IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF ILLINOIS
EASTERN DIVISION

ALISON FLOWERS, PHILIP ROGERS, CAROL MARIN, ROBIN AMER, LINDSEY DORCUS, YOHANCE LACOUR, and VICTORIA NASSIF, each individually and on behalf of all others similarly situated, Plaintiffs, v. MICROSOFT CORPORATION, a Washington corporation, Defendant.	)	Case No. **DEMAND FOR JURY TRIAL**

COMPLAINT

Plaintiffs, Alison Flowers, Philip Rogers, Carol Marin, Robin Amer, Lindsey Dorcus, Yohance Lacour, and Victoria Nassif, individually and on behalf of all others similarly situated, by their attorneys Loevy & Loevy, and for their complaint against Defendant Microsoft Corporation ("Microsoft"), allege as follows:

NATURE OF THE CASE

1. Microsoft Corporation built a suite of commercial voice AI products using the voices of real people, including some of the most accomplished broadcast journalists, investigative podcasters, audiobook narrators, and voice actors working in the United States today. To build the foundational voice

synthesis models that now power Microsoft's Azure AI Speech, Microsoft Foundry, Microsoft Copilot, and Dynamics 365 Contact Center voice offerings — and the speech assets Microsoft acquired through its $19.7 billion acquisition of Nuance Communications — Microsoft ingested hundreds of thousands of hours of human speech and extracted the unique biometric signatures, the voiceprints, of the speakers from those recordings. Plaintiffs' voices are among them. The full inventory of Microsoft's voice products and underlying foundational models at issue is at ¶ 31.

2. Plaintiffs are seven Illinois residents whose recorded voices are among the most distinguished in their fields. They include Carol Marin, a five-decade broadcast journalist, a three-time Peabody winner, and a recipient of the 2025 Order of Lincoln, Illinois's highest civilian honor; Yohance Lacour, whose investigative podcast *You Didn't See Nothin'* won the 2024 Pulitzer Prize for Audio Reporting; Alison Flowers, a 2021 Pulitzer finalist for the investigative podcast *Somebody*; Robin Amer, a three-time duPont–Columbia honoree and creator of USA Today's *The City*; Philip Rogers, an Emmy- and Murrow Award–winning broadcaster who covered four decades of Chicago news; Lindsey Dorcus, a SOVAS and Independent Audiobook Award–winning narrator of more than two hundred audiobooks for the major American publishers; and Victoria Nassif, a first-generation Lebanese-Palestinian American audiobook narrator and actor whose Arabic-accented narrations of works by Arab and Palestinian American authors have been commercially released by Penguin Random House, Hachette, and Simon & Schuster. None of them was told that their

voice was being used to train Microsoft's commercial voice AI. None of them was asked. None of them consented.

3. A voiceprint is a digital fingerprint of the human voice. It is a mathematical representation of the acoustic features — pitch, timbre, resonance — that arise from a person's distinctive physiology, combined with the speech patterns developed over a lifetime: accent, cadence, articulation. Like a fingerprint, a voiceprint identifies the individual and cannot be changed. A Social Security number can be reissued. A credit card can be canceled. A person whose voiceprint has been taken cannot recover it by altering their voice — the biological and behavioral patterns that produced the voiceprint are the same ones used to speak every day.

4. The Illinois General Assembly enacted the Biometric Information Privacy Act, 740 ILCS 14/1 *et seq.* ("BIPA"), to address this very danger. BIPA recognizes that biometric identifiers, expressly including voiceprints, are "biologically unique to the individual" and that, once compromised, "the individual has no recourse." 740 ILCS 14/5(c). Before any private entity may collect a voiceprint, BIPA requires written notice of the specific purpose and duration of collection, along with a written release. 740 ILCS 14/15(b). Microsoft failed to comply with any of those requirements with respect to the Plaintiffs.

5. Microsoft's noncompliance was not a misreading of the statute. Microsoft's product documentation, published research, and corporate disclosures describe its voice synthesis pipeline in biometric terms,

acknowledging that its systems "process biometric voice signatures," form "unique voice signatures" from extracted features, and process "Biometric Data" within the meaning of applicable data protection law, as alleged at ¶¶ 33-41. Microsoft built the pipeline, published the research, and shipped its products on that very capability.

6. Microsoft's noncompliance was, instead, selective. Microsoft built a consent infrastructure for the speakers with whom it had a commercial relationship and built nothing for the speakers it did not. For its customer-facing voice features, Microsoft built Limited Access registration, voice-talent acknowledgment workflows, speaker-verification checks, and contractual allocations of biometric-compliance duties to its customers. For its most powerful research models, Microsoft elected to keep the technology research-only with recommended speaker-approval protocols. Each of those measures, set forth at ¶¶ 42-48, reflects Microsoft's recognition that voice-AI products require speaker consent. Microsoft built each safeguard for the downstream uses of its products. Microsoft built none of them for the upstream voices, the recordings of Plaintiffs and Class members that Microsoft itself ingested to build the foundational voice models on which every Microsoft voice product depends.

7. Microsoft's noncompliance was a deliberate institutional decision. The scale of the unconsented corpus is documented in Microsoft's published research, cited at ¶ 50, and reflects training corpora measured in hundreds of thousands of hours and comprising the recordings of hundreds of thousands of

distinct speakers. Compliance with BIPA would have required Microsoft to identify the source speakers, provide written notice, and obtain a written release before ingesting each recording. That compliance burden would have constrained the speed and scale of Microsoft's voice AI development and the $281.7 billion in annual revenue and $75-billion-plus Azure cloud business that depend on it. Microsoft chose speed and scale over compliance. That was not an oversight. It was a business decision.

8. The voiceprints Microsoft extracted from Plaintiffs are not stored in a database that can be queried, redacted, and deleted on request. They are encoded in the parameters of Microsoft's commercial voice models and reproduced in the audio that those models generate. Those models, in turn, are deployed at global scale: through Azure AI Speech and the Microsoft Foundry Model Catalog to enterprise customers worldwide; through Microsoft Copilot to hundreds of millions of Microsoft 365 users; through Microsoft Teams, which is used by a substantial majority of the Fortune 500 and which has hundreds of millions of monthly active users globally; through Dynamics 365 Contact Center deployments; through Microsoft Copilot Daily and Copilot Podcasts experiences powered by MAI-Voice-1; and through the Voice Live API for real-time voice agents. Microsoft cannot delete Plaintiffs' voiceprints from these systems without destroying or retraining the products themselves. The biometric data and the product are, at this point, the same thing.

9. The technology Microsoft built using Plaintiffs' voices now competes with Plaintiffs in the markets where they earn their living. Microsoft's

commercial voice products generate professional-quality narration, broadcast voice work, podcast hosting, and contact-center voice services at per-character and per-minute rates orders of magnitude lower than the cost of the human performers whose recordings built those products. The market substitution is not coincidental. Each of Microsoft's voice products, as alleged at ¶¶ 75-82, was built using the vocal characteristics of the human performers it now displaces, including, on information and belief, the vocal characteristics of every Plaintiff in this case.

10. Plaintiffs' injuries are concrete and particularized. Microsoft extracted Plaintiffs' voiceprints without notice or consent, depriving them of the right BIPA guarantees to make an informed decision about the collection and use of their biometric data. Microsoft retains those voiceprints in its commercial models and continues to profit from them. Microsoft has further disseminated those voiceprints, encoded in model parameters, across its global Azure infrastructure, through its acquired Nuance speech services business, and to third-party affiliates, vendors, subprocessors, and service providers that operate Microsoft's voice products at scale. The voiceprints cannot be recovered or replaced. The technology built on those voiceprints now displaces Plaintiffs in the markets where they earn their living.

11. Plaintiffs bring this action under BIPA, 740 ILCS 14/15(a)–(e), alleging that Microsoft unlawfully collected, retained, commercialized, and disseminated their voiceprints, failed to protect them from disclosure, and did so without notice, informed written consent, a written release, or any publicly

available retention and destruction policy applicable to non-users. Plaintiffs also assert that Microsoft's commercial use of their voices and identities to build and sell AI products that mimic them violates the Illinois Right of Publicity Act ("IRPA"), 765 ILCS 1075/1 et seq., including its prohibitions on commercial use of identity and on the knowing distribution of unauthorized digital replicas. Plaintiffs further assert claims under the Illinois Consumer Fraud and Deceptive Business Practices Act ("ICFA"), the Illinois Uniform Deceptive Trade Practices Act ("IUDTPA"), and the common law of unjust enrichment.

12. Plaintiffs seek (i) statutory damages under 740 ILCS 14/20 for violations of BIPA's notice, consent, retention, profiting, dissemination, and protection requirements, computed on a per-person, per-subsection basis consistent with the statute as amended and as construed in *Clay v. Union Pacific Railroad Co.*, No. 25-2185 (7th Cir. Apr. 1, 2026); (ii) actual damages and disgorgement of profits Microsoft has earned from the commercial exploitation of Plaintiffs' biometric data; (iii) injunctive relief requiring Microsoft to (a) cease collecting biometric identifiers from voice recordings produced or recorded in Illinois without BIPA-compliant consent, (b) identify the sources of the voice training data used to build its foundational voice synthesis models, (c) destroy all voiceprints and biometric information unlawfully obtained from Plaintiffs and the Class, and (d) destroy or retrain, without the unlawfully obtained biometric data, the foundational voice synthesis models, and the downstream commercial products built on those models, in which the

unlawfully obtained biometric data are encoded; and (iv) reasonable attorneys' fees, costs, and expenses.

PARTIES

13. Plaintiff Alison Flowers ("Flowers") is a citizen of Illinois and resides in this District. Flowers is an investigative journalist and audio producer who produces her audio reporting from Chicago through her production company Spiralbound and was a 2021 Pulitzer finalist for the investigative podcast *Somebody*. Flowers's body of professional voice work, the public availability of her recordings, and the basis for Plaintiffs' allegation that Microsoft extracted her voiceprint are described at ¶¶ 83-88.

14. Plaintiff Philip Rogers ("Rogers") is a citizen of Illinois and resides in this District. Rogers is a broadcast journalist whose four-decade career was conducted in and from Chicago, primarily at WBBM Newsradio (CBS) and WMAQ-TV (NBC 5 Chicago). Rogers's body of professional voice work, the public availability of his recordings, and the basis for Plaintiffs' allegation that Microsoft extracted his voiceprint are described at ¶¶ 83-85, 89-91.

15. Plaintiff Carol Marin ("Marin") is a citizen of Illinois and resides in this District. Marin is a five-decade broadcast journalist whose career has been conducted primarily in Chicago, who has won three Peabody Awards (including a Personal Peabody), and who received the Order of Lincoln — Illinois's highest civilian honor — in 2025. Marin's body of professional voice work, the public availability of her recordings, and the basis for Plaintiffs' allegation that Microsoft extracted her voiceprint are described at ¶¶ 83-85, 92-94.

16. Plaintiff Robin Amer ("Amer") is a citizen of Illinois and resides in this District. Amer is a journalist, podcast creator, audio producer, and on-air host whose work has been produced substantially in and from Chicago, including as creator and host of USA Today's *The City* and as Managing Editor of *Love + Radio*. Amer's body of professional voice work, the public availability of her recordings, and the basis for Plaintiffs' allegation that Microsoft extracted her voiceprint are described at ¶¶ 83-85, 95-97.

17. Plaintiff Lindsey Dorcus ("Dorcus") is a citizen of Illinois and resides in this District. Dorcus is a professional audiobook narrator who has recorded more than two hundred audiobooks for major American publishers from her home recording studio in Chicago. Dorcus's body of professional voice work, the public availability of her recordings, and the basis for Plaintiffs' allegation that Microsoft extracted her voiceprint are described at ¶¶ 83-85, 98-100.

18. Plaintiff Yohance Lacour ("Lacour") is a citizen of Illinois and resides in this District. Lacour is a journalist, audio storyteller, writer, and playwright whose investigative podcast *You Didn't See Nothin'* — produced in Chicago through the Invisible Institute — won the 2024 Pulitzer Prize for Audio Reporting. Lacour's body of professional voice work, the public availability of his recordings, and the basis for Plaintiffs' allegation that Microsoft extracted his voiceprint are described at ¶¶ 83-85, 101-103.

19. Plaintiff Victoria Nassif ("Nassif") is a citizen of Illinois and resides in this District. Nassif is a first-generation Lebanese-Palestinian American

actor, audiobook narrator, voiceover artist, and intimacy director whose professional voice work is produced in Illinois, including Arabic-accented narrations of works by Arab and Palestinian American authors released by Penguin Random House, Hachette, and Simon & Schuster. Nassif's body of professional voice work, the public availability of her recordings, and the basis for Plaintiffs' allegation that Microsoft extracted her voiceprint are described at ¶¶ 83-85, 104-106.

20. Defendant Microsoft Corporation ("Microsoft") is a Washington corporation with its principal executive offices at One Microsoft Way, Redmond, Washington 98052.

JURISDICTION AND VENUE

21. This Court has subject-matter jurisdiction over this action under the Class Action Fairness Act, 28 U.S.C. § 1332(d). The amount in controversy exceeds $5,000,000 in the aggregate, exclusive of interest and costs. The proposed Class includes more than 100 members. Minimal diversity is satisfied: Plaintiffs are citizens of Illinois, and Defendant Microsoft Corporation is a Washington corporation with its principal place of business in Washington. None of the exceptions to CAFA jurisdiction set forth in 28 U.S.C. § 1332(d)(3)–(5) applies.

