case law, and certain international conventions. The author's first language is Finnish, and therefore, the AI-based DeepL Translator[54] application, Google Gemini AI[55] 2.5 and 3 LLM models, and Grammarly[56] writing tool have been used as translation tools in this research. Gemini has also been used for language maintenance, such as proofreading and sentence structure analysis.

[54] See DeepL Translator: https://www.deepl.com/en/translator.
[55] See Google Gemini AI: https://gemini.google/about/.
[56] See Grammarly: https://app.grammarly.com/.

2 Voice as a legal concept

This chapter discusses what the voice is, how it is formed and why protecting the voice is essential. A person's personality and identity are central to protection. Therefore, this chapter also discusses what personality and identity mean and why protecting those values is considered necessary.

2.1 What is voice, and why is it important to protect?

In this research, the starting point for defining voice is that the voice refers to naturally produced speech, singing, or other similar vocalisations coming from the human mouth.

Jody Kreiman and *Diana Sidtis* explain that defining voice isn't as easy as it seems. They point out that it's tough to find one definition that works for everyone because of the broad range of functions subserved by voice. They share the view with voice scientist *Johann Sundberg*, who has also noted that we all know what voice is until we try to really pin it down.[57]

The first idea for a definition of a voice would be the set of sounds produced by a normally functioning human vocal tract. The vocal tract is defined as basically the air passages in our head and neck, along with the parts we can move to change those passages' shapes and sizes. However, this definition would be too broad. It includes many sounds we don't usually think of as voice. For example, it would cover involuntary sounds such as coughs, sneezes, breathing noises, sniffs, hiccups, and chewing or swallowing sounds. It also includes small clicks and pops from our tongue or lips that aren't part of speaking.[58]

These reflexive sounds, such as a cough or a gasp, can be recognised as voices and sounds, and through this, they can be identified as part of someone's personality. Through those voices, one might instantly know who's coughing or sighing just by the sound of it. While these voices could deserve protection, especially if they are unique to an individual, they are outside the main focus of this research. This is because, unlike speech or singing, these reflexive sounds are not typically copied or imitated to the same extent.

---

[57] *Kreiman – Sidtis*, p. 5.
[58] *Watt et al.*, p. 139.

Thus, based on the aims of this research, a better definition for voice could be the set of sounds produced by the human vocal tract that are either intentional acts of communication (like speech and singing) or expressive vocalisations that carry identifiable characteristics of the individual, excluding involuntary or reflexive noises. This definition allows us to focus on the intentional and identity-bearing aspects of voice that are most relevant to the idea of protection.

Beyond just speaking or singing words, the voice can also make many other important sounds. Think about cries of alarm, anger, surprise, delight, or even pain. Laughter, sighs, and sounds such as "ums" and "ahs" that occur during pauses in speech are all part of the voice. Even humming a tune, sobbing, or weeping are ways our voice expresses something, even if there are no words involved. What is crucial here is that these sounds, even without words, can still tell a listener who is making them, especially if the listener knows the person well.[59] This shows how deeply connected our voice is to our own unique identity.

It is additionally important to note that voice is not just about the person making the sound. As important as making the sound is the listening, it is argued that the speech is not truly meaningful without someone to hear it.[60] The voice is argued to reflect how we present ourselves to others in social settings. It takes the whole brain and, by extension, the whole person to participate in producing and perceiving a voice.[61] So, a voice fully comes alive when a listener is there to experience it. Thus, it could be argued that a voice is a very personal and sensitive feature.

When we evaluate how the voices sound, the listener's judgment is at the centre. While assessing how similar two voices are, how unique another voice is, or how familiar a voice sounds, all these judgments are made by the listener. While recognising unauthorised use, the similarity of voices is the most crucial factor. Even if voices sound alike, it does not mean they are actually the same. It is important to note that, with over 8 billion people in the world, there is a high chance that some voices sound very similar.

Finally, the definition of voice that how we see it in this research is the following from the article by *Watts, Harrison* and *Cabot-King*:

---

[59] Ibid., p. 140.
[60] Ibid.
[61] *Kreiman – Sidtis*, p. 5.

> "The voice is the output in the sound domain of the coordinated movements of an individual human talker's speech organs, which is used principally but not exclusively by the talker for the acoustic encoding of linguistically meaningful utterances, and which after a period of exposure listeners may come to associate with that talker."[62]

The definition is well suited for the research and its focus on technological aspects and AI. It covers both productive and perceptual aspects, meaning it explains how a voice is physically produced and how listeners recognise it.[63] The definition fits the research well, especially since it specifically examines the unauthorised use of human-generated voices, which are voices that someone has control over.

2.2 Use of voice in emerging technologies

In the context of voice unauthorised use of a person's voice in the emerging technologies, it is essential to present the methods, technologies and techniques that are used to clone a voice. The first attempts to create human voices using technology began in the mid-1930s with voice synthesisers.[64] One of the first significant projects was "Voder" from Bell Labs at the 1939 World's Fair. At the beginning, producing recognisable voices required manual operation, which further emphasised the difficulty of creating a human voice in the early days.[65]

However, the tools continued to develop and become more realistic. In the 1980s, for example, the "DECtalk" speech synthesiser was created, which *Stephen Hawking* famously used.[66] "DECtalk" used acoustic-phonetic rules to create speech sounds. Additionally, the "Klatt" voice synthesiser was created in the 1980s by AT&T Bell Laboratories, which used a set of acoustic-phonetic rules to generate speech sounds, producing more natural-sounding output than its predecessors.[67]

The 1990s brought concatenative synthesis tools, which relied on pre-recorded speech segments to create new utterances.[68] Concatenative synthesis relied on a comprehensive database of recorded speech segments, which included phonemes, diphones, and triphones, to capture natural speech dynamics. When text input was received, the system analysed the

---

[62] Ibid., p. 144.
[63] Ibid.
[64] Speechify 2025.
[65] What Is the Voder 2024; Grundhauser.
[66] Digital Equipment Corporation, DECtalk DTC-01.
[67] *Klatt*, p. 737.
[68] *Schwarz*, pp. 3-22.

phonetic sequence and employed algorithms to select the best-matching speech units from the database. These units were then concatenated to form the final speech output, with smoothing techniques that ensured seamless transitions.[69] One such synthesiser was AT&T's Natural Voices text-to-speech (TTS) system.[70]

Concatenative synthesis tools were followed by Hidden Markov Models (HMMs) and HTS systems, which used statistical techniques to analyse and recreate the patterns of human speech, resulting in speech that sounded increasingly natural.[71] However, early voice replication tools were relatively limited, and the output was often robotic and unnatural. Moreover, they were slow to use, as concatenative synthesis required extensive manual processing and editing.[72]

Significant developments in deep learning technology in the 2010s addressed the shortcomings of previous systems. Deep learning systems proved their effectiveness by effectively learning the temporal dynamics of voice, including intonation and rhythm. This led to a more natural-sounding synthesised voice with improved prosody.[73]

WaveNet, introduced by Google in 2016, made a breakthrough in speech synthesis. It used a type of neural network called a convolutional neural network (CNN) to model raw audio waveforms directly. This allowed WaveNet to capture subtle details in voice, like intonation and pronunciation, resulting in much more natural-sounding synthetic voices. However, WaveNet had a significant drawback, as it was computationally intensive and slow to generate audio.[74]

The next significant development came in 2017 in the form of the Transformer architecture. Originally designed as a translation tool, Transformer proved to be groundbreaking in many areas of AI, including speech synthesis. Transformer was revolutionary due to its attention mechanism, which allows the model to focus on relevant parts of the input when generating output, leading to more accurate and natural-sounding speech. Unlike previous models that processed inputs in series, Transformers were able to process entire sentences at once, making them much faster to train and use. They are excellent at understanding context over longer

---

[69] *Fenwick et al.*, p. 8.
[70] *Juang – Rabiner*, pp. 2-24.
[71] *Tokuda et al.*, pp. 1-4.
[72] *Schwarz*, p. 1.
[73] *Fenwick et al.*, p. 9.
[74] See WaveNet: A Generative Model for Raw Audio, DeepMind 2025.

series, resulting in more coherent and contextually appropriate speech synthesis. Moreover, the Transformer architecture's versatility allowed it to be adapted for various tasks, including text-to-speech, voice cloning, and even controlling aspects like emotion and style in synthesised speech.[75]

These developments in Transformers were quickly adopted in AI. For example, Google's Tacotron 2[76] and NVIDIA's FastSpeech[77] used Transformer-based architectures to create high-quality voice synthesis that was not only natural-sounding speech but also able to mimic specific voices with remarkable accuracy. Later, for example, Meta's VoiceBox also implemented Transformers' techniques in its model. These hybrid approaches have pushed the boundaries even further, allowing unprecedented control over various aspects of synthesised speech, including voice style, emotion, and even background noise.[78]

Generative Adversarial Networks (GANs) have also made significant contributions to speech synthesis. GAN-based synthesisers, vocoders, like MelGAN[79] and HiFi-GAN[80] have demonstrated the ability to produce high-quality speech waveforms in real-time, addressing some of the speed limitations of earlier models like WaveNet. These GAN-based vocoders have become popular components in many state-of-the-art text to speech systems, offering a balance between audio quality and generation speed while working alongside other models that handle text analysis and acoustic feature prediction.[81]

Another notable advancement came with Google Research's AudioLM (2022), which applies language modeling techniques to audio generation. AudioLM represents a novel approach that can generate high-quality, coherent audio continuations given an audio prompt. Unlike previous models that focus solely on speech, AudioLM can handle a range of audio types, including speech, music, and sound effects.[82]

Furthermore, in March 2024, OpenAI released a new text-to-speech model as part of ChatGPT's voice chat feature.[83] The model represents a significant advancement in voice AI,

---

[75] *Vaswani et al*., pp. 1-11.
[76] See NVIDIA Tacotron 2, GitHub.
[77] *Yi Ren et al*., pp. 1-15.
[78] See Introducing Voicebox, Meta AI.
[79] *Kundan Kumar et al*., pp. 1-14.
[80] *Kong et al*., pp. 1-14.
[81] Ibid.
[82] *Borsos et al*., pp. 1-11.
[83] OpenAI 2024.

capable of generating natural-sounding speech in many different languages and accents from as little as a 15-second audio sample. It can also produce speech using different moods and styles, further blurring the lines between human and machine-generated voices.[84]

2.3 Use of voice in AI technologies

Artificial intelligence has played a key role in the development of voice cloning tools, even though the concept of artificial intelligence has not been precisely defined, nor has there been any consensus on its definition. Artificial intelligence researcher *John McCarthy* has described artificial intelligence as follows:

> "It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."[85]

In *Stuart Russell* and *Peter Norvig's* book *Artificial Intelligence: A Modern Approach*, the definition of artificial intelligence is divided into different approaches. The classification is based on what the creation of artificial intelligence aims to achieve.[86] First, artificial intelligence can mimic human behaviour or human thinking. Second, artificial intelligence can be capable of thinking or acting rationally. Rationality refers to logical reasoning and drawing conclusions.[87]

Artificial intelligence systems can be divided into narrow, general, and super-intelligent systems. Narrow artificial intelligence systems (ANI) can perform one limited task at a time. Examples of such systems include Apple's Siri and Amazon's Alexa. The intelligence of general artificial intelligence systems (AGI) is comparable to that of humans. Such systems are capable of independently solving problems, learning, and planning for the future. The intelligence and capabilities of Artificial Super Intelligence (ASI) systems would exceed that of the human brain.

Ismo Kallioniemi in his book *Law of Artificial Intelligence* has taken the position that artificial intelligence is an umbrella term covering other concepts that are central to the implementation of artificial intelligence.[88] Machine learning is a key technology in the implementation of

---

[84] *Fenwick et al.*, p. 11.
[85] *McCarthy*, p. 2.
[86] *Russell – Norvig*, pp. 1-4.
[87] Ibid.
[88] *Kallioniemi*, p. 14. An AI law book in Finnish that examines key legal issues related to AI systems.

artificial intelligence systems, based on the system's ability to adapt to different situations, even if the correct action in a given situation has not been specifically defined. In this case, the system or machine learns.[89]

Neural networks are a method used to implement machine learning, in which models that mimic brain activity are used in the functions of artificial intelligence systems. Deep learning is a way of teaching neural networks, where deep learning methods can be used to create depth in neural networks, which in turn enables the processing of larger amounts of data.[90]

Machine learning algorithms are more dependent on human activity, as they are simpler in their technical structure and operate more directly in accordance with the instructions set by the programmer. Deep learning algorithms operate more independently and do not require the same instructions set by the programmer. Deep learning is based on the machine learning from its own observations and mistakes, while teaching neural networks to operate rationally.[91]

Artificial intelligence systems require teaching material, which is fed into the system by the system developer. This material consists of data that is a collection of data elements that has been compiled, organised, and refined to teach the system in an appropriate manner. Data can include algorithms, texts, and images. In AI that models voice, these data sets could be audio recordings or other data that contains voice, for example.[92]

The operation of artificial intelligence systems is controlled by a computer program which aims to implement the algorithms set by the programmer or developer. In order for the computer program to be able to operate according to the algorithms, the algorithms must be expressed in a way that the computer program can understand in the form of a programming language. Coding refers to the translation of human language into a programming language that computer programs can understand.[93]

Generative AI models hold the most significance for the research and human voice perspectives. Unlike traditional AI, which is often used for classification or prediction, generative AI is designed to produce new content. The development of generative AI has been

---

[89] Ibid., p. 15.
[90] Ibid p. 16.
[91] IBM 2024.
[92] *Kallioniemi* 2022. p. 19.
[93] Ibid., p. 18.

rapid in recent years. This development has been based on innovations such as neural networks, deep learning, generative adversarial network models, and transformer models. To understand the operating principle and applicability of generative artificial intelligence, it is essential to understand the following two points. Firstly, the large language, image, and other models that are behind the AI have been trained on large amounts of text, images, or other data. Secondly, when AI produces or generates content, it does not cut and paste text, images, or voice clips from any database or the internet. Instead, the generative AI produces genuinely new and unique content. In other words, it does not copy existing content straight away, instead AI produces something new. [94]

Modern voice cloning systems can create highly accurate audio from just a few seconds of recorded speech.[95] This efficiency is made possible by powerful machine learning algorithms and the availability of those extensive computational resources. For example, LibriTTS[96] and Common Voice[97], enable AI models to learn from large amounts of speech data while improving overall performance and scalability. Advanced "Few-Shot" and "Zero-Shot Learning" techniques allow AI models to clone voices that the models have not encountered before, based on minimal reference material.[98] For example, NVIDIA's RAD-TTS (2021) can generate natural speech in the voice desired by the system user after hearing only a few utterances as a model.[99]

Although these systems have developed rapidly in recent years, further development is still coming. The rapid development of AI systems has also meant that the devices on which the systems are used must be high-performance. This affects, for example, the quality of output, generation speed, and latency. In addition, the systems still work most efficiently with the most commonly used languages. Languages with fewer speakers and, therefore, less data naturally pose challenges for the systems.[100]

Currently, there are many applications for voice generation. The most well-known systems for generating voices are generative AI applications such as ChatGPT, Copilot, and Gemini,

---

[94] *Salo*, p. 11. Finnish AI book, generative AI as a superpower in business.
[95] Ibid.
[96] *Zen et al*., pp. 1-7.
[97] See Common Voice, Mozilla 2025.
[98] *Zhang – Lin*, pp. 1-5.
[99] See Generate Natural Sounding Speech from Text in Real-Time, NVIDIA 2019.
[100] *Jia – Tadmor Ramanovich*, blog text.

which not only understand the user's speech but can also respond to it.[101] Additionally, there are also applications that specialise purely in voice and music, such as ElevenLabs[102] and Udio[103]. ElevenLabs was founded in 2022 and focuses on the use of artificial intelligence in voice production. ElevenLabs' AI models can produce speech in hundreds of different voices and in 29 different languages. ElevenLabs offers two options for creating synthetic voices: Instant Voice Cloning and Professional Voice Cloning. The first option allows voices to be created from short voice samples, meaning that a clip of speech lasting just over a minute without background noise is sufficient as a model. The Professional version, on the other hand, requires significantly more background material (2-6 hours), but the result is correspondingly better.[104]

Udio is an AI application for producing music, founded in 2024. In the application, the user gives a prompt and selects a music style, and after a moment, a sample of lyrics based on the prompt set to music is produced. Users can also enter their own lyrics or lyrics generated by ChatGPT, Gemini, Copilot, or similar tools into the application and request that the lyrics be set to music in the selected style. By default, Udio first creates a 30-second-long sample, which can be extended in the desired direction.[105]

2.4 Technology and personal rights facing challenges in protection

The rapid development of voice cloning technologies has exposed a fundamental weakness in legal frameworks. The main issue with the current relevant legal frameworks is that they have been designed to regulate markets, protect consumers, and facilitate competition, rather than safeguard the individual's personality. The human voice has therefore been left without robust safeguards against technological innovation. The voice and other personal traits historically attracted little legal attention precisely because there has not been a technological capability to replicate them on a mass scale. Following the rapid developments of generative artificial intelligence, the voice has become vulnerable to unauthorised reproduction, manipulation, and commercial exploitation.[106]

---

[101] Ibid., p. 145.
[102] See ElevenLabs: https://elevenlabs.io/about.
[103] See Udio: https://www.udio.com/about-us.
[104] *Salo*, p. 146.
[105] *Salo*, p. 147.
[106] *Fenwick et al.*, p. 15.

At the international law level, the risks of artificial intelligence have already been acknowledged in broader human rights discourse. AI systems deployed in surveillance, behaviour prediction, and governance contexts can directly undermine individual freedoms and privacy.[107] These concerns also apply to voice cloning technologies, when the possibility of replicating voice without authorisation might be a direct infringement towards individuals' autonomy and dignity. On the other hand, the international human rights framework may provide a legal framework for strengthening the basis of protection, but it has not been interpreted to address the risks created by generative AI.

The lack of existing regulatory frameworks becomes clear when viewed through this lens. Current doctrines, whether in copyright, consumer protection, or competition law, rest on the rationale of ensuring healthy competition, consumer welfare, and innovation, not on safeguarding the personal interests of individuals whose voices are being used in technologies without permission.[108] This reflects a structural bias in the regulation, as the law has historically prioritised market stability over the protection of personal identity. The result is a regulation that is capable of addressing issues such as consumer deception or market abuse, avoids addressing the core issue of voice cloning, namely the unauthorised use of a voice as a personal characteristic.[109]

AI-generated covers, where an artist's voice is used, without permission, to perform songs they never recorded, has raised widespread concerns broadly among the music industry. Singer *Cher's* description of the phenomenon as "just out of control"[110], captures the sense of powerlessness felt by artists when confronted with unauthorised imitations of their most personal attribute.[111] As *Antonios Baris* explains, the training of generative models on artists' discography highlights the tension between copyright law and voice rights.[112] While copyright may regulate the use of artists' works, it is structurally unsuited to address the appropriation of the voice itself. For example, the European Union has built its regulation on these bases without any success. That approach to image and personality rights has only proved this difficulty and left individuals without protection and legal remedies for their rights.

---

[107] *Liu*, p. 76.
[108] *Fenwick et al.*, p. 15.
[109] Ibid.
[110] Sky News 2023.
[111] *Baris*, p. 571.
[112] Ibid.

3 Copyright to protect the use of voice?

After stating the difficulties in protecting the voice in the emerging technologies, this chapter's purpose is to determine whether the human voice can be protected by copyright. The aim is to decide whether it is possible that a person is the author of their own voice and that the voice would therefore be protected by copyright. On the other hand, if that is not possible, the aim is to explain why the human voice itself cannot be copyrighted and whether it could be protected in some other way.

Copyrights are universally recognised intellectual property rights that create exclusive rights in "literary and artistic works" that are "original".[113] The scope of the copyright protection ranges from general art, such as poems and paintings, to usual information products, such as maps and almanacs. In general, the scope should be interpreted extensively, and it includes all kinds of different works that are regarded as original.[114] With technological progress, creative works today take many different forms and can be produced through a variety of methods. This increasing diversity has subsequently been recognised in several supplementary treaties. Given the wide range of works protected by copyright, it is therefore important to focus on the *originality* of a work when determining what qualifies for copyright protection. At first, a work must be original and demonstrate a certain degree of distinctiveness to be considered copyrightable.[115] There are, however, no universal standards specifying the degree of originality or uniqueness required. Each case must therefore be assessed individually, although the work typically needs to exhibit some level of distinctiveness. It should also be noted that ideas themselves are not protected by copyright. Rather, it is the *expression* of an eligible idea. The idea must be expressed in some form to fall within the scope of copyright.[116]

Copyright is generally granted to the author of the work, unless other arrangements are made before or after its creation. One of the key characteristics of copyright is that it arises automatically. Unlike many other forms of intellectual property, no application or registration is usually required for work to be protected and for exclusive rights to take effect.[117]

---

[113] Berne Convnetion, 1971 Paris Text Art. 2. See §6.1, below.
[114] *Goldstein – Hugenholtz*, p. 4.
[115] *Cass – Hylton*, p. 104.
[116] Ibid.
[117] *Burley*, p. 41.