22. The aggregate amount in controversy substantially exceeds $5,000,000. BIPA provides that an aggrieved person may recover the greater of liquidated damages or actual damages — $1,000 for a negligent violation or $5,000 for an intentional or reckless violation — and may recover those

statutory damages on a per-person, per-subsection basis where multiple distinct provisions of § 15 are violated, consistent with the statute as amended and as construed in *Clay v. Union Pacific Railroad Co.*, No. 25-2185 (7th Cir. Apr. 1, 2026). Plaintiffs seek recovery under five distinct BIPA subsections — § 15(a), (b), (c), (d), and (e) — each of which creates a distinct duty and supports a distinct per-person statutory or actual damages recovery, as well as recovery under IRPA, ICFA, IUDTPA, and common-law unjust enrichment. On information and belief, Microsoft extracted voiceprints and biometric information from voice recordings of hundreds of thousands of individuals, including a substantial number whose recordings were produced or recorded in Illinois. The aggregate damages across the proposed Class, together with the injunctive, equitable, and other relief sought, far exceed CAFA's jurisdictional threshold.

23. This Court has specific personal jurisdiction over Microsoft. Microsoft has purposefully directed its conduct at Illinois and this District in ways that bear directly on Plaintiffs' claims. Microsoft markets, sells, and delivers to Illinois customers the voice AI products at issue in this Complaint — including the prebuilt Azure neural voice catalog, Azure neural HD voices, Custom Neural Voice, Personal Voice, the Voice Live API, MAI-Voice-1, the Speaker Recognition API, Microsoft Teams voice attribution and transcription, and the voice products distributed through Azure AI Speech, Microsoft Foundry, Microsoft Copilot, Microsoft 365, and Dynamics 365 Contact Center — through subscription, consumption-based, and enterprise commercial

channels. Microsoft enters into recurring contractual relationships with Illinois residents and Illinois-based businesses through the Microsoft Services Agreement, the Microsoft Customer Agreement, Azure subscription terms, Microsoft 365 enterprise agreements, and the Custom Neural Voice and Personal Voice limited-access registration terms; accepts subscription and consumption payments from Illinois customers; and provides ongoing interactive service performance — including delivery of voice synthesis outputs in response to user requests — to users located in Illinois. Microsoft maintains office locations and a substantial operational presence in Illinois, including in the City of Chicago. On information and belief, Microsoft has substantial enterprise customer relationships in Illinois, including with Fortune 500 companies headquartered in this District and other Illinois-based enterprises across state and local government, healthcare, financial services, and education sectors, generating substantial Illinois revenue from Microsoft's $281.7 billion fiscal-year-2025 commercial operations.

24. Plaintiffs' claims arise directly from Microsoft's Illinois-directed conduct. Microsoft extracted the voiceprints of Plaintiffs and the Class — persons whose voice recordings were produced or recorded in Illinois — without the notice, consent, or written release BIPA requires. The voiceprints are now encoded in the parameters of Microsoft's commercial voice synthesis models, which Microsoft sells, delivers, and disseminates to Illinois customers and Illinois users through the channels identified above. Every commercial transaction Microsoft completes in Illinois involving those products monetizes

the biometric data Microsoft obtained from Plaintiffs and the Class without their consent. Microsoft's act of extracting voiceprints from voice recordings produced or recorded in Illinois is conduct directed at Illinois, regardless of the physical location of the servers on which the extraction occurred; Microsoft's continuing possession of those voiceprints, and its dissemination of them through its global Azure infrastructure and acquired Nuance speech-services business, is a continuing course of conduct directed at Illinois persons. Microsoft's Illinois contacts are not incidental to this Complaint; they are the downstream commercialization and ongoing dissemination of the upstream biometric extraction that gives rise to this Complaint.

25. Exercising specific personal jurisdiction over Microsoft in this District is consistent with the Illinois long-arm statute, 735 ILCS 5/2-209, and with due process requirements. Illinois has a strong interest in providing a forum for redress of unlawful biometric data collection from voice recordings produced or recorded in Illinois, an interest the Illinois General Assembly expressly identified in enacting BIPA, 740 ILCS 14/5. Plaintiffs, as Illinois residents whose biometric privacy was invaded in Illinois and whose recordings were produced or recorded in Illinois, have a corresponding interest in litigating their claims in their home forum. The burden on Microsoft of litigating in Illinois is not undue, given the resources alleged at ¶ 23 and Microsoft's experience as an active litigant in federal courts across the United States. Microsoft could reasonably anticipate being haled into Illinois courts for claims

like Plaintiffs' that arise from its extraction and commercial exploitation of biometric identifiers from voice recordings produced or recorded in Illinois.

26. Venue is proper in this District under 28 U.S.C. § 1391(b)(2) because a substantial part of the events giving rise to Plaintiffs' claims occurred here: Plaintiffs are residents of this District; their recorded vocal performances were produced and distributed from this District; the biometric privacy violations Plaintiffs allege occurred in this District; and the commercial exploitation of Plaintiffs' voiceprints continues through Microsoft's sale and delivery of the voice AI products at issue to customers in this District. Venue is independently proper under 28 U.S.C. § 1391(b)(1) because Defendant Microsoft is subject to personal jurisdiction in this District for the reasons set forth above and therefore "resides" in this District for venue purposes under 28 U.S.C. § 1391(c)(2).

FACTUAL BACKGROUND

Microsoft and Its Voice-Synthesis Business

27. Microsoft is one of the largest publicly traded technology companies in the world and one of the dominant suppliers of commercial AI products and infrastructure. Microsoft's commercial AI business spans cloud-platform subscriptions, enterprise software licensing, consumer software integrations, agent and assistant products, and an integrated portfolio of foundational AI models.

28. Microsoft's voice AI research, development, and commercial deployment are conducted across multiple internal organizations and one

wholly owned subsidiary, all operating under Microsoft Corporation. Microsoft Research operates the Text to Speech research project, established November 1, 2018, whose published outputs include the FastSpeech, AdaSpeech, NaturalSpeech, VALL-E, SpeechX, LiveSpeech, ELaTE, and DeepSinger model families. Azure AI Speech, a business unit of Azure within Microsoft Corporation, operates the commercial Speech service, which provides Microsoft's prebuilt Azure neural voice catalog, Azure neural HD voices, Custom Neural Voice, Personal Voice, the Voice Live API, and the Speaker Recognition API. Microsoft AI operates the MAI-Voice-1 model, which is deployed in Microsoft Copilot Daily, Copilot Podcasts, and Copilot Labs. Microsoft 365 includes Microsoft Teams, which implements speaker diarization and voice attribution features through its meeting transcription pipeline. Nuance Communications, Inc., a wholly owned subsidiary acquired by Microsoft in March 2022 in a $19.7 billion transaction, contributes additional speech-recognition and conversational-AI assets to Microsoft's voice-AI portfolio, including the Dragon family of dictation and conversational systems integrated into Microsoft's healthcare and contact-center offerings.

29. Microsoft has identified AI generally, and its voice and conversational-AI capabilities specifically, as core components of its commercial strategy in disclosures to the United States Securities and Exchange Commission and in public statements by its Chief Executive Officer, Satya Nadella. Microsoft publishes an annual Responsible AI Standard and an annual Responsible AI Transparency Report describing the company's claimed

approach to AI deployment, including voice technologies. Nadella has personally identified Microsoft Copilot, which integrates Microsoft's voice synthesis capabilities through MAI-Voice-1, Copilot Daily, and Copilot Podcasts, as the strategic centerpiece of Microsoft's AI commercialization in quarterly earnings calls, keynote presentations at Microsoft Ignite and Microsoft Build, and Microsoft's published corporate communications.

30. Microsoft distributes its commercial voice AI products and the underlying foundational voice synthesis models through an integrated commercial channel that Microsoft operates from the United States, including through Azure AI Speech (a paid Azure service available on subscription and consumption-based pricing); Microsoft Foundry and the Microsoft Foundry Model Catalog (Microsoft's AI development and deployment platform); Microsoft Copilot (integrated into Microsoft 365, Windows, GitHub, and Dynamics 365); Microsoft Teams (Microsoft's communications and collaboration platform, used by a substantial majority of the Fortune 500 and reported to have hundreds of millions of monthly active users globally); Dynamics 365 Contact Center (Microsoft's contact-center-as-a-service offering); and Microsoft 365 Copilot (Microsoft's productivity-suite AI agent). The voice AI research, model training, and commercial deployment alleged in this Complaint were carried out by Microsoft personnel in the United States, on infrastructure operated by Microsoft and its affiliated entities, in the ordinary course of Microsoft's business.

31. Microsoft's commercial voice AI products at issue in this Complaint include the following:

(a) The prebuilt Azure neural voice catalog and Azure neural HD voices, which provide commercial text-to-speech in over 150 languages and locales through Azure AI Speech.

(b) Custom Neural Voice, released in general availability in February 2021 and expanded into Dynamics 365 Contact Center in March 2026, a Limited Access feature that allows customers to train custom voice models from voice-talent recordings.

(c) Personal Voice, a Limited Access feature within Azure AI Speech that synthesizes replications of an individual's voice from short audio prompts in over 90 languages.

(d) The Voice Live API, announced in public preview at Microsoft Build in May 2025, provides low-latency speech-to-speech voice agent capabilities.

(e) MAI-Voice-1, announced by Microsoft AI on August 28, 2025, and made generally available in the Microsoft Foundry Model Catalog by April 2026, an expressive single- and multi-speaker voice generation model deployed in Microsoft Copilot Daily, Copilot Podcasts, and Copilot Labs.

(f) The Speaker Recognition API, which provides text-dependent and text-independent speaker verification and identification through Azure AI Speech.

(g) Microsoft Teams voice attribution and transcription, which deploys speaker-diarization technology in Microsoft's Teams meeting pipeline.

(h) The VALL-E, VALL-E X, VALL-E 2, NaturalSpeech, NaturalSpeech 2, NaturalSpeech 3, SpeechX, LiveSpeech, ELaTE, and DeepSinger research model families developed by Microsoft Research, certain of which (including VALL-E 2 and ELaTE) Microsoft has elected to retain on a research-only basis.

(i) The Dragon family and associated conversational and speech-recognition assets acquired through Microsoft's 2022 acquisition of Nuance Communications.

32. On information and belief, the training of Microsoft's foundational voice synthesis models occurred on Microsoft-operated computing infrastructure in the United States, including Azure GPU clusters owned and operated by Microsoft, and AI training infrastructure provisioned by Microsoft for its OpenAI partnership and Microsoft AI division. The production voice models that Microsoft deploys through its commercial products are hosted on Microsoft-controlled cloud infrastructure and served to United States users, including users in Illinois, from Azure data center facilities operated by Microsoft and its cloud and content-delivery partners. Microsoft reports its voice AI revenue within the Microsoft Cloud and Productivity and Business Processes segments of its consolidated financials.

How Microsoft's Voice AI Extracts and Encodes Voiceprints

33. Modern AI voice synthesis works by training neural networks on large quantities of recorded human speech. During training, the network learns to identify and reproduce the acoustic features that make individual voices distinctive — pitch, timbre, resonance, accent, cadence, articulation, and the

dynamics of emotional expression. The network encodes those features as mathematical representations and stores them in the parameters of the trained model. The same parameters are then used to generate new speech that exhibits the acoustic features of the training voices.[1]

34. Microsoft's published research and product documentation describe this process in technical detail, using terminology that aligns with the BIPA biometric definitions. Microsoft's data, privacy, and security documentation for Azure AI Speech states that Microsoft "may process biometric voice signatures" from a recorded acknowledgment statement and from "randomized audio from the training dataset(s)" to confirm same-speaker identity using Azure AI Speaker Verification. Microsoft's Speaker Recognition

[1] Microsoft's published research describes the architecture by which voice characteristics are extracted from speech recordings and encoded into the parameters of its voice models. Microsoft Research's Text to Speech project page describes the project's outputs across FastSpeech, AdaSpeech, NaturalSpeech, VALL-E, and DeepSinger; the underlying papers describe the architectures by which input audio is tokenized, encoded, and conditioned upon to produce speaker-faithful synthesis output. *See* Microsoft Research, Text to Speech, https://www.microsoft.com/en-us/research/project/text-to-speech/ (last visited on May 12, 2026); Yi Ren et al., *FastSpeech: Fast, Robust and Controllable Text to Speech*, arXiv:1905.09263 (May 22, 2019), https://arxiv.org/abs/1905.09263 (last visited on May 12, 2026); Mingjian Chen et al., *AdaSpeech: Adaptive Text to Speech for Custom Voice*, arXiv:2103.00993 (Mar. 2, 2021), https://arxiv.org/abs/2103.00993 (last visited on May 12, 2026); Zeqian Ju et al., *NaturalSpeech 3: Zero-Shot Speech Synthesis with Factorized Codec and Diffusion Models*, arXiv:2403.03100 (Mar. 5, 2024), https://arxiv.org/abs/2403.03100 (last visited on May 12, 2026); Chengyi Wang et al., *Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers*, arXiv:2301.02111 (Jan. 5, 2023), https://arxiv.org/abs/2301.02111 (last visited on May 12, 2026); Yi Ren et al., *DeepSinger: Singing Voice Synthesis with Data Mined From the Web*, arXiv:2007.04590 (July 9, 2020), https://arxiv.org/abs/2007.04590 (last visited on May 12, 2026). Plaintiffs allege on information and belief that the same architecture and processing pipeline was applied to the voice training data ingested for the foundational models that power Microsoft's commercial voice products.