The owner of the copyright enjoys robust rights that protect their work. These rights are usually divided into economic and moral rights.[118] Economic rights grant the copyright holder the exclusive right to authorise or prohibit specific methods of exploiting their work. The nature of these exclusive rights implies that no third party is permitted to exercise them without the owner's explicit consent. The main idea behind economic rights is to safeguard the author's or owner's financial interests and ensure they can benefit from potential commercial gains. Substantively, the scope of these rights encompasses the reproduction of the work in various formats, ranging from printed publications to digital files, as well as the distribution of copies and the public performance of the material. Furthermore, it extends to the broadcasting or other communication of the work to the public, including making it available online for on-demand access, alongside the rights of translation and adaptation, such as the transformation of a literary novel into a cinematic screenplay.[119]

Moral rights protect an author's creative integrity and reputation as expressed through the work. Moral rights consist of two main rights.[120] The first of those rights is to be named as the author of the work. When the work of an author is reproduced, published, made available or communicated to the public, or exhibited in public, the person responsible must make sure that the author's name appears on or in relation to the work, whenever reasonable.[121] The second right is the right to protect the integrity of the work. This prohibits the making of any changes to a work that would potentially damage the author's honour or reputation.[122]

3.1 Berne Convention and voice as a protected work

The Berne Convention for the Protection of Literary and Artistic Works (Berne Convention, as amended on September 28, 1979) is widely recognised as the cornerstone of international copyright law. It currently includes nearly two hundred member states forming a union and making it the most comprehensive and influential copyright treaty in force.[123] Established in the late nineteenth century[124], revised through the Paris Act of 1971 and with the latest amendments made in 1979, the Convention predates the digital and AI era by several decades.

---

[118] Ibid., p. 44.
[119] *Goldstein – Hugenholtz*, pp. 297-300.
[120] *Burley*, p. 46.
[121] *Goldstein – Hugenholtz* 2010, pp. 349-350.
[122] Ibid., p. 351.
[123] Berne Convention Article 1.
[124] *Goldstein – Hugenholtz*, pp. 33–35.

Despite its age, its principles continue to shape the framework for the copyright systems and remain central to international discussions on the protection of creative works.[125]

The Berne Convention is founded on three fundamental principles: national treatment, automatic protection, and independence of protection. The principle of national treatment requires that works originating in one member state receive the same protection in all other contracting states as those countries grant to works of their own nationals.[126] The principle of automatic protection, codified in Article 5(2), ensures that copyright protection is not subject to any formality, such as registration or deposit.[127] The independence of protection means that the existence or scope of protection in one member state is not dependent on whether the work is protected in its country of origin.[128]

Beyond these core principles, the Convention sets minimum standards for protecting both economic and moral rights. Economic rights grant authors control over how their works are used and exploited, such as through reproduction, translation, adaptation, and public communication. Moral rights, stated in Article 6bis, protect the personal and reputational connection between an author and their work, including the rights to claim authorship and to oppose any distortion, mutilation, or modification that could harm the author's honour or reputation. These moral rights stay with the author even after economic rights are transferred, highlighting the human and personal aspect of authorship.[129]

The Convention also provides for limitations and exceptions towards the copyright. Among these, Article 9(2) establishes the three-step test, which permits the reproduction of protected works only (1) in certain special cases, (2) provided that such use does not conflict with the normal exploitation of the work, and (3) does not unreasonably prejudice the legitimate interests of the author.[4] This test has become a cornerstone of international copyright law and has been incorporated, with minor variations, into subsequent treaties such as the TRIPS Agreement and the WIPO Copyright Treaty. Other exceptions, such as those for educational or press purposes, are typically left to national legislation but must comply with the principle of fair practice.[130]

---

[125] *Ricketson – Ginsburg*, p. 121.
[126] Berne Convention Article 5(1).
[127] Berne Convention Article 5(2).
[128] Berne Convention Article 5(3).
[129] Berne Convention Article 6bis.
[130] Berne Convention Article 9(2).

While the Berne Convention provides a robust framework for protecting creative expression, its nineteenth-century origins reveal significant limitations when applied to modern technologies, especially AI and its voice generation. The Convention's fixation requirement, reflected in Article 2(2) as works in general or any specified categories of works shall not be protected unless they have been fixed in some material form, allows member states to demand that works be expressed in a material form before protection applies. This requirement effectively excludes unrecorded human voices from copyright protection, as they do not exist in a fixed medium. Consequently, the unrecorded qualities of a person's voice and its tone, rhythm, and individuality remain beyond the reach of copyright law, even though they may represent distinctive aspects of personal identity and artistic expression. That means, where voice has been fixed in a recording, whether in analogue from on a wax cylinder or vinyl disc, on magnetic tape or as a digital copy held in whatever way, that recording is protected from being copied. However, it should be noted that the recording will be owned by the person who made it, not by the person whose voice is recorded.[131]

Furthermore, while moral rights provide authors with personal protection over their creative works, they presuppose the existence of a copyrightable work. In the context of AI-generated or imitated voices, this becomes problematic because when voices themselves are not recognised as works, individuals lack rights to claim moral rights against unauthorised or misleading reproductions. As a result, generated voice imitations created by AI systems fall outside the protective scope of the Berne framework. Since many national copyright laws are based on, or have been harmonised in accordance with, the Berne Convention, this limitation is commonly reflected at the national level as well. Consequently, such uses also remain unaddressed within domestic copyright systems. This is, however, a natural consequence of the Convention's age and historical context. As a product of the nineteenth century, the Berne Convention could not have anticipated modern technological developments such as artificial intelligence or synthetic voice generation. This technological gap requires a reassessment of the current legal framework to take into account the unique challenges posed by artificial intelligence, particularly with regard to personal rights and intellectual property rights.[132]

---

[131] *Watt et al*., p. 152.
[132] *Chopra et al*., p. 49.

3.2 WIPO Copyright Treaty

The WIPO Copyright Treaty (WCT), adopted in Geneva on December 20, 1996 and entering into force on March 6, 2002, was designed to supplement and modernise the Berne Convention in light of the digital revolution.[133] As one of the so-called "Internet Treaties", alongside the WIPO Performances and Phonograms Treaty (WPPT), the WCT aims to ensure that copyright protection remains effective and relevant in a technology-driven environment.[134] It builds upon the principles of the Berne Convention rather than replacing them, recognising that digital technologies and global networks introduced new modes of exploitation and new challenges to the control and enforcement of rights.

The WCT confirms that all provisions of the Berne Convention remain applicable to its contracting parties, with the exception of the moral rights provisions of Article 6bis, which it leaves unchanged.[135] Its principal contribution lies in clarifying and extending existing rights to cover digital uses and in establishing new obligations regarding technological protection measures and rights management information.[136] The treaty recognises early stages that digital reproduction, distribution, and communication demand legal certainty, particularly as creative works can now be shared globally with minimal cost and effort.

One of the most significant clarifications introduced by the WCT concerns the scope of the right of reproduction under Article 1(4) in connection with Article 9 of the Berne Convention.[137] The treaty confirms that storing a protected work in digital form constitutes reproduction, even when the storage is temporary or transient. This interpretation ensures that digital copies, whether in a computer's memory, on a server, or within an AI training dataset, are subject to the author's authorisation. In addition, the WCT explicitly recognises the right of communication to the public and the right of making available (Article 8), which together grant authors control over the interactive and on-demand sharing of their works through digital networks.[138] These provisions were intended to address the rise of online platforms and streaming services, establishing a legal foundation for digital content distribution long before such models became presented everywhere.

---

[133] WIPO, *WCT* (1996).
[134] WIPO, Guide to the Copyright and Related Rights Treaties, p. 219.
[135] WIPO, *WCT*, Art. 1(4).
[136] *Ricketson – Ginsburg*, p. 135.
[137] WIPO, *WCT*, Art. 1(4); Berne Convention, Art. 9.
[138] WIPO, *WCT*, Art. 8.

The WCT also introduces obligations related to technological measures and rights to management information (Articles 11–12).[139] Contracting parties must provide legal protection and effective remedies against the rotating of technological measures that authors use to control access to their works, and against the removal or alteration of electronic rights management information.[140] These obligations reflect the treaty's policy objective of supporting the digital economy by linking copyright protection with technological enforcement mechanisms. However, they also raise concerns regarding users' rights and the potential restriction of authorised uses, as digital locks may sometimes prevent access even for permitted exceptions.

In line with the Berne framework, the WCT maintains the three-step test for limitations and exceptions in the Article 10, confirming that exceptions must be confined to certain special cases, must not conflict with the normal exploitation of the work, and must not unreasonably harm the author's legitimate interests.[141] At the same time, the treaty acknowledges the need for flexibility by allowing countries to adapt these exceptions to digital and networked environments in the digital age.[142]

From the perspective of emerging technologies such as AI and synthetic voice generation, the WCT's relevance lies in its broad approach to digital reproduction and communication. By extending the scope of reproduction to digital copies, the treaty potentially covers the unauthorised use of protected sound recordings or performances in AI training datasets. However, the WCT, similarly to the Berne Convention, requires the existence of a protectable work or performance. It does not directly address the unauthorised use case or synthesis of human characteristics, such as voice, when those characteristics are not materialised in a copyright-protected work.[143] Consequently, while the WCT offers stronger protection for digital copies and networked sharing, it remains unable to address the unauthorised generation of human-like voices by AI systems.

The WCT thus represents a critical transitional step between the analogue principles of the Berne Convention and the digital realities of contemporary creativity. Its provisions

---

[139] WIPO, *WCT*, Art. 11–12.
[140] WIPO, Explanatory Notes on WCT, paras 10.10–10.20.
[141] WIPO, *WCT*, Art. 10; *Berne Convention*, Art. 9(2).
[142] *Gervais*, pp. 133-137.
[143] *Ginsburg*, p. 6–9.

modernised copyright for the online era, but they stop short of engaging with the modern age's emerging questions surrounding personal data and AI use.

3.3 The TRIPS Agreement and other international copyright treaties

The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), adopted in 1994 as Annex 1C to the Marrakesh Agreement establishing the World Trade Organization, is one of the most influential instruments shaping contemporary international copyright norms. Unlike the Berne Convention and the WIPO Internet Treaties, which are administered by WIPO, the TRIPS Agreement is part of the global trade regime and backed by the WTO's dispute settlement system.[144] Mainly, the TRIPS regulates trade related aspects of copyrights among other forms of intellectual property rights. Copyrights and trades are naturally closely linked.[145] The economic rights granted by copyright are transferable, and the TRIPS requires protection only for those economic rights. Recently, we have seen various artists selling these rights to companies, for example, in the music and entertainment industry. TRIPS has contributed significantly to the harmonisation of copyright standards globally.

TRIPS includes by reference the substantive provisions of the Berne Convention (1971) with the exception of moral rights.[146] As such, it then requires WTO members to respect the core Berne principles of national treatment, automatic protection, and independence of protection, while also adding new provisions that address trade and technology-driven issues.[147] The Agreement introduces further specificity regarding exclusive rights, including the reproduction right, rental rights for computer programs and cinematographic works, and protection for compilations of data.[148] These provisions reflect an early recognition of the growing economic relevance of digital technologies and related industries.