API documentation states that “Voice features are extracted from the audio recording to form a unique voice signature."[2]

35. Microsoft's VALL-E paper describes converting speech waveforms into discrete audio codec codes and treating text-to-speech as conditional language modeling on those codes such that, from a three-second enrolled recording, the system can synthesize personalized speech preserving "the speaker's emotion and acoustic environment of the acoustic prompt."[3]

36. Microsoft's NaturalSpeech 3 paper describes a factorized neural codec that disentangles speech into distinct attribute representations, including a "timbre" representation specifically capturing speaker identity, which the model conditions on at inference time to produce speech in a target voice.[4]

37. Microsoft's DeepSinger paper describes a singing-voice synthesis pipeline trained on data "mined from the web" — specifically, a self-built crawl-and-extract system that ingested vocal recordings from music websites, separated vocals from instrumental accompaniment, aligned lyrics with audio,

---

[2] *See* Microsoft, *Data, privacy, and security for text to speech*, Microsoft Learn, https://learn.microsoft.com/en-us/legal/cognitive-services/speech-service/text-to-speech/data-privacy-security (last visited on May 12, 2026); Microsoft, *Speaker recognition overview*, Microsoft Learn, https://learn.microsoft.com/en-us/azure/ai-services/speech-service/speaker-recognition-overview (last visited on May 12, 2026).

[3] *See* Chengyi Wang et al., *Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers*, arXiv:2301.02111 (Jan. 5, 2023), https://arxiv.org/abs/2301.02111 (last visited on May 12, 2026).

[4] *See* Zeqian Ju et al., *NaturalSpeech 3: Zero-Shot Speech Synthesis with Factorized Codec and Diffusion Models*, arXiv:2403.03100 (Mar. 5, 2024), https://arxiv.org/abs/2403.03100 (last visited on May 12, 2026).

and filtered samples. The DeepSinger architecture is built to learn singer-specific acoustic characteristics from training audio and to reproduce those characteristics in synthesized output.[5]

38. Microsoft's disclosure-for-voice-and-avatar-talent documentation acknowledges that Microsoft products are used to "process Biometric Data" within the meaning of applicable data protection requirements, and defines "Biometric Data" by reference to Article 4 of the General Data Protection Regulation — a definition that expressly includes "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data," and which the European Data Protection Board has confirmed includes voiceprints.[6]

39. BIPA's definitions confirm the legal classification. The statute defines "biometric identifier" to include "voiceprint." 740 ILCS 14/10. It separately defines "biometric information" to include "any information, regardless of how it is captured, converted, stored, or shared, based on an

[5] *See* Yi Ren et al., *DeepSinger: Singing Voice Synthesis with Data Mined From the Web*, arXiv:2007.04590 (July 9, 2020), https://arxiv.org/abs/2007.04590 (last visited on May 12, 2026).

[6] *See* Microsoft, *Disclosure for voice and avatar talent*, Microsoft Learn, https://learn.microsoft.com/en-us/azure/ai-foundry/responsible-ai/speech-service/text-to-speech/disclosure-voice-talent (last visited on May 12, 2026); Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), art. 4(14), 2016 O.J. (L 119) 1, https://eur-lex.europa.eu/eli/reg/2016/679/oj; European Data Protection Board, *Guidelines 02/2021 on Virtual Voice Assistants* (Version 2.0, adopted July 7, 2021), https://www.edpb.europa.eu/system/files/2021-07/edpb_guidelines_202102_on_vva_v2.0_adopted_en.pdf (last visited on May 12, 2026).

individual's biometric identifier used to identify an individual." *Id.* The biometric voice signatures, voice signatures, audio tokens, factorized timbre representations, speaker embeddings, and other speaker-acoustic-signature representations produced by Microsoft's voice synthesis pipeline fall within both definitions. The label Microsoft and the AI research community use for these representations does not determine their statutory classification; the function of the data does.

40. Microsoft's Custom Neural Voice transparency note confirms this functional understanding. The note describes Custom Neural Voice as enabling the customer to "create a synthetic voice that sounds like" the voice talent and, in its Pro configuration, to produce a voice that "even more closely resembles your voice talent's voice." It further requires Microsoft's customers to "obtain explicit written permission from voice talent" before creating a synthetic voice.[7] Microsoft's contractual treatment of voice-talent recordings acknowledges that Microsoft's voice cloning capability is built on the extraction and encoding of speaker-identifying acoustic characteristics — exactly the capability the Azure AI Speech, Personal Voice, Voice Live API, and MAI-Voice-1 documentation describes, and exactly the capability that BIPA's voiceprint and biometric-information definitions reach.

41. On information and belief, the same category of biometric processing that occurs in Microsoft's customer-facing voice features — Custom

[7] *See* Microsoft, *Transparency note and use cases for Custom Neural Voice*, Microsoft Learn, https://learn.microsoft.com/en-us/legal/cognitive-services/speech-service/custom-neural-voice/transparency-note-custom-neural-voice (last visited on May 12, 2026).

Neural Voice, Personal Voice, the Voice Live API, MAI-Voice-1, the Speaker Recognition API, and Microsoft Teams voice attribution and transcription — also occurs during base-model training. The customer-facing features are applications of capabilities built during training. The training pipeline is where extraction of voiceprints and biometric information occurs first, at the largest scale, and on the broadest set of voice recordings — including, on information and belief, the recordings of Plaintiffs and the Class.

Microsoft Built Safeguards Against Misuse,
Not Consent for the Voices It Took

42. Microsoft understands that voice synthesis technology built on human voice data creates the categories of harm that BIPA was enacted to prevent. Microsoft has built and publicly deployed five categories of safeguard measures over the period 2018-2026, each directed at the downstream misuse of Microsoft's commercial voice products. Microsoft built none of those measures for the upstream voices it took to construct those products in the first place.

43. Microsoft operates Custom Neural Voice as a Limited Access feature. Customers must apply for access, identify approved use cases, and warrant that they have obtained explicit written permission from voice talent before training a voice model. Each Custom Neural Voice voice talent must record a Microsoft-provided acknowledgment statement, and Microsoft uses Azure AI Speaker Verification to confirm that the same person who recorded the acknowledgment is the source of the customer-provided training data.

44. Microsoft operates Personal Voice as a Limited Access feature within Azure AI Speech, requiring each voice talent to record a verbal consent acknowledgment statement and pass a speaker-verification check before Microsoft will synthesize speech in that voice.

45. Microsoft's Enterprise AI Services Code of Conduct conditions external-user voice-model training on a recorded acknowledgment statement evidencing the speaker's consent. Microsoft's Custom Neural Voice transparency note instructs customers to "obtain explicit written permission from voice talent" before training a model on that talent's voice.

46. Microsoft Research elected to retain VALL-E 2 and ELaTE on a research-only basis. VALL-E 2's project page expressly states: "VALL-E 2 is purely a research project. Currently, we have no plans to incorporate VALL-E 2 into a product or expand access," and recommends a "speaker approval" protocol if speaker-conditioned voice synthesis is ever generalized to unseen speakers. ELaTE's project page contains a similar acknowledgment.

47. Microsoft's disclosure-for-voice-and-avatar-talent documentation expressly allocates BIPA-equivalent compliance obligations to Microsoft's customers. The documentation states: "If you are using Microsoft products or services to process Biometric Data, you are responsible for: (i) providing notice to data subjects, including with respect to retention periods and destruction; (ii) obtaining consent from data subjects; and (iii) deleting the Biometric Data, all as appropriate and required under applicable Data Protection Requirements."

48. Each of the five measures described in ¶¶ 43-47 is a constraint at the downstream end of Microsoft's commercial pipeline, intended to restrict how customers use Microsoft's products or to impose the upstream-rights obligation on Microsoft's customers. None is directed at the upstream issue: that Microsoft's foundational voice synthesis models — including the prebuilt Azure neural voice catalog, Azure neural HD voices, MAI-Voice-1, and the VALL-E and NaturalSpeech research models — were trained on voiceprints extracted without the knowledge or consent of the speakers whose voice recordings were ingested during training. Microsoft built each of these safeguards to prevent downstream misuse of its products and to protect the customers and voice talent with whom it had a commercial relationship. It built none of them for the people whose voices it took to train those products in the first place.

Microsoft Trained on Data It Knew It Had No Right to Use

49. Microsoft has never published a comprehensive disclosure of the sources, scale, or provenance of the voice data used to train its commercial voice synthesis models. It has not published a model card, data sheet, training data manifest, licensing inventory, or transparency report identifying all of the voice recordings used to train the prebuilt Azure neural voice catalog, the Azure neural HD voices, MAI-Voice-1, the Voice Live API foundational models, or the foundational models underlying Custom Neural Voice and Personal Voice, where it obtained those recordings, or whether any of the speakers consented.

50. What Microsoft has disclosed for some of its research models is partial and self-incriminating. Microsoft's VALL-E paper publicly states that the model was "trained with LibriLight, a corpus consisting of 60K hours of English speech with over 7000 unique speakers", a corpus derived from LibriVox audiobook recordings without speaker-by-speaker consent. Microsoft's NaturalSpeech 2 paper reports training on approximately 44,000 hours of speech and singing data. Microsoft's NaturalSpeech 3 paper reports a 1-billion-parameter model trained on approximately 200,000 hours of data. Microsoft's DeepSinger paper expressly identifies the training data as "data mined from the web" through a self-built pipeline that crawled music websites, separated vocals from instrumental accompaniment, aligned lyrics, and filtered audio samples.

51. Microsoft's silence about the bulk of its commercial voice training data is made more conspicuous by the contrast with these research-model disclosures. If Microsoft's prebuilt Azure neural voices, Azure neural HD voices, and MAI-Voice-1 were trained entirely on properly licensed or consented data, Microsoft would have an obvious commercial incentive to say so. It has not.

52. The contemporaneous public record establishes that Microsoft obtained substantial portions of its training data, across multiple foundational-model families, from publicly accessible sources without the consent of the source speakers, authors, or content creators, and that Microsoft did so at industrial scale.

53. Microsoft is a defendant in two pending putative class actions concerning the training-data provenance of its other foundational AI models. In *The New York Times Co. v. Microsoft Corp.*, No. 1:23-cv-11195 (S.D.N.Y. filed Dec. 27, 2023), The New York Times alleges that Microsoft trained large language models, including the models that power Microsoft Copilot, on Times journalism without notice, consent, or compensation. In *Doe v. GitHub, Inc.*, No. 4:22-cv-06823 (N.D. Cal.), software developers allege that Microsoft, its GitHub subsidiary, and OpenAI trained the GitHub Copilot code-generation model on millions of public code repositories in violation of the open-source licensing terms under which the code was distributed.[8]

54. A parallel BIPA action concerning Microsoft's voice-biometric practices is pending in the Western District of Washington. *Basich et al. v. Microsoft Corp.*, No. 2:26-cv-00422 (W.D. Wash. filed Feb. 5, 2026), alleges that Microsoft Teams' speaker-diarization and transcription features capture and analyze voiceprints from meeting participants in violation of BIPA, and that Microsoft's collection and storage of those voiceprints occurred without the notice, written consent, or publicly available retention policy that BIPA requires.[9]

55. Microsoft's DeepSinger paper, taken on its own public terms, demonstrates that Microsoft maintains the infrastructure for the large-scale

[8] Plaintiffs cite the public allegations in *The New York Times* and *Doe v. GitHub* as circumstantial evidence of Microsoft's training-data acquisition practices across its foundational model families, not as adjudicated fact.

[9] Plaintiffs cite the *Basich* allegations as circumstantial evidence relating to Microsoft's voice-data practices, not as adjudicated fact.

ingestion and processing of human voice audio from publicly accessible platforms; that Microsoft has willingly ingested identifiable voice recordings from those platforms without obtaining the consent of the speakers whose voices appear in the recordings; and that Microsoft's voice-audio acquisition pipeline operates at a scale and on a category of source material, music recordings drawn from the public web, that aligns with the long-form, single-speaker, professionally produced audio on which Microsoft's commercial voice synthesis models depend.

56. Microsoft has not published, and on information and belief does not maintain, any mechanism that would allow speakers like Plaintiffs, whose voices were ingested into Microsoft's foundational voice synthesis models, to discover that fact, withdraw consent, or request deletion. Microsoft's general Privacy Policy provides certain rights to users with an existing relationship with Microsoft, but it does not address voice data used in AI model training and does not apply to non-users like Plaintiffs.

57. Microsoft's contemporaneous training-data acquisition practices, documented in published research, internal disclosures, and pending litigation across multiple foundational-model families, reflect a sustained institutional pattern of ingesting identifiable source material at industrial scale without the consent of the people whose work the source material represents. Microsoft operates the infrastructure required to do so for voice audio, has used it for the DeepSinger corpus, and has provided no mechanism by which the speakers whose voices were ingested can verify, withdraw consent, or request deletion.

Plaintiffs allege, on the basis of these public facts, that Microsoft trained its foundational voice synthesis models on voice recordings that included Plaintiffs' recordings, obtained from publicly accessible platforms, and extracted voiceprints from those recordings without the knowledge or consent of the speakers.

The Biometric Data Was Generated in Illinois

58. The conduct giving rise to Plaintiffs' claims is localized in Illinois in three independent and mutually reinforcing respects: (i) the biometric data at issue was generated in Illinois; (ii) Microsoft's acquisition of that biometric data targeted material that was identifiably Illinois-origin; and (iii) Microsoft's ongoing retention, dissemination, and commercial exploitation of the biometric data is directed at, or felt in, Illinois. Each ground supports the application of BIPA, IRPA, and the related Illinois statutes invoked in this Complaint to Microsoft's conduct as to Plaintiffs and the Class.

59. Plaintiffs' voices, the biological characteristics from which voiceprints and biometric information are derived, were produced by Plaintiffs while they were physically present in Illinois. Each voice came into existence in Illinois, in the body of an Illinois person speaking in Illinois. The audio recordings embodying Plaintiffs' voices, from which Microsoft is alleged to have extracted voiceprints and biometric information, were created in Illinois: Plaintiffs were physically present in Illinois at the time of the recordings, and the recordings captured the biometric characteristics of speakers who were in Illinois at the time of recording.

60. The recordings were published from Illinois to publicly accessible platforms. The biometric source material — the voice itself, the audio encoding of the voice, and the publicly distributed recordings embodying both — is therefore of Illinois origin, and the biometric data Microsoft is alleged to have extracted from those recordings is Illinois-origin biometric data regardless of the physical location of the servers on which Microsoft's extraction occurred.

61. Microsoft accessed Plaintiffs' voice recordings from publicly accessible third-party platforms — including YouTube, Spotify, Apple Podcasts, Audible, iHeartRadio, Amazon Music, and similar major audio distribution platforms — on which the recordings were hosted with metadata identifying each Plaintiff by name and identifying the recordings' Illinois origin. Each of those platforms displays creator profile information including the creator's name, geographic location, channel or artist description, and content catalog. This information was visible and accessible to Microsoft at the time of data acquisition. To a person with knowledge of Illinois broadcast, journalism, audiobook, podcast, and voice-performance markets, the metadata accompanying Plaintiffs' recordings identified each Plaintiff as a speaker whose work originated from Illinois.

62. Microsoft's act of accessing and processing those recordings was therefore not the passive receipt of an anonymized dataset assembled by a third party. It was the affirmative acquisition of identifiable, attributed voice content — voice content that, at the time of acquisition, was publicly associated with Illinois-based speakers and Illinois-origin production. The

Illinois affiliation of each Plaintiff and of each Plaintiff's recordings was knowable from publicly visible information at the time Microsoft accessed and processed the recordings.

63. Microsoft sells subscriptions and AI-generated voice outputs to Illinois subscribers and Illinois-based commercial users on a continuing basis through Azure AI Speech, Microsoft Foundry, Microsoft Copilot, Microsoft Teams, Dynamics 365 Contact Center, the Voice Live API, MAI-Voice-1 deployments, Custom Neural Voice and Personal Voice services, and acquired Nuance speech offerings. The commercial value of those subscriptions and outputs derives, on information and belief, from voice models developed through unconsented extraction of biometric data, including biometric data generated in Illinois. Each Illinois transaction in those products is an exercise of commercial value that Microsoft derived from the same biometric extraction Plaintiffs allege is unlawful.

64. Microsoft's continuing dissemination of the foundational voice models — encoded in model parameters and deployed across Microsoft's global Azure infrastructure, through Microsoft's acquired Nuance speech-services business, and through the third-party affiliates, vendors, subprocessors, and service providers that operate Microsoft's voice products at scale — is likewise Illinois-directed. The injury suffered by Plaintiffs and the Class is felt in Illinois: the privacy interest BIPA protects — the right to control the collection, retention, and commercial use of one's biometric identity — is exercised by Plaintiffs in Illinois. Plaintiffs' loss of control over their biometric data, their

loss of the licensing and consent rights BIPA preserves, and the dilution of the commercial value of their voices in markets in which they participate are all injuries Plaintiffs sustain in connection with their personal and professional activities, including activities in Illinois.

65. BIPA's notice and consent obligations under § 15(b) are duties that run to the subject whose biometric data is collected. The statute requires that the collecting entity inform "the subject" in writing and receive "a written release executed by the subject" before collection. 740 ILCS 14/15(b). The duty is owed to the subject, not to the location of the collecting entity's computational infrastructure.

66. The same is true of BIPA's retention obligations under § 15(a), profiting prohibitions under § 15(c), dissemination prohibitions under § 15(d), and reasonable-care obligations under § 15(e), each of which protects the persons from whom biometric data is taken. The subjects of the biometric data at issue in this Complaint, Plaintiffs and the Class, are persons whose voices and recordings were generated in Illinois. Microsoft's continuing possession of voiceprints and biometric information generated in Illinois runs afoul of obligations BIPA imposes for the benefit of the subjects of that data, including Microsoft's failure to publish a retention schedule and destruction policy compliant with 740 ILCS 14/15(a) covering biometric data obtained from non-user training-data sources, and Microsoft's failure to provide any mechanism by which Plaintiffs or any other non-user training-data subject may seek removal or destruction of their biometric data.

Microsoft's Monetization of the Voice Models

67. Microsoft's commercial exploitation of the foundational voice synthesis models is not limited to a single product line. Microsoft monetizes the voice models through a vertically integrated commercial chain spanning Azure cloud subscriptions and consumption-based pricing, Microsoft Foundry deployments, Microsoft Copilot integrations across Microsoft 365 and Windows, Microsoft Teams premium licensing tiers and meeting-transcription features, Dynamics 365 Contact Center deployments, real-time voice agent licensing through the Voice Live API, Limited Access deployments of Custom Neural Voice and Personal Voice, and the conversational and speech-recognition assets of Microsoft's acquired Nuance Communications subsidiary. The voice characteristics learned during training, the same characteristics extracted from Plaintiffs' and Class members' voice recordings, enable each of these revenue streams.

68. Microsoft monetizes the voice models through, among other channels:

(a) *Azure AI Speech subscription and consumption revenue.* Azure AI Speech is a paid Azure service available on tiered consumption-based pricing, with enterprise commitments and committed-use discounts for high-volume customers. Microsoft charges Illinois customers and enterprises worldwide for Azure AI Speech text-to-speech, custom voice training, and speaker verification on a per-character, per-API-call, and per-deployed-model basis. Each

transaction generates revenue that Microsoft captures because the voice models exist.

(b) *Microsoft Foundry deployment revenue.* The Microsoft Foundry Model Catalog distributes Microsoft's voice synthesis models — including MAI-Voice-1 — to enterprise customers for production deployment on Azure. Microsoft charges enterprise customers for Foundry deployments and the associated compute consumption.

(c) *Microsoft Copilot subscription revenue.* Microsoft Copilot, into which Microsoft's voice synthesis capabilities are integrated through MAI-Voice-1 (powering Copilot Daily and Copilot Podcasts), is licensed to Microsoft 365 customers on per-user subscription pricing as part of Microsoft 365 Copilot, Copilot Pro, and Copilot for individual products. Microsoft's voice synthesis capabilities are a marketed component of the Copilot subscription value proposition.

(d) *Microsoft Teams premium licensing.* Microsoft Teams Premium and the Microsoft Teams voice and transcription features generate licensing revenue across Microsoft's enterprise customer base. Microsoft Teams' speaker-diarization and voice-attribution features rely on the foundational speaker-recognition and voice-processing capabilities at issue in this Complaint.

(e) *Dynamics 365 Contact Center revenue.* Microsoft launched custom neural voices for Dynamics 365 Contact Center in March 2026, integrating Microsoft's voice synthesis capabilities into enterprise contact-center deployments on a per-seat and per-deployment licensing basis.

(f) *Voice Live API consumption revenue.* The Voice Live API, launched in public preview at Microsoft Build in May 2025 with pricing announced June 30, 2025, generates per-minute and per-session revenue for real-time speech-to-speech voice agent deployments.

(g) *Custom Neural Voice and Personal Voice Limited Access revenue.* Limited-access deployments of Custom Neural Voice and Personal Voice generate per-deployment, per-voice-model, and per-character synthesis fees for customer-trained custom voices, all of which rely on the same foundational voice-synthesis architecture at issue in this Complaint.

(h) *Acquired Nuance speech-services revenue.* Microsoft's acquired Nuance Communications business operates enterprise speech-recognition and conversational-AI offerings — including the Dragon dictation and conversational families integrated into Microsoft's healthcare, customer-service, and contact-center products — that contribute additional revenue to Microsoft's consolidated voice-AI business. Microsoft paid $19.7 billion to acquire this business, completed in March 2022.

69. Each monetization channel depends on the foundational voice models. Each foundational voice model depends on the training data used to build it. And the training data, on information and belief, includes the voice recordings of Plaintiffs and Class members from which Microsoft extracted voiceprints without notice or consent. The commercial exploitation Microsoft captures across the full chain is, in operational substance, the monetization of biometric data Microsoft obtained from Plaintiffs without authorization.

Microsoft's Dissemination of the Voice Models
Across Its Corporate, Cloud, and Acquired-Entity Infrastructure

70. The voice synthesis models in which Plaintiffs' and Class members' voiceprints are encoded are not stored in a single location or accessed by a single legal entity. They are deployed, served, processed, and transferred across multiple categories of infrastructure Microsoft operates in the ordinary course of its voice AI business.

71. Microsoft conducts its global voice AI business through a corporate structure that includes its wholly owned subsidiary Nuance Communications, Inc. and additional international affiliates including Microsoft International Holdings B.V. and Microsoft Ireland Operations Limited. The development, training, refinement, and commercial deployment of Microsoft's voice models requires, on information and belief, the transfer of those models — and the voiceprints encoded within them — among Microsoft's affiliated entities and acquired Nuance entities in the ordinary course of business. Plaintiffs and Class members did not consent to any such transfer.

72. Microsoft's voice synthesis pipeline operates within a global subprocessor framework that includes Azure cloud infrastructure providers, content delivery network operators, hardware partners, and third-party vendors. The pipeline's training, evaluation, deployment, and operation involve, on information and belief, the transmission of model parameters — and of the voiceprints encoded within those parameters — among Microsoft's vendors and service providers in the ordinary course of business. The delivery of voice synthesis services to enterprise customers, including Illinois customers,

necessarily involves transmitting model parameters and inference outputs over that infrastructure.

73. Microsoft Research distributes certain of its speech-synthesis research outputs through GitHub and similar platforms, including reference implementations and demonstration code associated with the VALL-E, NaturalSpeech, and DeepSinger model families. To the extent such research releases include or enable the reconstruction of model parameters trained on Plaintiffs' voice recordings, those releases constitute additional dissemination of biometric data encoded in those parameters.

74. The pattern of Microsoft's dissemination practices, taken together, places the voiceprints and biometric information of Plaintiffs and Class members in the hands of Microsoft's affiliated entities, acquired Nuance entities, vendors, subprocessors, service providers, and, to the extent any research-model release exposes the underlying parameters, third parties who may have downloaded such releases. Plaintiffs and Class members did not consent to any of these disseminations. No enumerated exception to BIPA's dissemination prohibition under 740 ILCS 14/15(d) applies to the disseminations described in this subsection.

The Voice Products Compete With the People Whose Voices Built Them

75. The voice AI products built on Plaintiffs' biometric data are now sold and deployed into the markets where Plaintiffs and the Class earn their livelihoods.

76. Professional audiobook narration historically has been performed by human narrators at industry rates of approximately $250 to $400 per finished hour, meaning a typical ten-hour novel costs $3,000 to $4,000 to narrate. Microsoft's voice synthesis products, including Custom Neural Voice, Personal Voice, the prebuilt Azure neural voice catalog, and the Azure neural HD voices, generate professional-quality narration at a fraction of those rates and can render an entire audiobook in hours. The major American audiobook publishers, including Penguin Random House, Hachette, Simon & Schuster, and Macmillan, for whom Plaintiffs Dorcus and Nassif have narrated books, have begun to integrate AI-generated narration into their production pipelines.

77. Microsoft's MAI-Voice-1 generates expressive single- and multi-speaker speech at scale and is deployed by Microsoft in Copilot Daily (an automated daily news briefing experience) and Copilot Podcasts (an AI-generated podcast feature within Microsoft Copilot). These deployments substitute for the long-form audio journalism, podcast hosting, narrative non-fiction production, and on-air reporting that Plaintiffs Marin, Rogers, Lacour, Flowers, and Amer have spent their careers developing and producing, until now, by human reporters and narrators at substantial professional cost.

78. Microsoft's Voice Live API and Custom Neural Voice deployments into Dynamics 365 Contact Center provide real-time, multilingual, AI-generated voice agents and contact-center voices for enterprise customers at per-character and per-minute pricing orders of magnitude lower than the cost of human voice talent. These deployments substitute directly for the broadcast,

contact-center, advertising voiceover, and customer-service voice work for which Plaintiffs Marin and Rogers built their careers and for which Plaintiffs Dorcus and Nassif provide their voice acting and voiceover services.

79. Microsoft's prebuilt Azure neural voice catalog and Azure neural HD voices provide commercial text-to-speech in over 150 languages and locales. Microsoft's Custom Neural Voice and Personal Voice support synthesized speech in over 90 languages. Microsoft's voice synthesis capability across multiple languages and accents includes, on information and belief, the capability to generate Arabic-accented English narration, Levantine Arabic narration, and other culturally and linguistically specific voice content. This capability competes specifically with Plaintiff Nassif's distinctive professional market position as a Lebanese-Palestinian American audiobook narrator providing authentic Arabic-accented narration of works by Arab and Palestinian American authors.

80. Microsoft's voice synthesis products are marketed for podcast and audiobook production, studio dubbing, e-learning narration, advertising voiceover, contact-center voice work, and conversational-AI deployments. Each of these markets historically has been served by professional voice actors, narrators, journalists, and broadcasters — including Plaintiffs. Microsoft's products substitute for that human work at per-character and per-audio-second pricing that is orders of magnitude lower than human voice talent rates.