One of the most significant contributions of TRIPS is its elaboration of enforcement obligations. Part III of the Agreement requires member states to provide effective civil, criminal, and border enforcement mechanisms capable of preventing infringements.[149] This inclusion of enforcement standards differs TRIPS from earlier copyright treaties, making it not only a harmonisation instrument but also a framework for regulatory compliance. Given

---

[144] *Gervais*, p. 89–92.
[145] *Goldstein – Hugenholtz*, p. 71.
[146] TRIPS Agreement, Art. 9(1).
[147] *Gervais*, pp. 94–95.
[148] TRIPS Agreement, Arts. 11–14.
[149] *Gervais*, pp. 285-289.

the prominence of global trade in digital goods and the circulation of content across platforms, the TRIPS enforcement regime remains central to discussions on the effectiveness of international copyright protection in the digital age.

TRIPS also includes the three-step test for limitations and exceptions to exclusive rights (Article 13), mirroring the Berne Convention's formulation in Article 9(2).[150] While the test originally applied only to reproduction rights under the Berne Convention, TRIPS extends it to all exclusive rights, thereby establishing a global normative benchmark for the permissible scope of user freedoms. As with the WCT, however, the three-step test has been criticised for its undefined nature and for potentially narrowing the space for domestic copyright exceptions, particularly in relation to new technologies.[151]

Again, from the perspective of AI-generated voice uses, TRIPS does not directly address issues of personality, likeness, or human characteristics. Consistent with the Berne and WCT frameworks where it is also based, TRIPS requires the existence of a copyright-protectable work and therefore offers no protection for individuals whose voices are imitated unless the imitation involves an underlying protected work such as a voice recording or performance.[152] While the Agreement requires protection for performers' rights in recordings, it does not regulate personal attributes or identifiers that are not captured or recorded in such works.[153] As a result, issues related to the research remain outside the scope of TRIPS.

In addition to TRIPS, several other international copyright agreements are relevant to the broader legal context. The Rome Convention (1961) established protection for performers, producers of phonograms, and broadcasting organisations, formed the historical foundation for neighbouring rights.[154] The WPPT (1996) modernised this protection for the digital environment by granting performers rights over the digital use and sharing of their performances.[155] The Beijing Treaty on Audiovisual Performances (2012) further extended those rights in audiovisual works, though its global implementation has been inconsistent.[156] Together, these treaties describe the progressive expansion of neighbouring rights, but none of them directly address non-work-based characteristics such as human voice.

---

[150] TRIPS Agreement, Art. 13.
[151] *Griffiths*, pp. 10-11.
[152] *Ginsburg*, p. 7–9.
[153] TRIPS Agreement, Art. 14.
[154] Rome Convention (1961).
[155] WPPT (1996).
[156] Beijing Treaty (2012).

Overall, the TRIPS Agreement and the broader network of international copyright treaties have significantly shaped the global copyright landscape, especially related to digital technologies and global enforcement. However, like the Berne Convention and the WCT, TRIPS is also primarily oriented toward the protection of works and performances, rather than personal attributes such as voice.

3.4 The fixation requirement of the voice and voice as identity

The previous analyses of the Berne Convention, the WIPO Copyright Treaty, and the TRIPS Agreement have consistently revealed a structural limitation of tendency within the international copyright framework. This framework, designed to protect "literary and artistic works" and "performances", is fundamentally not able to provide direct protection to the human voice as a personal characteristic. This subchapter will first identify the primary legal mechanism, the "fixation" requirement. It will then define why even the moral rights doctrine that is very close to the author's personality and protects honour and reputation, is not able to provide any protection for the voice. Finally, the analysis will present an alternative *sui generis* framework, the right of publicity that protects voice not as a work, but as an attribute of identity.

The common factor among the international treaties examined in this chapter is that they protect expressions such as literary and artistic works, phonograms, and fixed performances rather than the personal attributes of their authors. The human voice itself, in its unrecorded state, is not considered as a work in the copyright sense. The central legal barrier that forms this barrier is the requirement of fixation. This barrier originates in the Berne Convention itself. While the Convention is known for establishing automatic protection under Article 5(2), which states that the enjoyment and exercise of rights shall not be subject to any formality[157], it then provides a critical, discretionary exception for national legislations. Article 2(2) of the Berne Convention explicitly states: "It shall, however, be a matter for legislation in the countries of the Union to prescribe that works in general or any specified categories of works shall not be protected unless they have been fixed in some material form".[158]

[157] Guide to the Berne Convention for the Protection of Literary and Artistic Works, p. 33.
[158] Ibid.

This requirement is not based on a desire to exclude individuals' personal characteristics from protection. Instead, it was a 19th-century practical and evidentiary compromise. Fixation in a material form provided tangible proof of a work's existence and content, which was necessary to resolve infringement disputes in a court of law.[159] This historical, evidentiary based rule has now, in the 21st century, become the central legal loophole that technologies, such as AI voice cloning, are able to exploit. The law does not take into account harm related to a person's voice identity itself.

For example, the United Kingdom's Copyright, Designs and Patents Act 1988, which forms UK copyright law, is a prime example of a Berne member state that explicitly requires fixation, such as in a recording, for a work to have protection.[160] Consequently, the unrecorded voice, with its unique timbre, pitch, and cadence that defines a person's vocal identity, is not a work under the international copyright system. It fails at this first and most fundamental requirement, not by oversight, but as a structural feature of the entire framework.

At first glance, the infringement caused by AI voice cloning appears to be a moral rights violation. For instance, an inappropriate song that uses the person's voice without permission could be seen as a direct attack on that individual's public persona. This aligns conceptually with the right of integrity under Article 6bis of the Berne Convention, which grants an author the right to object to any distortion, mutilation, or other modification of the work, which would be prejudicial to the author's honour or reputation.[161] This right seems ideally suited to protect the voice. However, that legal doctrine fails after a closer inspection. The text of Article 6bis itself contains the reason for its inapplicability: the rights of paternity (right to claim authorship) and integrity (right to protect the work) are granted to the author in relation to the said work.[162] Moral rights, just like economic rights, are connected to and utterly dependent upon the existence of a copyrightable work.[163] If there is no work to protect, especially when an unfixed voice is not considered as a work, then the moral rights protection is not applicable.

This creates some legal irony as moral rights were introduced into the Berne framework to protect the deep, personal, non-economic link between an author and their creative output,

---

[159] Ibid.
[160] Copyright, Designs and Patents Act 1988, para. 3(2).
[161] Berne Convention Article 6bis (1).
[162] Ibid.
[163] *Makovec Petrik*, pp. 366-367.

recognising that a work is an extension of the author's personality.[164] The harm from voice cloning is a direct attack on that very personality and the individual's honour or reputation.[165] Following that, because the legal doctrine is completely connected to a fixed work, the regulation protects the expression of personality in the form of recording or a script but is not able to protect the source of the personality, the voice itself.

3.5 Right of publicity to safeguard the voice

After presenting the failure of the international copyright regime to protect the individual's voice, the research presents the *sui generis* legal doctrine explicitly developed to address this gap. The most developed and relevant framework is the American right of publicity doctrine that exists largely outside of federal copyright law.[166] The doctrine is seen as a patchwork to fill the gap left by copyright law.[167]

The right of publicity is defined as an intellectual property right that protects against the unauthorised commercial use of a person's identity. Its scope is crucial as it explicitly protects the signs of identity, which include a person's name, likeness, image, signature, and, most significantly for the research, their voice.[168]

The doctrine's unique practicality is based on its clear legal origins. The right of publicity did not evolve from copyright or the right to privacy, instead of that it evolved from property law, recognising that an individual has a right to control and profit from the commercial value inherent in their own identity.[169] It is, therefore, seen as a commercial tort and enforceable when one appropriates the commercial value of a person's identity by using without consent the person's name, likeness, or other signs of their identity for purposes of trade.[170]

As seen in the previous subchapter the copyright failed to protect the voice because it is not a work. On the other hand, the right of publicity success in the protection because a voice is an sign of identity that has commercial value. The viability of the right of publicity doctrine is

---

[164] *Chisolm*, pp. 457-459.
[165] EditMojo.
[166] INTA 2019.
[167] *Post – Rothman*, p. 89.
[168] INTA 2019.
[169] *Lapter*, pp. 10-12.
[170] Free Speech Center, Misappropriation 2024.

described in 1988 by its landmark case of *Midler v. Ford Motor Co.*, which established voice as a protected attribute of identity.[171]

In this case, the Ford Motor Company and its advertising agency, Young & Rubicam, produced a series of commercials (the "Yuppie Campaign") designed to evoke 1970s nostalgia.[172] They sought to have singer *Bette Midler* sing her hit song "Do You Want to Dance", but Midler, who had a long-standing policy against appearing in commercials, refused the offer. Subsequently, the agency hired *Ula Hedwig*, one of Midler's former backup singers, and instructed her to sound as much as possible like the Bette Midler record. Following that, the successful commercial was made, and numerous listeners, including friends of Midler, were misled into believing it was genuinely Midler singing. [173]

The U.S. Court of Appeals for the Ninth Circuit was first forced to confront the copyright dilemma. It explicitly stated that Midler could not sue the company for copyright infringement. A voice, the Court defined, is not copyrightable because the sounds are not fixed in a tangible medium of expression. The Court stated that the Midler sought to protect something "more personal than any work of authorship".[174] However, the Court decided not to leave Midler without a remedy. It ruled that the defendants had committed a tort under California common law. In the case the Court's language was revolutionary when it stated that: "A voice is as distinctive and personal as a face".[175] By deliberately imitating her unique voice to sell their product, the defendants had appropriated what is not theirs.[176] The decision's most-cited line established the new principle, stating that "to impersonate her voice is to pirate her identity".[177]

This principle was tested and clarified in the following case, *Waits v. Frito-Lay, Inc.* (1992).[178] Singer *Tom Waits*, who was known for his distinctive gravelly voice, sued after an agency used a soundalike to mimic his singing style in a snackfood commercial. The Waits

---

[171] *Midler v. Ford Motor Co*., 849 F.2d 460.
[172] Ibid.
[173] Ibid.
[174] Chester-Taxin, p. 4.
[175] Ibid.
[176] *Midler v. Ford Motor Co*., 849 F.2d 460.
[177] Ibid.
[178] *Waits v. Frito-Lay, Inc*., 978 F.2d 1093.

Court not only upheld the Midler case claim but also found a separate violation under the federal Lanham Act, a law governing trademarks and unfair competition.[179]

The Waits case helped to establish the legal principle. The infringement is complete when the imitation is used for a commercial purpose, and it creates a likelihood of confusion. It does not matter whether the voice used in the copy is actually the person's own voice or an attempt to imitate it, the principle focuses on whether the public is led to believe that the original artist has endorsed or is associated with the product.[180] The combination of Midler and Waits cases provides a robust legal test. When Midler establishes the *sui generis* property right in one's vocal identity, Waits clarifies the infringement as misleading use of that person's identity for commercial gain.