81. Microsoft has publicly demonstrated the use of its voice synthesis models for commercial applications that historically required human voice talent, including in Microsoft Copilot Daily, Copilot Podcasts, Copilot Labs experiences, and enterprise contact-center deployments. Each of these deployments converts work that historically commanded professional voice talent rates into output produced at machine-scale costs.

82. Each of Microsoft's voice products was built using the vocal characteristics of the human performers it now displaces. Plaintiffs allege, on information and belief, that the voiceprints of every Plaintiff in this case are among the voiceprints encoded in the foundational voice synthesis models on which these products depend. The market substitution is therefore not merely temporal but causal. The machines generate the audio they generate because the performers' voiceprints were extracted from their recordings and embedded in the model parameters that produce the output.

Named Plaintiffs

Allegations Common to Each Named Plaintiff

83. On information and belief, the voice recordings of each named Plaintiff were among the audio that Microsoft ingested to train its foundational voice synthesis models, and the voiceprints and biometric information derived from those recordings are encoded in the parameters of Microsoft's commercial voice models and reproduced in the audio those models generate. Each named Plaintiff's catalog of professional voice work matches the profile of training audio that Microsoft's technical documentation identifies as optimal — long-

form, single-speaker, studio-quality, professionally produced, identifiable by name and source, and continuously available on publicly distributed audio platforms, and corresponds, in form and content, to the categories of training material identified in Microsoft's published research, including the audiobook narration corpus on which Microsoft's VALL-E paper specifically reports training (60,000 hours of single-speaker audiobook narration drawn from LibriVox via the LibriLight corpus, comprising over 7,000 unique speakers).

84. Microsoft has not disclosed, and has refused to disclose, the sources of the voice training data used to develop its commercial voice synthesis models. Microsoft's DeepSinger paper publicly describes the self-built crawl-and-extract pipeline Microsoft maintains for the large-scale ingestion of human voice audio from publicly accessible platforms. The records identifying which specific voice recordings Microsoft ingested into the training corpora that produced the prebuilt Azure neural voice catalog, the Azure neural HD voices, MAI-Voice-1, and the foundational models underlying Custom Neural Voice and Personal Voice are within Microsoft's exclusive control.

85. No named Plaintiff created an account with Microsoft for the purpose of voice synthesis. No named Plaintiff uploaded a voice recording to Azure AI Speech, Speech Studio, Microsoft Foundry, Custom Neural Voice, Personal Voice, the Voice Live API, or any other Microsoft voice-synthesis product. No named Plaintiff accepted Microsoft's Custom Neural Voice or Personal Voice limited-access registration terms, and no named Plaintiff recorded a voice-talent acknowledgment statement for Microsoft. No named

Plaintiff received notice that Microsoft had collected his or her voiceprint, any disclosure of the specific purpose or duration of that collection, or executed a written release. Microsoft's collection of each named Plaintiff's voiceprint, and Microsoft's continuing possession and commercial exploitation of those voiceprints, occurred and continues without that Plaintiff's knowledge or consent.

Alison Flowers

86. Alison Flowers is an investigative journalist, podcast producer, and audio storyteller whose career has been conducted in Chicago. Flowers is the founder of Spiralbound, a Chicago-based production company operating at the intersection of investigative journalism and immersive storytelling, and previously served as Head of Production at the Invisible Institute, a nonprofit investigative journalism organization on the South Side of Chicago. Flowers was producer and reporter of the seven-part investigative podcast *Somebody*, which premiered March 31, 2020 and investigated the 2016 murder of Courtney Copeland. *Somebody* was a 2021 Pulitzer Prize finalist for Audio Reporting, won the National Magazine Award (Ellie) for Podcasting, the Scripps Howard Award for Excellence in Radio/Podcast Coverage, the International Documentary Association's Best Audio Documentary award, a National Headliner Award, a Gracie Award, and the 2020 Third Coast International Audio Festival Award for Best Serialized Story; it was named the No. 1 true-crime podcast of all time by Rolling Stone, ranked first on The New York Times' list of true-crime podcasts at the intersection of race, ranked second on The

Atlantic's "The 50 Best Podcasts of 2020," and reached the number-two position among true-crime podcasts on the Apple Podcasts chart. Flowers's reporting and on-air work additionally appear on Reveal from the Center for Investigative Reporting, The Heist from the Center for Public Integrity, Vox, Dateline NBC, and Democracy Now!.

87. Flowers's audio catalog comprises hundreds of hours of single-speaker, studio-quality recorded vocal performance — investigative narration, on-air reporting, and recorded interviews — continuously available through Apple Podcasts, Spotify, iHeartRadio, YouTube, Stitcher, and the major podcast distribution platforms on which *Somebody* and Flowers's other audio work has been published. *Somebody* was additionally distributed through Google Podcasts before Google shut that service down in 2024. The platforms display metadata identifying Flowers as the producer, reporter, and on-air voice; identifying Chicago and Spiralbound (or, for earlier work, the Invisible Institute) as the production origin; and identifying the date and series of publication. Flowers's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

88. Flowers's injury is concrete and particular to her. Microsoft extracted her voiceprint from recordings she produced over the course of her career as an investigative journalist, podcast producer, and audio storyteller in Chicago, encoded that voiceprint in commercial models, and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its

operational form, it is the same biological and behavioral signature Flowers uses to speak every day. The technology Microsoft built using Flowers's voiceprint — including MAI-Voice-1 deployed in Microsoft Copilot Daily and Copilot Podcasts, the Voice Live API, Custom Neural Voice, and Personal Voice — now operates in the same long-form investigative audio journalism and podcast-production market in which Flowers continues to work, against the same colleagues with whom she collaborates, in a competitive position Flowers neither chose nor authorized.

Philip Rogers

89. Philip Rogers is a broadcast journalist whose four-decade career was conducted in and from Chicago, primarily at WBBM Newsradio (CBS) and WMAQ-TV (NBC 5 Chicago). His on-air work spans radio reporting at WBBM Newsradio, television reporting and anchoring at WMAQ-TV, and live broadcast coverage from conflict zones, disaster scenes, the Olympic Games, mass shootings, corruption trials, and major national and international events. Rogers has been recognized with a National Emmy Award, the Edward R. Murrow Award, five Associated Press Best Reporter honors, and multiple Peter Lisagor Awards from the Chicago Headline Club. The Lisagor Awards, conferred by Chicago's professional journalism organization, recognize excellence in journalism conducted within the Chicago metropolitan region.

90. Rogers's broadcast catalog comprises thousands of hours of single-speaker, studio-quality audio, continuously available through the NBC 5 Chicago digital archive at nbcchicago.com, where his on-air reports,

investigative segments, and broadcast news stories are archived and searchable, and on YouTube, including a career-retrospective interview conducted by the Illinois News Broadcasters Association in which Rogers reflects on four decades of on-air reporting from war zones, mass shootings, corruption trials, the Olympic Games, and other major events. The platforms display metadata identifying Rogers as the speaker and Chicago as the location of production. Rogers's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

91. Rogers's injury is concrete and particular to him. Microsoft extracted his voiceprint from recordings he produced over four decades of professional broadcast work, encoded that voiceprint in commercial models, and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its operational form, it is the same biological and behavioral signature Rogers used to speak every day during his career. Although Rogers is now retired, the technology Microsoft built using Rogers's voiceprint — including the prebuilt Azure neural voice catalog, Azure neural HD voices, MAI-Voice-1 deployed in Microsoft Copilot Daily and Copilot Podcasts, and custom neural voices deployed in Dynamics 365 Contact Center — now operates in the same broadcast journalism, contact-center, and voiceover markets where he built his career, against the same colleagues with whom he worked, in a competitive position Rogers neither chose nor authorized.

Carol Marin

92. Carol Marin is a five-decade investigative broadcast journalist whose career has been conducted substantially in and from Chicago. Her on-air work has aired on NBC (WMAQ-TV), CBS News (60 Minutes, 60 Minutes II, the CBS Evening News), WTTW (Chicago Tonight), CNN, and the Discovery Channel, and includes television anchoring, investigative reporting, debate moderation, documentary narration, and long-form interviewing. Marin has been recognized with three George Foster Peabody Awards (including a Personal Peabody), two Alfred I. duPont–Columbia University Awards, two National Emmy Awards, fifteen Regional Emmy Awards, the George Polk Award, the Gracie Award, and the Sigma Delta Chi Ethics in Journalism Award. She has been inducted into the Chicago Journalism Hall of Fame. In 2025, the Governor of Illinois designated Marin a Lincoln Laureate and awarded her the Order of Lincoln, the State of Illinois's highest civilian honor.

93. Marin's broadcast catalog comprises thousands of hours of single-speaker, studio-quality audio. Substantial portions of that catalog are continuously and publicly available through the Media Burn Independent Video Archive (mediaburn.org), a Chicago-based nonprofit archive that preserves Marin's broadcast investigative reporting and documentary work; through the WTTW digital archive at news.wttw.com; through the NBC 5 Chicago digital archive at nbcchicago.com; and on YouTube, where Marin's Peabody acceptance speeches, debate moderation, and archived broadcast segments are publicly accessible. Each platform displays metadata identifying

Marin as the speaker, the producing program, and Chicago as the production location. Marin's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

94. Marin's injury is concrete and particular to her. Microsoft extracted her voiceprint from recordings she produced over five decades of professional broadcast and documentary work in Chicago, encoded that voiceprint in commercial models, and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its operational form, it is the same biological and behavioral signature Marin used to speak every day during her career. Although Marin is no longer producing daily on-air work, the technology Microsoft built using Marin's voiceprint — including the prebuilt Azure neural voice catalog, Azure neural HD voices, MAI-Voice-1 deployed in Microsoft Copilot Daily and Copilot Podcasts, and custom neural voices deployed in Dynamics 365 Contact Center — now operates in the same broadcast journalism, investigative reporting, and documentary narration markets where she built her career, against the same colleagues with whom she worked, in a competitive position Marin neither chose nor authorized.

Robin Amer

95. Robin Amer is a journalist, podcast creator, audio producer, and on-air host whose work has been produced substantially in and from Chicago. Amer is the creator, host, narrator, and showrunner of USA Today's *The City*, a critically acclaimed investigative podcast whose first season peaked at No. 6 on

the Apple Podcasts charts and was named Best Podcast of the Year by The New Yorker, The New York Times, Quartz, and Apple Podcasts. Amer was a producer for WBEZ Chicago Public Media, the NPR affiliate in Chicago; contributed to the podcast *Gravy*, produced by the Southern Foodways Alliance, which won the 2015 James Beard Award for Best Podcast; served three years as Senior Producer for Audio Features at The Washington Post, where she edited *Post Reports* and the standalone narrative series *Field Trip* and won or was a finalist for the Alfred I. duPont–Columbia University Award three consecutive years; and currently serves as Managing Editor of *Love + Radio*, whose ten-part narrative series *Blood Memory* won the Tribeca Festival Audio Storytelling prize for Best Independent Non-Fiction in June 2025 and was shortlisted for the Whickers Prize. Amer's audio reporting and production work has also aired on Reveal from the Center for Investigative Reporting, The Heist from the Center for Public Integrity, and Vox.

96. Amer's audio catalog comprises hundreds of hours of single-speaker and multi-speaker, studio-quality recorded vocal performance — on-air hosting, recorded interviews, podcast narration, and field reporting — continuously available through Apple Podcasts, Spotify, iHeartRadio, Amazon Music, YouTube, and the major podcast distribution platforms on which *The City*, *Blood Memory*, *Gravy*, *Post Reports*, *Field Trip*, and Amer's other audio work has been published. *The City* and certain of Amer's other work were additionally distributed through Google Podcasts before Google shut that service down in 2024. The platforms display metadata identifying Amer as the

host, producer, and narrator; the producing organization; the date and series of publication; and the Chicago or Washington production location for the relevant segments. Amer's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

97. Amer's injury is concrete and particular to her. Microsoft extracted her voiceprint from recordings she produced over more than twenty years of professional audio journalism and podcast production work, encoded that voiceprint in commercial models, and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its operational form, it is the same biological and behavioral signature Amer uses to speak every day. The technology Microsoft built using Amer's voiceprint — including MAI-Voice-1 deployed in Microsoft Copilot Daily and Copilot Podcasts, the Voice Live API for real-time voice agents, Custom Neural Voice, and Personal Voice — now operates in the same long-form audio journalism, podcast hosting, and audio production market in which Amer continues to work, against the same colleagues with whom she collaborates, in a competitive position Amer neither chose nor authorized.

Lindsey Dorcus

98. Lindsey Dorcus is a professional audiobook narrator and voice actor based in Chicago. Dorcus has recorded more than two hundred audiobooks for major American publishers including Penguin Random House, Simon & Schuster, Macmillan, Hachette, Disney Hyperion, Audible Studios,

Blackstone Publishing, Tantor Media, Harper Audio, Podium, and Scribd, working from a professional home recording studio in Chicago equipped with industry-standard audiobook production equipment. Dorcus won the Society of Voice Arts and Sciences (SOVAS) Voice Arts Award in 2020 as part of the full-cast voice artist ensemble for the audiobook anthology *Wild Monsters Dance About: Stories from an Unruly Mind*, and the Independent Audiobook Award in 2021 for LGBTQ+ audiobook narration. AudioFile Magazine has reviewed Dorcus's narration favorably. Dorcus's professional range spans General American, British (Received Pronunciation, Estuary, and Cockney), Scottish, Irish (Dublin and Northern Irish), French, American Southern, and Greek for main characters and narration, and New England, New York, German, Indian, and Russian for supporting characters.

99. Dorcus's audiobook catalog comprises thousands of hours of single-speaker, studio-quality, professionally produced narration, continuously available through Audible (a subsidiary of Amazon), Apple Books, Google Play Books, Spotify, Libro.fm, Chirp, Scribd/Everand, and the other major audiobook retailers and subscription services that distribute her work. Her complete catalog of narrated titles is searchable and accessible on Audible.com and the other major audiobook platforms. The platforms display metadata identifying Dorcus as the narrator, the publisher, the title and series, and the publication date for each title. Dorcus's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

100. Dorcus's injury is concrete and particular to her. Microsoft extracted her voiceprint from recordings she produced over a career narrating more than two hundred audiobooks for the major American publishers, encoded that voiceprint in commercial models, and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its operational form, it is the same biological and behavioral signature Dorcus uses to speak and to perform every day. The technology Microsoft built using Dorcus's voiceprint — including the prebuilt Azure neural voice catalog (over 150 languages and locales), Azure neural HD voices, Custom Neural Voice, Personal Voice (over 90 languages), and MAI-Voice-1 — now operates in the same audiobook narration market where Dorcus earns her livelihood, generating synthetic narration that the major American audiobook publishers for whom Dorcus works have begun to integrate into their production pipelines, in direct competition with the human narration Dorcus provides and at per-character pricing orders of magnitude lower than the industry rates that have historically supported professional audiobook narration.

Yohance Lacour

101. Yohance Lacour is a journalist, audio storyteller, writer, playwright, and entrepreneur from the South Side of Chicago who has dedicated his professional work to telling stories of Black Chicago. Lacour is the creator, host, writer, and lead reporter of the seven-part investigative and memoir podcast *You Didn't See Nothin'*, a production of the Invisible Institute and USG Audio in which Lacour revisits the 1997 hate-crime attack on Lenard

Clark, which he had covered as a young community journalist, tracks down key players in the case a quarter-century later, and examines how the case and its aftermath shaped his own life. *You Didn't See Nothin'* won the 2024 Pulitzer Prize for Audio Reporting and the 2024 Peabody Award; was named to Apple Podcasts' "100 Best Podcasts of All Time" and "Podcasts We Love" lists; received nominations at the Signal Podcasting Awards in the categories of Limited Series & Specials – Best Host and Limited Series & Specials – Documentary; and received four nominations at the Black Podcasting Awards. Lacour serves as the on-air voice, narrator, and lead reporter throughout the series, conducting recorded interviews and providing oral narration. Lacour has additionally appeared as a featured guest and interview subject on NPR's Fresh Air with Tonya Mosley, NPR's All Things Considered with Adrian Florido, the Canadian Broadcasting Corporation's Crime Story podcast with Kathleen Goldhar, and the Pulitzer Prize Board's *Pulitzer on the Road* podcast produced by Audacy's Pineapple Street Studios.

102. Lacour's audio catalog comprises hundreds of hours of single-speaker, studio-quality recorded vocal performance — investigative narration, on-air reporting, recorded interviews, and broadcast guest appearances — continuously available through Apple Podcasts, Spotify, iHeart Podcasts, YouTube, Amazon Music, Overcast, and the major podcast and public-radio distribution platforms on which *You Didn't See Nothin'* and Lacour's other audio work has been published. The platforms display metadata identifying Lacour as the host, narrator, and reporter; the producing organization (the Invisible

Institute and USG Audio for the podcast itself); the date and series of publication; and the Chicago production location. Lacour's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

103. Lacour's injury is concrete and particular to him. Microsoft extracted his voiceprint from recordings he produced over the course of a career as a journalist, audio storyteller, and podcast creator focused on telling stories of Black Chicago, encoded that voiceprint in commercial models, and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its operational form, it is the same biological and behavioral signature Lacour uses to speak every day. The technology Microsoft built using Lacour's voiceprint — including MAI-Voice-1 deployed in Microsoft Copilot Daily and Copilot Podcasts, the Voice Live API, Custom Neural Voice, and Personal Voice — now operates in the same long-form investigative audio journalism and podcast-production market in which Lacour continues to work, against the same colleagues with whom he collaborates, in a competitive position Lacour neither chose nor authorized.

Victoria Nassif

104. Victoria Nassif is a first-generation Lebanese-Palestinian American actor, audiobook narrator, voiceover artist, and intimacy director whose professional voice work is produced in Illinois. Nassif is a trained mezzo-soprano singer whose audiobook narration has been commercially released by

Penguin Random House (through its Random House Audio imprint), Hachette Book Group (through its Little, Brown Young Readers imprint), and Simon & Schuster. Her notable narrations include *The Next New Syrian Girl* by Ream Shukairy (Hachette/Little, Brown Young Readers), for which she served as the solo narrator performing multiple characters with authentic Levantine Arabic accents; *The Skin and Its Girl* by Sarah Cypher (Random House Audio), a novel featuring a queer Palestinian American protagonist that was shortlisted for the Ursula K. Le Guin Prize and named a Best Book of the Year by *them* magazine; *Gulf* by Mo Ogrodnik (Simon & Schuster); *The Jasad Crown* by Sara Hashem; and *Every Moment is a Life*, a bilingual Arabic-English anthology compiled by Susan Abulhawa. Nassif's on-camera work includes multiple episodes of *Chicago PD* on NBC (Season 12) and nationally broadcast commercials. Her professional range spans General American, British (Received Pronunciation and Cockney), Persian, Arabic (Levantine dialect), and American Southern. As a first-generation Lebanese-Palestinian American who speaks basic Arabic, Nassif brings native cultural and linguistic authenticity to her narration of works featuring Middle Eastern characters and settings — a distinctive professional asset in the audiobook narration industry.

105. Nassif's audiobook catalog comprises dozens of hours of single-speaker, studio-quality, professionally produced narration, continuously available through Audible, Apple Books, Spotify, Libro.fm, and the other major audiobook retailers and subscription services that distribute her work. Her broadcast television work is continuously available through NBC's distribution

channels for *Chicago PD*. The platforms display metadata identifying Nassif as the narrator or performer, the publisher or production company, the title and series, and the publication date. Nassif's body of work has been continuously available through these channels for years preceding Microsoft's training of the foundational voice synthesis models alleged in this Complaint.

106. Nassif's injury is concrete and particular to her. Microsoft extracted her voiceprint from recordings she produced over the course of a career as a professional actor, audiobook narrator, and voiceover artist whose distinctive market position depends on authentic Arabic-accented narration of works by Arab and Palestinian American authors. Microsoft encoded that voiceprint in commercial models and continues to profit from those models. The voiceprint cannot be recovered or replaced; in its operational form, it is the same biological and behavioral signature Nassif uses to speak, sing, and perform every day. The technology Microsoft built using Nassif's voiceprint — including the prebuilt Azure neural voice catalog (over 150 languages and locales), Personal Voice (over 90 languages), Custom Neural Voice, MAI-Voice-1, and Microsoft's multilingual voice synthesis capabilities marketed for audiobook narration, dubbing, advertising voiceover, and contact-center voice work — now operates in the same audiobook narration, voice acting, and on-camera commercial markets where Nassif earns her livelihood, generating synthetic narration in Microsoft's expansive language and locale catalog that includes the capability to produce Arabic-accented English narration and other culturally and linguistically specific voice content in direct competition with the

distinctive professional market position Nassif has built as a Lebanese-Palestinian American narrator providing authentic Arabic-accented performance of works by Arab and Palestinian American authors.

Microsoft Acted Willfully and Recklessly

107. Microsoft's collection, retention, commercial exploitation, and dissemination of Plaintiffs' voiceprints without notice or consent was not the result of inadvertence or unfamiliarity with BIPA. Microsoft acted with knowledge of, or at a minimum, reckless disregard for, its obligations under Illinois law.

108. By the time Microsoft commercialized Custom Neural Voice in February 2021, deployed VALL-E in January 2023, and announced MAI-Voice-1 in August 2025, BIPA had been the law of Illinois for more than fifteen years and had generated some of the largest privacy settlements in American history. Those settlements include *In re Facebook Biometric Information Privacy Litigation*, No. 3:15-cv-03747 (N.D. Cal.) (approximately $650 million); *Rivera v. Google LLC*, Cook County Cir. Ct. No. 2019-CH-00990 (approximately $100 million); and *In re TikTok, Inc., Consumer Privacy Litigation*, No. 1:20-cv-04699 (N.D. Ill.) (approximately $92 million). In late 2025, Google paid $1.375 billion to the State of Texas to resolve parallel claims under Texas's biometric identifier statute, following Texas's $1.4 billion settlement with Meta in 2024. By any measure, Microsoft was on notice of the biometric-privacy obligations that govern the extraction of biometric identifiers from human source material

throughout the period in which it ingested Plaintiffs' voiceprints into its foundational voice synthesis models.

109. Voice-AI-specific litigation sharpened the notice further. In May 2024, voice actors filed *Lehrman v. Lovo, Inc.*, No. 1:24-cv-03770 (S.D.N.Y. filed May 16, 2024), alleging that an AI voice company created and commercialized unauthorized voice clones trained on the plaintiffs' recordings. In August 2024, additional voice actors filed *Vacker v. Eleven Labs, Inc.*, No. 1:24-cv-00987 (D. Del. filed Aug. 29, 2024), asserting analogous claims against another AI voice company. Microsoft itself was a prior BIPA defendant in *Vance v. Microsoft Corp.*, No. 2:20-cv-01082 (W.D. Wash.). And on February 5, 2026, while Microsoft continued to deploy and expand its voice products, Illinois residents filed *Basich et al. v. Microsoft Corp.*, No. 2:26-cv-00422 (W.D. Wash.), alleging that Microsoft Teams' speaker-diarization and transcription features capture and analyze voiceprints from meeting participants in violation of BIPA. By the time Microsoft expanded Custom Neural Voice into Dynamics 365 Contact Center in March 2026 and made MAI-Voice-1 generally available in the Microsoft Foundry Model Catalog by April 2026, AI voice companies were active defendants in litigation alleging the same conduct: building commercial voice synthesis products on voice recordings ingested without the speakers' consent. Microsoft was itself one of those defendants.

110. Microsoft is one of the largest publicly traded technology companies in the world, with a global privacy and compliance apparatus that includes a Chief Privacy Officer, dedicated privacy teams, an annual

Responsible AI Standard, and an annual Responsible AI Transparency Report. Microsoft's own product and research documentation describes its voice synthesis pipeline in biometric terms, as alleged at ¶¶ 33-41. A company that publicly acknowledges the biometric implications of its technology in its own documentation, while proceeding without implementing the notice, consent, retention, and destruction measures that Illinois law requires, has, at a minimum, acted with reckless disregard.

111. Microsoft cannot credibly plead ignorance that its voice synthesis pipeline extracts and stores speaker-specific representations capable of identifying the source speakers. Microsoft built the pipeline, published the research, and shipped the products on that very capability.

112. Microsoft has, moreover, confirmed its understanding that voice synthesis training implicates the legal rights of source speakers in two distinct ways: first, by building safeguards against the downstream misuse of its own voice synthesis products; second, by placing on its enterprise customers the upstream-rights obligation that Microsoft failed to discharge with respect to its own training data.

113. Microsoft built each of the downstream safeguards described at ¶¶ 42-48 because Microsoft understood that voice-AI products require speaker consent. Each safeguard operates at the downstream end of Microsoft's commercial pipeline, constraining how end users use Microsoft's products. None addresses the upstream issue: that Microsoft's foundational voice models were trained on voiceprints extracted without the knowledge or consent of the

speakers whose recordings were ingested. The selective application of consent infrastructure across Microsoft's product line is itself the willfulness fact. Microsoft's knowledge that voice-AI products require speaker consent, proven by the downstream infrastructure it built for some products and the research-only status it preserved for others, applies with equal force to the foundational training pipeline for which Microsoft built no consent infrastructure at all.

114. Microsoft's contractual allocation of biometric-compliance obligations to its customers, alleged at ¶ 47, and Microsoft's instruction to customers that they must "obtain explicit written permission from voice talent," alleged at ¶ 40, are Microsoft's own acknowledgments that the commercial use of an individual's voice to generate synthetic speech requires that individual's consent. Microsoft imposed those obligations on its downstream enterprise customers. Microsoft did not discharge the same obligation, in symmetric form, with respect to its own ingestion of training data.

115. The pattern of Microsoft's conduct between 2014 and 2026 reflects a deliberate, contemporaneous choice. Microsoft built and publicly deployed customer-facing consent and rights-acknowledgment infrastructure for its voice products. It built none of the corresponding upstream consent infrastructure for the speakers whose voices populated the training corpora — corpora documented in Microsoft's own research disclosures as measured in tens and hundreds of thousands of hours, as alleged at ¶ 50. The same Microsoft Research organization, Azure AI Speech business, Microsoft AI division, and acquired Nuance Communications business that developed the

customer-facing consent infrastructure also developed the foundational voice models trained on those unconsented corpora. The same legal and privacy apparatus that drafted the Custom Neural Voice limited-access terms, the Personal Voice acknowledgment workflow, and the talent-disclosure documentation also presided over the absence of corresponding consent for non-user training audio.