The legal problems addressed in the Midler and Waits cases involved human imitators and a commercial creation process that was costly, difficult, and time-consuming. The development of AI, deep learning, and neural networks has democratised and scaled this threat exponentially.[181] AI-driven voice cloning can now create realistic replicas of a person's voice from seconds of audio, creating, for example, a song or a commercial that poses a threat not only to the commercial interests of celebrities but to the personal security and reputation of any individual.[182]

As this subchapter has described, the established international framework, the Berne Convention, WCT, and TRIPS are not able to address this issue. It should be noted, however, that the American-based right of publicity is not widely recognised internationally, which creates a gap in global protection. There is no international treaty for the right of publicity and a harmful act of voice cloning that could be protected in a jurisdiction in America may have no legal remedy in a country that relies purely on the Berne framework for its IP protections. For example, there is no comparable regulation in Europe. It is also important to note that in these cases, the voice has been highly distinctive and identifiable, and it is public figures who have been granted protection, while ordinary people remain completely unprotected.[183]

Additionally, it is important to note that this legislative issue has not gone unnoticed. The problem of deepfakes and AI-driven identity appropriation is forcing a global re-evaluation.

---

[179] Ward & Smith 2025.
[180] Ibid.
[181] Modern Diplomacy 2025.
[182] Prolific North, AI vs the human voice 2025.
[183] *Watt et al.,* pp. 166-167.

WIPO, the administrator of the core copyright treaties, is now at the center of this conversation. In its issues paper on AI and intellectual property, WIPO explicitly poses the question of whether rights of publicity, personality rights or other *sui generis rights* should be developed in countries that currently lack them, specifically to protect against unauthorised creations of deep fakes. WIPO is also examining complex issues, such as the blending of multiple individuals' likenesses to create a new, non-existent persona, and the legal status of randomly generated voices that may, by pure chance, resemble a real person.[184]

[184] WIPO, Artificial Intelligence: Call for Comments on Intellectual Property Policy.

4 Human rights approach to the protection of the voice

4.1 Human rights and the current technological era

As explained in the previous chapter, the protection of the human voice cannot be based on copyright law. In this chapter, divided into subchapters, the aim is to determine whether the protection of an individual's voice in the age of AI can be grounded in the broader human rights framework that defines the dignity, autonomy, and identity of every person. The issue is examined through two major human rights instruments that provide a foundation: the European Convention on Human Rights (ECHR) and the Universal Declaration of Human Rights (UDHR). While the UDHR is not a binding treaty, it remains the cornerstone of modern human-rights thinking. The UDHR articulates the universal values, and its principles have shaped the development of subsequent binding instruments, including the ECHR and national constitutional frameworks.[185] On the other hand, ECHR supplies the concrete legal standards which are interpreted through the case law of the European Court of Human Rights (ECtHR).[186] This combination of conventions allows the research's human rights analysis to connect fundamental human-rights principles with the practical jurisprudence needed to evaluate the challenges faced in modern technological era.

The ECtHR has stated that the ECHR convention is a living instrument which must be interpreted in the light of present-day conditions. That has allowed the Court to interpret the provisions of the convention in modern daylight and provide new freedoms under the existing provisions of the ECHR.[187]

The living instrument doctrine was established in Strasbourg's case law since the very early days in the judgement of *Tyrer v. United Kingdom* delivered in 1978. In the case, the ECtHR had to decide whether corporal punishment of young people was degrading punishment within the meaning of Article 3 of the ECHR.[188] The punishment, which consisted of corporal punishment on bare skin carried out by police officers at a police station, was provided for by law and practised in the Isle of Man, a territory of the United Kingdom with considerable legislative autonomy. At that time, legal corporal punishment had been abandoned elsewhere in the United Kingdom and was not in use in most other states parties to the Convention. The

---

[185] *Hallo de Wolf – Moerland*, p. 186.
[186] *Milošević*, p. 211.
[187] *van der Sloot*, p. 39.
[188] *Letsas*, p. 2.

Minister of Justice of the Isle of Man stated in his opinion that the punishment was generally accepted on the Isle of Man.[189] However, the Court rejected this argument, stating that the general acceptance of legal corporal punishment cannot be a criterion for assessing its degrading nature, because the reason why people support such punishment may well be that corporal punishment is so degrading that it acts as a deterrent. It added:

> "The Court must also recall that the Convention is a living instrument which, as the Commission rightly stressed, must be interpreted in the light of present-day conditions. In the case now before it the Court cannot but be influenced by the developments and commonly accepted standards in the penal policy of the member States of the Council of Europe in this field."[190]

The Court ruled on purely substantive grounds, stating that physical violence used by one person against another constitutes an attack on human dignity and physical integrity, which the Article of the agreement is specifically intended to protect. It added that the institutionalised nature of the punishment, the fact that it was carried out by persons completely unknown to the offender, and the fact that it was carried out on the bare buttocks made the punishment degrading. The Court therefore found that Article 3 of the Convention had been violated.[191] However, what is significant for our research is what the Court ruled on the ECHR and how it applied Article 3 to a form of punishment that was entirely legal in other places.

Shortly after the preceding case, another significant case followed, creating a solid foundation for the living instrument doctrine concerning the ECHR. In the case of *Marckx v Belgium*, the applicants' child was born out of wedlock.[192] Belgian law did not recognise the general principle of *mater semper certa est* in child law and therefore did not recognise the mother as the child's parent. The mother could only be recognised as the parent through voluntary recognition or a Court decision. The applicants considered this to be a violation of their right to family life under Article 8 of the ECHR and discrimination against them under Article 14 of the ECHR.[193] The Court ruled as follows, stating that Belgian legislation was behind the curve and the Article 8 of ECHR had been breached:

> "The domestic law of the great majority of the member States of the Council of Europe has evolved and is continuing to evolve, in company with the relevant

---

[189] *Tyrer v. United Kingdom*, para. 31.
[190] Ibid., para. 31.
[191] Ibid., para. 33.
[192] *Letsas*, p. 3.
[193] *Marckx v Belgium*, para. 31.

> international instruments, towards full juridical recognition of the maxim *mater semper certa est*."[194]

What made the case even more interesting was the significant difference between the Tyrer and Marckx cases. In the latter case, the Court specifically referred to two international conventions (The Brussels Convention on the Establishment of Maternal Affiliation of Natural Children and the European Convention on the Legal Status of Children born out of wedlock) to demonstrate that generally accepted norms exist. However, these two international conventions had not been signed by the majority of the contracting states at that time. The Court justified this by stating that "the existence of these two conventions shows that there is a clear common basis among modern societies in this area." It added that Belgian legislation itself showed signs of this development of rules and attitudes.[195] Although this case is far removed from the whirlwind of technology, the interpretation by the Court is very interesting in that it relies on certain known agreements, legal norms that were not yet universally accepted at the time. By analogy, much can be deduced from this for the subject of the research.

Also, the case of *Dudgeon v. United Kingdom*, which dealt with the criminalisation of homosexuality in Northern Ireland, further supports the living instrument doctrine in Court interpretations. Criminal penalties were found to be contrary to Article 8 of the ECHR, which guarantees the right to respect for private life.[196] In the case, the Court stated that compared to when the legislation was enacted, homosexual behavior is better understood today, which has increased tolerance towards it. This has led to a social change in which the social consensus is shifting, in this case towards tolerance, and as a result, a law that was previously acceptable has become a violation of human rights.[197] However, I would find that the logic of the Court's interpretation in this case is quite applicable to the subject of the research. In this case, better understanding led to the removal of prohibitions and decriminalisation. In the case of AI and voice, it could be argued that such a better understanding and comparing it to the time when the conventions were signed, the significance of voice for identity and the potential harmfulness of AI now requires increased protection.

---

[194] Ibid., para. 41.
[195] *Letsas*, p. 4.
[196] Ibid.
[197] Dudgeon v United Kingdom, para. 60.

In these very early cases where the Court characterised the ECHR as a living instrument, a certain pattern can be observed. Firstly, the Court is called upon to resolve an issue that is morally sensitive in the respondent state: corporal punishment, children born outside of marriage, and homosexuality. Secondly, the Court recognises the importance of the prevailing moral climate in the respondent state for the interpretation of the Convention. Thirdly, the Court refers to broader developments within the Council of Europe and the prevailing moral climate in the respondent state. Fourthly, the Court also examines a number of substantive issues that are not related to the prevailing moral climate or the common standards of the Council of Europe. Rather, they relate directly to the purpose of the protected right and why the government's actions do not fulfil that purpose in the case in question. In most cases, this formula has led to a finding of violation.[198]

Following the interpretation of the Court and developments in case law, the research topic of unauthorised use of person's voice and happened infringement could be brought to ECtHR for a ruling. Firstly, protecting the voice against technologies and their ability to both copy and infringe the voice is a sensitive issue and a genuine concern at the present. Cases related to this issue are constantly coming up. Secondly, the public climate recognises this issue and has acknowledged the challenges it poses. Thirdly, this problem is universal, but no significant developments have been made in this field, meaning that there is very little regulation on the subject. However, the problem is recognised in other regions, but there is little support from other regulatory areas to resolve it. Fourthly, in such a case, the Court might note that infringements of a person's voice would prevent that person from enjoying their potentially protective human rights. Furthermore, the living instrument doctrine is supported by the fact that, at the time the agreements were concluded, it was not possible to foresee the challenges that technology might pose to the realisation of human rights.

4.2 Voice protection under the right to private life and privacy

As stated in the chapter on the relationship between voice and copyright, it was argued that the protection of a person's voice cannot be structured on that ground. In this subchapter, the aim is to argue that the protection of a person's voice can be structured within the framework of human rights.

---

[198] *Letsas*, p. 5.

The rights that are closest to protecting human's personal identity, authenticity, and dignity are found in the following provisions of the mentioned conventions of UDHR and ECHR. Under Article 12 of the Universal Declaration of Human Rights:

> "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

As seen in the UDHR provision, it provides protection to the person's privacy, family, home or correspondence, which could be argued to mean also the personality. Additionally, the provision protects against attacks upon a person's honour and reputation, which could be argued to mean the infringement of the voice as a part of a person's personality.

Article 8 of the European Convention on Human Rights protects the right to respect for private and family life:

> "1. Everyone has the right to respect for his private and family life, his home and his correspondence.
>
> 2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others."

Again, as seen in ECHR provision's first paragraph, it provides protection to everyone's private and family life, home and correspondence, which could be argued to mean the personality. The second paragraph shows the possible limitations of the right by public authorities. However, to fully understand how these provisions protect the human voice from unauthorised use in technologies, it is necessary to look beyond the literal text and examine the interpretation of provisions and the obligations of the state.

The European Court of Human Rights has repeatedly stated that the essential object of Article 8 is to protect the individual against arbitrary interference. While the wording might suggest a negative right to be left alone, the case law has significantly expanded this into a positive dimension.[199] The Court interprets private life as a broad concept that cannot be precisely defined, covering the area of personal autonomy where individuals can freely seek to develop and fulfil their potential personality.[200] The concept of personal autonomy protection is central

---

[199] *Schabas*, p. 736.
[200] Ibid., p. 739.

to the argument for voice protection, implying that an individual has the right to establish and control the details of their identity as a human being.[201]

The Court, while interpreting the convention, has explicitly recognised that this protection extends to a person's physical and psychological integrity and additionally to their social identity. It is highlighted that this includes attributes such as a person's name and, crucially, their image.[202] The Court has viewed a person's image as one of the chief attributes of their personality because it reveals unique characteristics and distinguishes the person from their peers. Consequently, individuals are entitled to control the use of their image, including the right to refuse its reproduction by others. [203] Following the analogy, the human voice, which is arguably as unique and distinguishing as an image of a person, does that mean the face of a person or the appearance of it, could be afforded the same level of protection. Suppose the unauthorised recording and reproduction of an image violates Article 8 because it places an essential attribute of personality into the hands of a third party. In that case, the unauthorised synthesis of a voice by AI constitutes a parallel violation of Article 8 and the protected right.

Furthermore, the protection of the voice under Article 8 is strengthened by the doctrine of positive obligations. It is argued that the word "respect" in Article 8 implies that the state must do more than simply abstain from interference. That means the state must also protect individuals from infringements of their rights by other private individuals.[204] This is particularly relevant to the threats of the unauthorised use of technologies such as AI, where the infringement typically comes from private technology companies or individual users rather than the state itself. The state should therefore be required to have a sufficient regulatory framework in place to protect against these acts. When it comes to fundamental values and essential aspects of private life, such as the integrity of one's own voice, the state's margin of discretion should be very limited, and effective legal safeguards are essential. Understandably, creating such safeguards takes time, and it is impossible to prepare for every eventuality, but the current situation is demonstrably challenging, and leaves people vulnerable to infringement.

Finally, the issues of emerging technologies such as AI voice infringements often involve the creation of deepfakes or other kinds of misleading content, which touches upon the protection

---

[201] Ibid., p. 745.
[202] Ibid., p. 745. See *Aksu v. Turkey [GC]*, nos 4149/04 and 41029/04, § 58, ECHR 2012.
[203] Ibid., p. 746. See *Reklos and Davourlis v. Greece*, no. 1234/05, § 40, 15 January 2009.
[204] Ibid., p. 737.

of honour and reputation. Although Article 8 does not explicitly mention honour and reputation (unlike Article 12 of the UDHR), it is noted that the Court has recognised the right to protection of reputation as part of the right to respect for private life.[205] To fall within the scope of Article 8, the damage to a person's reputation must be of a certain level of severity[206], but, for example, the unauthorised use of a voice, even in an inappropriate song or context, which has been created using technology such as AI, could seriously affect a person's moral integrity and social standing. Thus, the unauthorised use of a voice is not merely a technical breach of security, but a severe violation of a person's privacy and reputation.

4.3 Case law of the ECtHR related to the Article 8

To support the theoretical framework and delineate the scope of rights under the conventions, the following subchapter examines relevant ECtHR case law. Specifically, to justify the argument that the human voice falls under the strict protection of Article 8 ECHR, it is necessary to analyse the Court's extensive jurisprudence regarding the right to private life and the right to one's own image.

The following analysis examines three pivotal cases: *Von Hannover v. Germany (No. 1)*[207], which delineates the scope of privacy in the public sphere; *P.G. and J.H. v. the United Kingdom*[208], which specifically addresses the recording of the human voice as the processing of personal data; and *Reklos and Davourlis v. Greece*[209], which establishes the necessity of prior consent for the capture of biometric identifiers.

The jurisprudence of the ECtHR regarding Article 8 has been significantly shaped by the landmark case of *Von Hannover v. Germany (No. 1)*. This case is particularly distinct because it establishes the extent to which the state is obligated to protect an individual’s personality rights, specifically their image, even when that individual is a public figure situated in a public space.

The litigation was initiated by Princess Caroline von Hannover, the eldest daughter of Prince Rainier III of Monaco, who sought to restrain the publication of photographs depicting her daily life. These images, published by German magazines, captured the applicant engaging in

---

[205] Ibid., p. 753. See *Chauvy and Others v. France*, no. 64915/01, § 70, ECHR 2004-VI.
[206] Ibid., p. 754.
[207] *Von Hannover v. Germany (no. 1)*, no. 59320/00, ECHR 2004-VI.
[208] *P.G. and J.H. v. the United Kingdom*, no. 44787/98, ECHR 2001-IX.
[209] *Reklos and Davourlis v. Greece*, no. 1234/05, 15 January 2009.

purely private activities such as dining, shopping, and horseback riding in locations that were technically accessible to the public.[210] The applicant argued that the unauthorised dissemination of these images infringed upon her right to the protection of her private life and her own image.[211]

The German domestic courts dismissed her claims by the reasoning that relied on the classification of the applicant as a figure of contemporary society par excellence. Under this classification, the courts held that she was obliged to tolerate the publication of photographs taken outside her secluded private residence without her consent.[212] Essentially, the domestic courts posited that her right to privacy was forfeited once she entered the public sphere.[213]

In this case, the ECtHR ruled in favour of the applicant, finding a violation of Article 8. A critical aspect of this judgment was the Court's clarification of the state's role. The case did not involve a direct interference by the state, but rather a failure to protect the applicant from private media entities. The Court clarified that the purpose of Article 8 is not limited to a negative obligation requiring the state to refrain from arbitrary interference but also covers positive obligations relating to effective respect for private life. This requires the state to adopt provisions designed to secure respect for private life even in the sphere of relations between individuals.[214]

In defining the scope of this protection, the Court considered that the concept of private life includes elements of personal identity, such as a person's name, image, and physical and psychological integrity. Crucially, the Court stated that the protection of Article 8 is intended to ensure the unrestricted development of the personality in relationships with others. Therefore, the Court famously stated that there is "a zone of interaction of a person with others, even in a public context, which may fall within the scope of private life".[215]

The core of the Court's decision was based on finding a fair balance between the right to privacy and the freedom of expression (Article 10). The Court drew a fundamental distinction between reporting on facts that may contribute to a debate of general interest in a democratic society, such as the conduct of politicians or public officials and reporting that merely

---

[210] *Von Hannover v. Germany*, 2004, § 11-18.
[211] Ibid., § 3.
[212] Ibid., § 72.
[213] Ibid., § 53.
[214] Ibid., § 57.
[215] Ibid., § 50.

concerns the private details of an individual who does not hold public office. The Court found that the photographs served only to satisfy the curiosity of a specific readership and did not contribute to public debate.[216] The Court reasoned that the public does not possess a legitimate interest in knowing the whereabouts or behavioural details of the applicant's private life, even if she is a well-known figure. Consequently, the commercial interests of the magazines and the curiosity of the public had to give way to the applicant's right to effective protection. The Court found that the criteria established by the national courts were too vague to protect the rights of individuals and confirmed that every person, regardless of their public status, has a legitimate expectation that their private life will be protected.[217]

While Von Hannover established the boundaries of privacy regarding a person's image and voice by following the analogy in the case, the case of *P.G. and J.H. v. the United Kingdom* is important for the research, as it deals with the issue of the legal status of the human voice. The case involved two applicants who were arrested on suspicion of planning a robbery. During their detention at the police station, they refused to answer questions or provide voice samples for forensic comparison. Following that, the police covertly installed listening devices in their cells and in the custody area to record their conversations. These recordings were then analysed by voice experts to compare the applicants' speech patterns with incriminating voice recordings obtained from the crime scene.[218]

The applicants held that the covert recording of their voices while in police custody violated their right to respect for private life. The government argued that the applicants had no reasonable expectation of privacy within a police cell, a location where they were under constant supervision. In the case the legal issue was not merely the listening to the conversations, but the creation of a permanent record of the voice for the specific purpose of forensic analysis.[219]

In this case, the ECtHR found a violation of Article 8 and its reasoning, the Court made a critical distinction between the immediate monitoring of a detainee's actions and the permanent recording of their voice, interpreted as data. The Court stated that while a person in a police cell might expect to be observed or listened to for security reasons, this does not

---

[216] Ibid., § 65.
[217] Ibid., § 78.
[218] *P.G. and J.H. v. the United Kingdom*, 2001, § 9-15.
[219] Ibid., § 35.

automatically permit the recording of their voice.[220] Additionally, the domestic law of the United Kingdom did not provide a clear statutory framework regulating the use of such covert listening devices for this purpose and thus the police actions were not in accordance with the law.[221]

The importance of this judgment for the protection of voice against emerging technologies such as AI is the Court's classification of the voice itself. The Court ruled that the recording of the applicants' voices had to be regarded as concerning the processing of personal data. The Court emphasised that even if the information, in this case the information that was spoken, is not inherently private in the sense of containing secrets, the systematic collection and storage of that data triggers Article 8 protection. The Court stated that "the recording of such data and the systematic or permanent nature of the record may give rise to such issues".[222]

The legal principles supported in P.G. and J.H., firstly, confirm that the human voice is legally recognised and classified as personal data. Which means that it is something individual, associated with a person, and can even be used to identify somebody.

Secondly, the distinction between listening and recording the voice of a person is directly applicable to emerging technologies. In this case the police were entitled to listen the voice but not to record the voice for identification without a legal basis. For example, in the AI technology, the system or its operator might hear a publicly available voice, but creating a permanent record of a voice violates the rights protected by Article 8.

Thirdly, the Court emphasised in this case that police actions must be strictly in accordance with the law. On the other hand, the police is an authority, so this demand is strongly underlined. It must still be noted that, at the moment the use of recorded voice in emerging technologies, for example in AI, is not strictly regulated, which causes the issue that such activities cannot be explicitly in accordance with the law.

While P.G. and J.H. established the voice as personal data, the case of *Reklos and Davourlis v. Greece* addresses the upstream issue of capturing biometric data without consent.

---

[220] Ibid., § 57.
[221] Ibid., § 63.
[222] Ibid., § 57.