116. Microsoft's allocation of biometric-compliance obligations to its customers is itself probative of awareness, and at the same time underscores Microsoft's own breach. As to the voiceprints Microsoft itself extracted from third-party voice recordings, including those of Plaintiffs and Class members, there is no customer in the chain to assume the allocated responsibility. Microsoft is the controller, the entity that collected the data, and the entity in possession of the data. The duty Microsoft's talent-disclosure documentation allocates to others falls, with respect to Plaintiffs' voiceprints, on Microsoft itself. No contractual allocation between Microsoft and its customers can run to, or relieve Microsoft of duties owed to, the non-user speakers from whose recordings Microsoft extracted voiceprints. Microsoft's obligation to Plaintiffs arises directly from BIPA and cannot be discharged by reassignment to downstream parties who have no operational relationship with Plaintiffs.

117. Microsoft's contemporaneous training-data acquisition practices across multiple foundational-model families, alleged at ¶¶ 53-57, confirm that Microsoft's noncompliance reflects a sustained institutional pattern of ingesting identifiable source material at an industrial scale without the consent of the

people whose work the source material represents. Microsoft operates the infrastructure required to do so for voice audio, has used it for the DeepSinger corpus, has provided no mechanism by which the speakers whose voices were ingested can verify, withdraw consent, or request deletion, and has chosen that path over the available alternative — licensed or consented voice data — that would have constrained the speed and scale of its voice AI commercialization.

118. Each of the violations alleged in this Complaint was therefore committed by Microsoft with knowledge of, or in reckless disregard for, BIPA's requirements. Plaintiffs are entitled to liquidated damages of $5,000 per violation under 740 ILCS 14/20(2), or, in the alternative, $1,000 per violation under 740 ILCS 14/20(1), recoverable on a per-person, per-subsection basis where multiple distinct provisions of § 15 are violated, consistent with the statute as amended by P.A. 103-769 effective August 2, 2024 and as construed in *Clay v. Union Pacific Railroad Co.*, No. 25-2185 (7th Cir. Apr. 1, 2026).

CLASS ACTION ALLEGATIONS

119. Plaintiffs bring this action individually and on behalf of all others similarly situated, pursuant to Federal Rules of Civil Procedure 23(b)(2) and 23(b)(3), as the following Class: All natural persons whose voice recordings were produced or recorded in Illinois, and from whose recordings Microsoft extracted, derived, or otherwise obtained voiceprints or biometric information in connection with the development, training, fine-tuning, evaluation, or operation of Microsoft's foundational voice synthesis models, or commercial voice products derived from those models, during the Class Period.

120. The Class Period runs from the earlier of (a) the date Microsoft first ingested any voice recording into the training pipeline for any of its foundational voice synthesis models, or (b) January 1, 2014, through the date of judgment in this action. Discovery will establish the operative start date of the Class Period.

121. Excluded from the Class are: (i) Microsoft Corporation and each of its parents, subsidiaries, affiliates, and controlled entities, including without limitation Nuance Communications, Inc., Microsoft International Holdings B.V., and Microsoft Ireland Operations Limited; (ii) all current and former officers and directors of Microsoft; (iii) Microsoft's employees, contractors, agents, and counsel; (iv) the Court, the Court's staff, and any jurors assigned to this action; (v) the immediate family members of any person excluded above; and (vi) any person who executed a written release authorizing Microsoft's use of their voice recordings to train, fine-tune, or operate Microsoft's foundational voice synthesis models, in compliance with 740 ILCS 14/15(b), including registered Custom Neural Voice voice talent who provided voice data pursuant to Microsoft's voice-talent acknowledgment workflow, and registered Personal Voice users who provided voice data pursuant to a written or recorded acknowledgment statement.

122. Plaintiffs reserve the right to amend or refine the Class definition based on facts learned through discovery. Nothing in the Class definition limits or disclaims claims or remedies available under any statute or theory asserted in this Complaint.

123. *Ascertainability.* Class membership is defined by objective criteria and can be determined from records that exist or will be produced in discovery. Whether a particular voice recording entered Microsoft's training pipeline is a binary factual question — the recording is either in the pipeline's ingestion logs or it is not — and the records that answer the question are within Microsoft's exclusive control. Microsoft's published documentation establishes the existence of these records. Microsoft's Speech Studio architecture organizes voice work into projects that contain version-controlled datasets, training runs, and model versions, with file-level identifiers and provenance tracking. Microsoft's data, privacy, and security documentation for Azure AI Speech describes data flow, processing, and retention for voice data within Microsoft's pipelines, and confirms that Microsoft "may process biometric voice signatures" from "randomized audio from the training dataset(s)" — a representation that necessarily implies dataset-level identification of the audio so processed. Microsoft's Speaker Recognition API documentation describes enrollment as the formation of a "unique voice signature" from extracted voice features, and Microsoft's Speech service release notes track dataset and model versions over time. On information and belief, the training-data manifests, ingestion logs, source-URL records, dataset-version records, and associated speaker- and file-level identifiers for the foundational voice models that power the prebuilt Azure neural voice catalog, Azure neural HD voices, Custom Neural Voice, Personal Voice, the Voice Live API, MAI-Voice-1, and Microsoft's research models — including the VALL-E and NaturalSpeech families — exist in Microsoft's

pipeline records and are within its exclusive control. Class membership can be further confirmed through publicly available distribution metadata for voice recordings (audiobook platform records, podcast directory records, streaming service catalogs, and broadcast archive records), through speaker-identification technology applied to Microsoft's training corpus, and through voice-matching analysis comparing Microsoft's voice model outputs against publicly available recordings of class members.

124. *Numerosity.* Joinder of all members of the Class is impracticable. Fed. R. Civ. P. 23(a)(1). Microsoft has publicly disclosed, in research papers describing models built within the same research enterprise that produced its commercial voice products, that VALL-E was trained on approximately 60,000 hours of English speech from over 7,000 unique speakers; that NaturalSpeech 2 was trained on approximately 44,000 hours of data; that NaturalSpeech 3 was trained on approximately 200,000 hours of data; and that DeepSinger was trained on singing-voice data "mined from the web" through a crawl-and-extract pipeline that crawled music websites, separated vocals from accompaniment, aligned lyrics, and filtered samples. The number of distinct individuals whose voice recordings were produced or recorded in Illinois and ingested into the training pipelines that produced Microsoft's commercial voice products plainly satisfies the threshold for numerosity. The Class includes not only the professional broadcast journalists, audiobook narrators, podcasters, voice actors, voiceover artists, and other voice professionals who have produced work in Illinois during the Class Period — themselves a population numbering

in the thousands — but also the interviewees, guests, panelists, witnesses, callers, public officials, and other persons whose voices were captured in Illinois-produced or Illinois-recorded broadcast, podcast, audiobook, archival, and other publicly distributed audio content, and whose recordings are accessible on the publicly distributed audio platforms from which Microsoft, on information and belief, sourced training data. On information and belief, the Class population numbers in the hundreds of thousands or more. The exact number is within Microsoft's exclusive control and will be established through discovery.

125. *Commonality.* Common questions of law and fact apply to every member of the Class. Fed. R. Civ. P. 23(a)(2). Microsoft did not engage in any individualized notice, consent, retention-policy disclosure, release, or biometric-data-protection process with respect to any non-user whose voice recordings were ingested into Microsoft's foundational voice training pipelines. Microsoft's conduct was uniform: the same training pipelines ingested the same categories of voice recordings under the same absent-consent posture, applied to all class members through the same automated and standardized process. The questions whether Microsoft complied with BIPA's notice, consent, retention-policy, profiting, dissemination, and privacy-protection requirements before and after extracting the voiceprints of class members can be answered classwide because Microsoft's compliance, or noncompliance, was identical as to every class member.

126. Common questions of law and fact include, without limitation:

(a) whether the computational representations of vocal characteristics that Microsoft extracts during voice model training, including representations Microsoft itself describes in its own product and research documentation as "biometric voice signatures," "voice signatures," and speaker embeddings, constitute "voiceprints" or "biometric information" within the meaning of 740 ILCS 14/10;

(b) whether Microsoft informed class members in writing that their biometric identifiers were being collected, of the specific purpose and duration of collection, and obtained a written release, as 740 ILCS 14/15(b) requires;

(c) whether Microsoft developed and made publicly available a written retention and destruction policy applicable to class members' biometric identifiers, as 740 ILCS 14/15(a) requires;

(d) whether Microsoft sold, leased, traded, or otherwise profited from class members' biometric identifiers in violation of 740 ILCS 14/15(c);

(e) whether Microsoft disclosed or disseminated class members' biometric identifiers to third parties — including affiliates, vendors, subprocessors, and service providers across the multiple jurisdictions in which Microsoft operates its global cloud business, and through Microsoft's acquired Nuance Communications speech-services business — without consent and outside any enumerated exception, in violation of 740 ILCS 14/15(d);

(f) whether Microsoft stored, transmitted, and protected class members' biometric identifiers using the reasonable standard of care required by 740 ILCS 14/15(e)(1) and in a manner equally or more protective than its

protection of other confidential and sensitive information as required by 740 ILCS 14/15(e)(2), including in light of (i) the asymmetry between Microsoft's robust safeguards for customer-uploaded acknowledgment recordings under Custom Neural Voice and Personal Voice limited-access controls and the absence of corresponding safeguards for the biometric identifiers Microsoft itself extracted from non-user training audio, and (ii) the asymmetry between Microsoft's protection of its own confidential and proprietary commercial information — including source code, enterprise customer data, and trade secrets — under closed-system controls, restricted access, encryption at rest and in transit, and customer-managed encryption keys, and the absence of corresponding protections for the biometric identifiers Microsoft extracted from non-user training audio;

(g) whether Microsoft's conduct was willful or reckless within the meaning of 740 ILCS 14/20(2);

(h) whether Microsoft used class members' voices and identities for commercial purposes without prior written consent in violation of 765 ILCS 1075/30(a);

(i) whether Microsoft's voice products generate, distribute, or make available unauthorized digital replicas within the meaning of 765 ILCS 1075/30(b), as amended by P.A. 103-836 effective January 1, 2025, and whether Microsoft materially contributes to or facilitates their distribution within the meaning of 765 ILCS 1075/30(d);

(j) whether Microsoft's conduct constitutes deceptive trade practices likely to cause confusion in violation of the Illinois Uniform Deceptive Trade Practices Act, 815 ILCS 510/2(a)(2) and 510/2(a)(3);

(k) whether Microsoft's conduct constitutes an unfair practice under the Illinois Consumer Fraud and Deceptive Business Practices Act, 815 ILCS 505/2, under the test articulated in *Robinson v. Toyota Motor Credit Corp.*, 201 Ill. 2d 403, 417–18 (2002);

(l) whether Microsoft was unjustly enriched by its unauthorized use of class members' voice data; and

(m) the appropriate measures of damages, restitution, and injunctive relief.

127. *Typicality.* The claims of the named Plaintiffs are typical of the claims of the Class. Fed. R. Civ. P. 23(a)(3). Each named Plaintiff produced voice recordings in Illinois — broadcast journalism, audio reporting, audiobook narration, podcast production, voice acting, and related professional voice work, recorded in Chicago broadcast facilities, podcast and audio production studios, and home recording studios in this District. Microsoft, on information and belief, ingested those recordings into the training pipelines for its foundational voice synthesis models and extracted voiceprints and/or biometric information from them without notice, consent, or written release. The voiceprints are now embedded in Microsoft's foundational voice models and commercial voice products alleged at ¶ 31— and reproduced in the audio those models generate. The legal theories asserted on behalf of the Class —

that Microsoft's extraction, retention, commercial exploitation, and dissemination of voiceprints from Illinois-recorded voice work without BIPA-compliant consent violates 740 ILCS 14/15(a)–(e), and that Microsoft's commercial use of voices and identities without consent violates the Illinois Right of Publicity Act — apply with equal force to each named Plaintiff and to every other class member. Each named Plaintiff has the same interest as every other class member in establishing Microsoft's liability and obtaining the relief sought in this Complaint.

128. *Adequacy.* Plaintiffs will fairly and adequately protect the interests of the Class. Fed. R. Civ. P. 23(a)(4). Plaintiffs' interests are aligned with, and not antagonistic to, the interests of the absent class members; each named Plaintiff has the same incentive as every other class member to maximize recovery and to obtain comprehensive injunctive relief addressing the unlawful extraction of voiceprints. Plaintiffs are represented by counsel experienced in complex class action litigation, privacy litigation, and BIPA litigation, with the resources to prosecute this action vigorously on behalf of the Class.

129. *Rule 23(b)(2) Certification.* Certification under Rule 23(b)(2) is appropriate because Microsoft has acted on grounds generally applicable to the Class, such that final injunctive and corresponding declaratory relief is appropriate as to the Class as a whole. Microsoft's training pipelines operated uniformly across every class member's voice recordings; Microsoft's failure to obtain BIPA-compliant consent was uniform; Microsoft's failure to publish a retention and destruction policy applicable to non-user training data subjects

was uniform; and Microsoft's continuing possession and commercial exploitation of class members' voiceprints in commercial models — and ongoing deployment of those voiceprints through the deployment channels alleged at ¶ 31, ¶ 68(a)-(h), and ¶ 8 — is uniform. Plaintiffs seek classwide injunctive relief under 740 ILCS 14/20, 815 ILCS 510/3 (IUDTPA, which authorizes only injunctive relief), and the equitable jurisdiction of this Court — including the destruction or retraining, without the unlawfully obtained voiceprints and biometric information, of the foundational voice synthesis models in which class members' voiceprints are encoded — that necessarily applies on the same terms to every class member.