The applicants were the parents of a newborn baby who had been placed in a sterile unit at a private clinic immediately after birth. A professional photographer entered the restricted unit and took close-up photographs of the infant's face without the parents' prior knowledge or consent. When the parents later demanded that the photographs be handed over, the clinic and the photographer refused. The applicants sought damages in the Greek civil courts, claiming a violation of their child's personality rights without success. The courts reasoned that the infant's personality had not yet developed sufficiently to suffer damage and that the photographs were not offensive or derogatory in nature. Domestic ruling based on that since the images were not published or used to degrade the child, no harm had occurred.[223]

The ECtHR disagreed with the domestic courts' ruling and found a violation of Article 8. The Court emphasised that a person's image constitutes one of the chief attributes of their personality, as it reveals the person's unique characteristics and distinguishes them from their peers. The Court stated that the violation of privacy happens at the moment of recording, regardless of the publication of the photographs later or not. The Court stated that "the effective protection of the right to control one's image presupposes obtaining the consent of the person concerned at the time the picture is taken and not just when it comes to possible publication". [224]

The Court rejected the argument that the retention of the images was harmless. It held that by taking the photographs, the photographer retained the ability to use the image for subsequent purposes against the wishes of the person concerned. This situation resulted in an essential attribute of personality being retained in the hands of a third party, which the Court deemed a violation of Article 8 rights.[225]

This ruling also provides a significant basis for the topic of this research. Firstly, the ruling confirms that a violation of Article 8 occurs at the time of recording and does not even require the publication of the work. This creates an interesting basis for, for example, training AI with voice samples, in which case damage is already caused when the training takes place, samples are fed to the AI, and various works are created, even if they are not published. Reklos effectively disproves the claim of harmless training by arguing that legal damage occurs at the moment the personal data is collected.

---

[223] *Reklos and Davourlis v. Greece*, 2009, § 8-14.
[224] Ibid., § 40.
[225] Ibid., § 42.

Second, and perhaps most importantly in current discussions, the Court's requirement that effective protection requires consent at the time the image is taken is directly opposed to the opt-out models favoured by AI developers. Reklos points out that a system in which the voice is collected first and the subject must later request its removal, is fundamentally in conflict with Article 8. This would therefore require a strict opt-in mechanism to confirm the recording of personality traits, to which willing individuals could give their consent for the use of their characteristics.

Finally, just as the storage of photographs by photographers created a permanent opportunity for misuse, the storage of voice data without consent leaves an essential feature of personality in the hands of third parties. This creates a hidden but permanent threat to personal autonomy, which, according to the convention, must be prevented right from the start.

4.4 The conflict of rights: AI creativity and the right to private life and privacy

While the preceding chapter established the human voice's position as a subject of right to respect for private and family life protection. However, this right is not absolute. In the digital environment, those rights immediately clash with the competing fundamental right of freedom of expression, which is enjoyed by media and developers creating content with emerging technologies, for example with generative AI. The right to freedom of expression defines the scope of the protection provided by the right to privacy. For example the song made with the AI system by using the voice samples without permission, whether for a musical parody, or a satirical video, is frequently defended as a form of creative or artistic speech. Consequently, the legal analysis must navigate the conflict between the individual’s right to control their personal identity and the creator’s right to perform the ideas and artistic expressions.

The right to freedom of expression is stated in Article 19 of the Universal Declaration of Human Rights:

> "Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers."

A similar right is stipulated in Article 10 of the European Convention of Human Rights:

> "1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not

> prevent States from requiring the licensing of broadcasting, television or cinema enterprises.
>
> 2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary."

Article 10 of the European Convention on Human Rights protects freedom of expression, which is described by the ECtHR as one of the essential foundations of a democratic society and one of the basic conditions for its progress and for each individual's self-fulfillment. As noted in the commentary by Schabas, the right is not limited to information that is favorably received, as it extends to ideas that could offend, shock, or disturb. [226]

Article 10 guarantees freedom of expression to everyone. It applies to both natural and legal persons and has frequently been invoked by the latter, notably by publishing houses and media outlets. No distinction is made with respect to the role played by those who exercise freedom of expression, or whether the aim pursued is to make a profit.[227]

Although Article 10 was drafted in a pre-digital era, its contemporary application is significant from the rise of the internet. As the Court has noted, the internet has become a central platform for the performance of freedom of expression, significantly lowering barriers to participation and enabling individuals to spread and access information with unprecedented ease and speed. This development has effectively expanded the practical reach of Article 10 in modern society.[228] However, those attributes also intensify risks to other Convention rights, particularly those protected under Article 8. The Court has acknowledged that the potential harm caused by online content may be greater than that associated with the traditional press. That has opened the discussions for justifying the distinct regulatory approaches related to the issue.[229]

---

[226] *Schabas*, p. 859. See *Palomo Sánchez and Others v. Spain [GC]*, nos 28955/06, 28957/06, 28959/06, and 28964/06, § 53, ECHR 2011; *Lingens v. Austria*, 8 July 1986, § 41, Series A no. 103.
[227] Ibid., p. 863.
[228] Ibid., p. 870.
[229] Ibid., p. 871.

4.5 Case law of the ECtHR related to the Article 10

In the following subchapter, the case law of the ECtHR will be presented to highlight how the right to freedom of expression has been approached in relation to other legal rights, especially related to the research topic. The cases that are examined are *Von Hannover v. Germany (No. 2)*[230], *Bohlen v. Germany*[231] and *Sousa Goucha v. Portugal*[232] to evaluate the balance between the rights.

While the first Von Hannover judgment defined the scope of privacy, the important judgment in *Von Hannover v. Germany (No. 2)* is a significant decision on how the conflicting rights of Article 8 regarding privacy and Article 10 regarding freedom of expression could be balanced. The case establishes a structured test to determine when the public's right to receive information overrides an individual's right to control their personal data.

The applicants, Princess Caroline of Monaco and her husband Prince Ernst August von Hannover, sought an injunction against the publication of photographs taken of them during a skiing holiday. Unlike the photos in the first case, which were purely taken in free time, these images were published in an article reporting on the Prince's poor health during the vacation. The applicants argued that the publication violated their Article 8 rights, while the publishing companies claimed that the publication contributed to a debate of general interest regarding the health of a public figure. The German Federal Court of Justice had refused to grant an injunction for these specific photos, arguing that the article's text regarding the Prince's illness constituted a matter of public concern.[233]

The ECtHR found no violation of Article 8, ruling that the German courts had struck a fair balance between the competing interests. In the ruling, the Court applied the case law into five specific criteria that had to be applied when balancing Article 8 and Article 10:

1) the contribution to a debate of general interest, noting that whether a matter is of public interest depends on the specific circumstances;

2) the degree of public recognition of the person concerned and the subject of the report, meaning that public figures must tolerate a higher level of scrutiny;

---

[230] *Von Hannover v. Germany (no. 2)*, nos. 40660/08 and 60641/08, ECHR 2012.
[231] *Bohlen v. Germany*, no. 53495/09, ECHR 2015.
[232] *Sousa Goucha v. Portugal*, no. 70434/12, ECHR 2016.
[233] *Von Hannover v. Germany (No. 2)*, 2008, § 30-35.

3) the prior conduct of the person concerned, including whether they have themselves exposed aspects of their private life;

4) the content, form and consequences of the publication, meaning the manner and context in which the information is presented and disseminated; and

5) the circumstances in which the photographs were taken, particularly whether the person consented or whether the material was obtained through intrusion or harassment.[234]

Applying these criteria, the Court reasoned that the Prince's illness was a matter of general interest and that the photos, when combined with the article, contributed to the public discussion on this topic. Consequently, the freedom of the press outweighed the applicants' right to privacy in this specific instance.[235]

The principle of general interest, as it is described in the test in *Von Hannover v. Germany (No. 2)*, could serve as the primary legal filter for separating the harmful and the permissible use of a person's voice in AI. When applying the first criterion of the test, an unauthorised AI voice use can likely only claim Article 10 protection if it contributes to a debate of general interest. However, the vast majority of commercial AI uses, such as AI cover songs, hardly fulfil this threshold. These applications serve primarily to entertain or defraud, rather than to inform the public. As the Court noted in the case, when the publication serves merely to satisfy the curiosity of a particular readership regarding the details of private life, freedom of expression calls for a narrower interpretation.[236] Therefore, under the *Von Hannover v. Germany (No. 2)* framework, functional or entertainment-based AI cloning that lacks a clear public interest contribution should yield to the Article 8 rights of the voice owner.

In the case *Bohlen v. Germany*, the applicant, Dieter Bohlen, was a prominent German musician and television personality, known for his polarised public image and commercially successful but criticised autobiography. A tobacco company launched a large-scale advertising campaign that used the applicant's first name and a play on words referring to his book without his consent. The advertisement did not use his image but capitalised on his recognition for promoting cigarettes. The applicant sought an injunction and damages,

---

[234] Ibid., § 108-113.
[235] Ibid., § 118-124.
[236] Ibid., § 110.

arguing that the unauthorised use of his name for commercial purposes violated his Article 8 rights. The German Federal Court of Justice dismissed his claim, reasoning that the advertisement was a satirical comment on a public figure and that the applicant's personality right had to yield to the tobacco company's freedom of expression.[237]

The ECtHR dismissed the application, finding no violation of Article 8. In its reasoning, the Court had to determine whether the use of the applicant's identity was purely commercial or if it contained elements of protected expression. The Court acknowledged that the advertisement was indeed commercial in nature, intended to sell products.[238] However, the Court emphasised that the advertisement also contained also a communicative element, namely a satirical comment on the quality of the applicant's book and his public character.[239]

The Court reasoned that the applicant, who had actively sought media attention through provocative behavior, had to display a higher degree of tolerance towards public criticism and satire.[240] Crucially, the Court noted that the domestic courts had correctly classified the advertisement as satire rather than a mere appropriation of the name. Because the ad contained additionally a critical, communicative message about a public figure, it was entitled to a degree of protection under Article 10, even though it appeared in a commercial context.[241] The Court concluded that the domestic authorities had found a fair balance, as the use of the name did not suggest the applicant had endorsed the product. Instead, the reference to his name was clearly intended as a joke at his expense, not as a commercial use of his personal identity.[242]

The reasoning in *Bohlen v. Germany* provides an interesting viewpoint for regulating the use of voice in AI and even the commercial aspect of it. Firstly, the case establishes that commercial use does not automatically negate Article 10 protection if there is a communicative or satirical element involved. This could suggest that, for example, an AI-generated parody song, even if monetised on a platform like YouTube or Spotify, could be protected if it serves as a satirical element.

---

[237] *Bohlen v. Germany*, 2015, § 6-13.
[238] Ibid., § 48.
[239] Ibid., § 50.
[240] Ibid., § 54.
[241] Ibid., § 52.
[242] Ibid., § 56.

However, the *Bohlen v. Germany* judgment also implies a limit. The Court protected the advertisement specifically because it was about the applicant and it was a joke at his expense. By contrast, most of the unauthorised AI voice-using cases are lacking this communicative element. Such infringements usually do not make a satirical statement about the owner of the voice, and they are most often abusing the value of the voice to improve their own intentions.

While *Von Hannover v. Germany* and *Bohlen v. Germany* addressed the possible unauthorised use of images and names, the case of *Sousa Goucha v. Portugal* concerns the imitation of a person's voice, gestures, and identity for parody purposes. The judgment is important in defining the boundary between protected artistic satire and the violation of a person's dignity.

The applicant, Manuel Luís Goucha, was a well-known television host in Portugal. He complained about a comedy television show that featured a weekly sketch parodying him. In these sketches, an actor dressed in women's clothing imitated the applicant's voice, gestures, and mannerisms. The applicant argued that the imitation was degrading, attacked his honour, and mocked his sexual orientation and thereby violated his right to private life under Article 8. The Portuguese domestic courts dismissed his claim for damages, reasoning that the sketches fell within the scope of permissible criticism and satire regarding a public figure.[243]

The ECtHR dismissed the application, finding no violation of Article 8. In its reasoning, the Court emphasised that satire is a form of artistic expression and social commentary which, by its inherent features of exaggeration and distortion of reality, naturally aims to provoke and agitate.[244]

In the case, the Court applied a specific test regarding the likelihood of confusion. It observed that the show was explicitly comedic and that the imitation was a caricature. Consequently, a reasonable observer would not have mistaken the actor's performance for the real applicant, nor would they interpret the satirical portrayal as a factual statement about the applicant's life.[245] Furthermore, the Court repeated that as a public figure, the applicant was expected to display a greater degree of tolerance towards artistic criticism and parody.[246] The Court

[243] *Sousa Goucha v. Portugal*, 2016, § 6-14.
[244] Ibid., § 49.
[245] Ibid., § 55.
[246] Ibid., § 50.

concluded that the domestic courts' assessment that the sketches were not personal attacks but humorous expressions was not unreasonable and did not infringe the applicant's rights.[247]

Therefore, the legal principles established in Sousa Goucha provide protection on behalf of certain forms of AI usage, but also arguments against it. Firstly, the judgment confirms that AI-generated parody could be protected under Article 10, if it is provided as recognisable satire. Created voice output could get protection if it is clearly presented and labeled as a joke and a reasonable observer would not confuse it with reality.[248] The reasonable observer is specified in ECtHR's case law as a hypothetical, reasonably critical member of the public who understands the context, is able to distinguish between normal means of expression, and does not interpret the message in a way that cannot reasonably be imposed on the exercise of freedom of expression.[249] On the other hand, the unreasonable observer would be the theoretical opposite of the reasonably critical member of the public. Such an observer would be characterised by a hyper-sensitive or literalist interpretation of a message, failing to distinguish between irony, satire, or common metaphors, and ignoring the broader societal or political context in which the expression occurs. By rejecting the unreasonable observer, the boundaries of acceptable discourse are not dictated by the most easily offended or irrational interpretations, but by a balanced and context-aware public understanding.

However, this case highlights the specific issue with the false material created by emerging technologies, such as AI. The Court's tolerance for the imitation in *Sousa Goucha v. Portugal* relied heavily on the fact that it was an obvious performance by an actor, which means that there was no risk of the public believing the applicant actually said those words. On the other hand, emerging modern technologies and AI aim for hyper-realism that often removes the element of caricature and makes the content, in this research voice, hardly distinguishable from the original. Thus, it can be argued that because high-fidelity and realism aiming AI voice cloning easily eliminates the distinction from reality that the Court relied upon in Sousa Goucha, it forfeits the protection of satire. If a reasonable observer cannot recognise the AI voice from the original real human voice, the risk of reputational damage is far higher than in a comedy sketch, shifting the balance back in favor of Article 8 protection.

---

[247] Ibid., § 56.
[248] Ibid., § 55.
[249] See *Nikowitz and Verlagsgruppe News GmbH v. Austria*, no. 5266/03, ECHR 2007 and *Vereinigung Bildender Künstler v. Austria*, no. 68354/01, ECHR 2007.

4.6 Right to property and a voice as a possession

While the preceding analysis has focused on the protection of the voice as a matter of dignity and moral integrity, the unauthorised use of a person's voice in AI has additionally has a significant economic aspect. For many individuals, particularly professional singers, voice actors, and public figures, the voice is not only a component of their personality but also the primary instrument of their livelihood. Consequently, it is necessary to examine the protection of the voice through the lenses of international conventions that guarantee the right to property. Due to the limited scope of the research, the discussion of the subject area has been presented by bringing up possible approaches to interpretation, but a broader coverage of the comparison of possible personality traits to property would require more extensive research.

Article 17 of the Universal Declaration of Human Rights states the following:

> "1. Everyone has the right to own property alone as well as in association with others.
>
> 2. No one shall be arbitrarily deprived of his property."

European Convention on Human Rights in Article 1 of Protocol offers the following protection:

> "Every natural or legal person is entitled to the peaceful enjoyment of his possessions. No one shall be deprived of his possessions except in the public interest and subject to the conditions provided for by law and by the general principles of international law.
>
> The preceding provisions shall not, however, in any way impair the right of a State to enforce such laws as it deems necessary to control the use of property in accordance with the general interest or to secure the payment of taxes or other contributions or penalties."

Article 1 of Protocol No. 1 states that every natural or legal person is entitled to the peaceful enjoyment of his possessions and shall not be deprived of his possessions except of the state's right to enforce laws as it deems necessary to control the use of property in accordance with the general interest. In the context of the research, it is remarkable to investigate whether the voice, or the rights attached to it, could be considered as an enjoyable possession. The possession has been seen as a property, asset that at least has at least a legitimate expectation of obtaining effective enjoyment of a property right.[250] However, the possessions have an

[250] *Schabas*, p. 1726.

autonomous meaning in the ECtHR's jurisprudence that is not limited to the ownership of physical goods. The Court has clarified that certain other rights and interests constituting assets can also be regarded as property rights, and thus as possessions.[251] Respecting the non-physical assets, the Court has considered whether the legal position in question gave rise to financial rights and interests and thus had an economic value.[252] *Van Rijn* has additionally presented the interpretation that the right or interest can only be regarded as a possession if it is sufficiently established to be enforceable. If the interference with a property right is brought up, the person complaining has to show that such a right exists.[253] While arguing that the voice might has an economic value, it is essential to show that. For example, a professional singer or voice actor has a legitimate expectation that their specific vocal timbre and style could be seen as a professional asset that generates income. Such examples are songs, records, and performances, as well as the intellectual property rights associated with them. For the professionals, it might be easier to show the economic value and the enforceable right than for ordinary people.

The protection of property is strengthened by the state's positive obligations, which duty the state to secure to everyone the rights in the Convention. When an interference with the peaceful enjoyment of possessions is perpetrated, even in disputes between private individuals, the obligations arise for the state to ensure in its domestic legal system that property rights have been sufficiently protected by law.[254] Currently, the lack of regulations regarding the AI use make individuals vulnerable to infringements. Without the relevant laws, the state fails to secure a fair balance between the technological demands and the protection of the individual's proprietary interests.

---

[251] Ibid., p. 1727.
[252] Ibid.
[253] *van Rijn*, p. 869.
[254] *Schabas*, p. 1728.

5 Conclusions

This thesis aimed to answer the research question: "To what extent can copyright law and human rights frameworks effectively protect individuals from the unauthorised commercial use of their voice in AI-generated content?"

To provide the answer to the question, at first, the meaning of the voice, AI and technologies to create the AI-generated content were explored. Secondly, it was examined whether the voice would be able to get protection under copyright and copyright law. Thirdly, the human rights framework was used to see if the voice could apply as a part of the human rights and therefore get the human right protection.

At the beginning of the research, I thought it would be clear that a voice, being such a fundamental attribute of identity, would provide sufficient protection against emerging technologies and the use of voice-cloning AI and therefore the unauthorised use of a voice would be clearly prohibited. However, the findings of the research present a concerning reality. There seems to be a fundamental weakness in current international legal frameworks regarding the protection of personality rights. The research indicates that while international laws are well-designed to regulate markets, protect consumers, and facilitate fair competition, they are structurally powerless to safeguard the individual's personality and the parts of it during the digital age. Additionally, the concern arising from the research is the challenges in the protection afforded to public figures and ordinary citizens. While celebrities may refer to the economic value of their voice and other personal characteristics to claim protection under the human rights protection or possible rights of publicity doctrine, ordinary people remain vulnerable to infringements. As AI technologies can clone the voices of ordinary people, they face real threats to their privacy and autonomy, without the enforceable rights or legal standing to claim the protection.

Firstly, the analysis of the international copyright framework, including the Berne Convention, the WIPO Copyright Treaty, and the TRIPS Agreement, leads to the conclusion that copyright is an ineffective tool for protecting the human voice itself. The limitation of copyright protection requires fixation of the work that exists. As analysed in the research, the voice is an attribute of the person, not a copyrightable work.

Secondly, compared to copyright regulation, the human rights framework provided a better outcome. ECHR and ECtHR's case law has established the living instrument doctrine that

allows the broader interpretation of rights that are related to the new technological threats. Case law related to voice as a part of Article 8's right to private life and privacy showed that the voice is legally classified as personal data. Its systematic processing triggers protection regardless of the content of the speech as seen in the *P.G. and J.H. v. the United Kingdom*. Under *Reklos and Davourlis v. Greece*, the legal infringement of the person's image, including a voice potentially, occurs at the moment of capturing it without permission, not merely at the moment of publication. This supports the conclusion that opt-out models for AI training are insufficient and a strict opt-in consent mechanism is required to satisfy human rights standards.

The conflict between AI creativity and personality rights was tried to be resolved through the general interest test established in *Von Hannover (No. 2)*, and the research found a conclusion that most of AI voice cloning serves commercial or entertainment purposes, which are not a part of the debate regarding the general interest. Therefore, privacy rights should protect voice in those cases. Additionally, in the case of *Sousa Goucha*, which protects satire, the protection relies on the audience recognising the content as a caricature. AI cloned voice content, which aims for hyper-realism and possible confusion and misleading, forfeits this satire protection. If a reasonable observer cannot distinguish the AI from the human, the satire exemption fails.

Thirdly, voice as a possible possession and property under Article 1 of Protocol No. 1, presented the finding that suggests that voice, especially for professionals, could constitute an asset with a legitimate economic expectation of generating income. When AI systems challenge this asset to create competing outputs, they may interfere with the peaceful enjoyment of this possession. However, the chapter presented possible outcomes in the scope of the thesis while it left room for more detailed research in the future.

Finally, based on the findings of the research, it is safe to say that the current situation poses challenges for the protection of human personality and its specific aspects, such as voice. At the end of the research, it is proposed that stricter international binding rules should be established for the use of AI. Such rules could be a special *sui generis* regulation that would protect individuals' personal characteristics and establish an opt-in-based system for the use of other individuals' characteristics. Unauthorised use would result in compensation for the infringed. Sounds good on paper, especially when the issues are relatively new and attitudes towards them vary at the international level.