130. *Rule 23(b)(3) Certification.* Certification under Rule 23(b)(3) is appropriate for the Class on the damages and restitutionary claims asserted in this Complaint, including the claims under BIPA, IRPA, ICFA, IUDTPA, and Illinois common law for unjust enrichment, because common questions of law and fact predominate over questions affecting only individual members of the Class, and a class action is superior to other available methods for fairly and efficiently adjudicating the controversy.

131. *Predominance.* The questions that drive this litigation are common to the Class and predominate over individual questions. Whether Microsoft's voice model training pipelines extract voiceprints within the meaning of BIPA is a common technical question with a common answer — Microsoft's own data, privacy, and security documentation for Azure AI Speech describes the pipelines as processing "biometric voice signatures" from "randomized audio

from the training dataset(s)"; Microsoft's Speaker Recognition API documentation describes enrollment as "extract[ing] voice features to form a unique voice signature"; and Microsoft's VALL-E paper describes the system as preserving "the speaker's emotion and acoustic environment" from input audio. Whether Microsoft complied with BIPA's notice, consent, retention, profiting, dissemination, and privacy-protection requirements is a common legal question with a common answer — Microsoft did not, with respect to any non-user whose voice recordings were ingested into the training pipelines. Whether Microsoft's conduct was willful or reckless turns on Microsoft's institutional knowledge, decision-making, and conduct, all of which are common to the Class. The principal individual question — whether a specific class member's voice recordings entered the training pipelines — is binary and resolvable from Microsoft's own records, including the training-data manifests, ingestion logs, pipeline metadata, and source-identifier records in Microsoft's exclusive control. Individual damages calculations under BIPA's per-person, per-subsection liquidated-damages framework, 740 ILCS 14/20, as amended and as construed in *Clay*, do not predominate over the common liability questions because the per-violation amounts are statutorily fixed and do not require individualized proof of actual damages.

132. *Superiority.* A class action is the superior method for adjudicating these claims:

(a) *Class members' interest in individual control.* Class members are voice professionals — journalists, audiobook narrators, podcasters, voiceover

artists — and incidental speakers whose voiceprints were extracted without their knowledge. Many class members are unaware their biometric identifiers were ever taken. Even those who become aware face the prospect of individual litigation against one of the largest publicly traded technology companies in the world, with statutory damages amounts that, while meaningful in the aggregate, are likely too modest in individual cases to justify the cost and burden of independent representation. A class action is the only realistic vehicle for redressing the violations alleged in this Complaint.

(b) *Existing related litigation.* Plaintiffs are unaware of any other action asserting these claims against Microsoft on behalf of persons whose voice recordings were produced or recorded in Illinois and used to train Microsoft's foundational voice synthesis models. One related action concerning Microsoft's biometric voice practices is currently pending in the Western District of Washington: *Basich et al. v. Microsoft Corp.*, No. 2:26-cv-00422 (W.D. Wash. filed Feb. 5, 2026), in which Illinois residents allege that Microsoft Teams' speaker-diarization and transcription features capture and analyze voiceprints from meeting participants in violation of BIPA. *Basich* arises from a different Microsoft product feature unrelated to foundational voice synthesis models, Microsoft Teams' real-time voice attribution. The two actions involve different products, data sources, biometric-processing pipelines, alleged uses, proof, and have non-overlapping class definitions. Microsoft was previously a BIPA defendant in *Vance v. Microsoft Corp.*, No. 2:20-cv-01082 (W.D. Wash.). Neither *Basich* nor *Vance* asserts claims on behalf of voice professionals —

broadcast journalists, podcasters, audiobook narrators, voiceover artists, and other speakers — whose voice recordings were ingested into Microsoft's foundational voice synthesis training pipelines.

(c) *Desirability of concentration in this forum.* Concentrating this litigation in this District is appropriate. Plaintiffs are Illinois residents whose recordings were produced or recorded in Illinois. The claims arise under Illinois statutes. The injuries were suffered in Illinois. Microsoft conducts substantial commercial business in Illinois, including through enterprise licensing of its voice AI products to Illinois-based customers and through office locations and operational presence in the City of Chicago. This Court is well-suited to adjudicate BIPA and other Illinois statutory claims arising from Microsoft's commercial conduct in Illinois.

(d) *Manageability.* The case is manageable as a class action. Microsoft's conduct was automated, uniform, and standardized; the common questions identified above are susceptible to common proof; class membership can be determined from Microsoft's records, supplemented as needed by publicly available distribution metadata and voice-matching analysis; and BIPA's per-person, per-subsection liquidated-damages framework eliminates the need for individualized damages calculations on the principal claim. No unusual management difficulties are anticipated.

133. To the extent any portion of the Class Period predates the limitations period applicable to any claim asserted in this action, Plaintiffs allege that the limitations periods are equitably tolled by Microsoft’s

concealment of its training-data sources, by Microsoft's failure to provide any notice of its collection of biometric identifiers, by Plaintiffs' inability through reasonable diligence to discover that Microsoft had ingested their recordings into its training pipelines, and by the continuing nature of Microsoft's violations. Independently, each retention of the unlawfully obtained biometric data, each operation of the foundational audio models in which the unlawfully obtained biometric data is encoded, each disclosure or transmission of that biometric data, and each public release of model parameters encoding that biometric data is a separate violation of BIPA under *Cothron v. White Castle System, Inc.*, 2023 IL 128004 (Ill. 2023), each accruing a separate limitations period from the date of the discrete violative act.

CLAIMS FOR RELIEF

Count I

Violation of the Illinois Biometric Information Privacy Act, 740 ILCS 14/15(b)

Brought on behalf of the Class

134. Plaintiffs reallege and incorporate by reference all allegations in this complaint.

135. Plaintiffs bring this Count individually and on behalf of the Class.

136. BIPA defines "biometric identifier" to include a "voiceprint" and "biometric information" to include "any information, regardless of how it is captured, converted, stored, or shared, based on an individual's biometric identifier used to identify an individual." 740 ILCS 14/10. Section 15(b) prohibits a private entity from collecting, capturing, purchasing, receiving

through trade, or otherwise obtaining a person's biometric identifier or biometric information unless the entity first informs the subject in writing that biometric data is being collected or stored, informs the subject in writing of the specific purpose and length of term of collection, and receives a written release executed by the subject. 740 ILCS 14/15(b).

137. Microsoft is a "private entity" within the meaning of BIPA. 740 ILCS 14/10.

138. Microsoft collected, captured, and otherwise obtained voiceprints and biometric information from the voice recordings of Plaintiffs and Class members by ingesting their voice recordings into its training pipelines and extracting from those recordings computational representations capable of identifying the speakers, as alleged in detail at ¶¶ 33-41. The resulting representations, described in Microsoft's product and research documentation as "biometric voice signatures," "voice signatures," speaker embeddings, audio tokens, and factorized timbre representations, are voiceprints and biometric information within the meaning of BIPA.

139. Microsoft did not, before extracting Plaintiffs' or Class members' voiceprints, inform any Plaintiff or Class member in writing that their biometric identifiers were being collected or stored, did not inform any of them in writing of the specific purpose or length of term of collection, and did not receive a written release executed by any of them. Microsoft obtained no consent of any kind, in any form, from any Plaintiff or Class member.

140. Microsoft's violations of § 15(b) were intentional or reckless, as alleged at ¶¶ 107-118. In the alternative, Microsoft's violations were negligent.

141. Plaintiffs and the Class seek all relief available under 740 ILCS 14/20, including, for each Class member, the greater of liquidated damages or actual damages on a per-person, per-subsection basis consistent with the statute as amended by P.A. 103-769 effective August 2, 2024 and as construed in *Clay v. Union Pacific Railroad Co.*, No. 25-2185 (7th Cir. Apr. 1, 2026), in the amount of $1,000 per negligent violation or $5,000 per intentional or reckless violation; injunctive relief; and reasonable attorneys' fees, costs, and any other relief the Court deems just and proper.

Count II

Violation of the Illinois Biometric Information Privacy Act, 740 ILCS 14/15(a)

Brought on behalf of the Class

142. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

143. Plaintiffs bring this Count individually and on behalf of the Class.

144. Section 15(a) of BIPA requires a private entity in possession of biometric identifiers to develop a written policy, made available to the public, establishing a retention schedule and guidelines for permanently destroying biometric identifiers when the initial purpose for collecting or obtaining such identifiers has been satisfied or within three years of the individual's last interaction with the private entity, whichever occurs first. 740 ILCS 14/15(a). Because Class members never interacted with Microsoft in connection with the

collection of their biometric data, the operative destruction prong for Class members is that biometric identifiers be destroyed when the initial purpose for their collection has been satisfied.

145. Microsoft has been, and remains, in possession of voiceprints and biometric information extracted from recordings of Plaintiffs and Class members, encoded in the parameters of the foundational voice synthesis models that power Microsoft's commercial voice products.

146. Microsoft has not developed and made publicly available a retention and destruction policy applicable to voiceprints extracted from non-user training data and embedded in the parameters of Microsoft's foundational voice synthesis models, as alleged at ¶¶ 49-57. Microsoft's Azure AI Speech privacy documentation addresses customer-provided acknowledgment recordings used for Azure AI Speaker Verification ("retained only as long as necessary to perform speaker verification"); it does not address voiceprints extracted from non-user training audio. Microsoft's general Privacy Policy does not address voice data used in AI model training and does not apply to non-users like Plaintiffs. Microsoft has not provided, and on information and belief does not maintain, any mechanism by which Plaintiffs or other non-user training-data subjects can request access to, correction of, or deletion of their biometric data.

147. Microsoft's violations of § 15(a) were intentional or reckless, as alleged at ¶¶ 107-118. In the alternative, Microsoft's violations were negligent.

148. Plaintiffs and the Class seek all relief available under 740 ILCS 14/20 on the terms set forth in Count I.

Count III

Violation of the Illinois Biometric Information Privacy Act, 740 ILCS 14/15(c)

Brought on behalf of the Class

149. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

150. Plaintiffs bring this Count individually and on behalf of the Class.

151. Section 15(c) of BIPA provides that no private entity in possession of a biometric identifier may sell, lease, trade, or otherwise profit from a person's biometric identifier. 740 ILCS 14/15(c). The phrase "otherwise profit from" is a statutory catch-all that extends beyond the enumerated forms of selling, leasing, and trading.

152. Microsoft has profited, and continues to profit, from Plaintiffs' and Class members' voiceprints by using those voiceprints to develop, train, and commercially operate the foundational voice synthesis models that power Microsoft's commercial voice products, and by monetizing those products through each of the channels alleged at ¶ 68(a)-(h), which are incorporated here. The voice quality, expressiveness, multilingual capability, speaker similarity, and realism that Microsoft sells through those channels exist because of the voiceprints encoded in Microsoft's foundational voice synthesis models. Microsoft's commercial exploitation of Plaintiffs' and Class members' voiceprints does not fall within any exception to § 15(c).

153. Microsoft's violations of § 15(c) were intentional or reckless, as alleged at ¶¶ 107-118. In the alternative, Microsoft's violations were negligent.

154. Plaintiffs and the Class seek all relief available under 740 ILCS 14/20 on the terms set forth in Count I.

Count IV

Violation of the Illinois Biometric Information Privacy Act, 740 ILCS 14/15(d)

Brought on behalf of the Class

155. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

156. Plaintiffs bring this Count individually and on behalf of the Class.

157. Section 15(d) of BIPA provides that no private entity in possession of a biometric identifier may disclose, redisclose, or otherwise disseminate a person's biometric identifier unless an enumerated exception applies. 740 ILCS 14/15(d).

158. Microsoft has disclosed, redisclosed, and otherwise disseminated Plaintiffs' and Class members' voiceprints in three categories of conduct, each independently sufficient to violate § 15(d) and as alleged at ¶¶ 70-74.

159. *First*, Microsoft has, on information and belief, transferred the foundational voice synthesis models — and the voiceprints encoded within them — among Microsoft's affiliated corporate entities, including Nuance Communications, Inc., Microsoft International Holdings B.V., and Microsoft Ireland Operations Limited, in the ordinary course of Microsoft's global voice AI business.

160. *Second*, Microsoft has, on information and belief, transmitted the foundational voice synthesis models — and the voiceprints encoded within them — to Azure cloud-infrastructure subprocessors, content-delivery-network operators, hardware partners, and other vendors in the ordinary course of training, evaluating, deploying, and operating its commercial voice products. The delivery of Microsoft's voice synthesis services to enterprise and consumer customers, including Illinois customers, necessarily involves transmitting model parameters and inference outputs over that infrastructure.

161. *Third*, to the extent Microsoft Research has published reference implementations or model artifacts associated with the VALL-E, NaturalSpeech, DeepSinger, or other research model families on GitHub or similar platforms that enable the reconstruction of or downstream training on model parameters trained on Plaintiffs' voice recordings, those releases constitute additional dissemination of the biometric data encoded in those parameters.

162. Plaintiffs and Class members did not consent to any of the disseminations alleged in this Count. No enumerated exception under § 15(d) applies.

163. Microsoft's violations of § 15(d) were intentional or reckless, as alleged at ¶¶ 107-118. In the alternative, Microsoft's violations were negligent.

164. Plaintiffs and the Class seek all relief available under 740 ILCS 14/20 on the terms set forth in Count I.

Count V

Violation of the Illinois Biometric Information Privacy Act, 740 ILCS 14/15(e)

Brought on behalf of the Class

165. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

166. Plaintiffs bring this Count individually and on behalf of the Class.

167. Section 15(e) of BIPA requires a private entity in possession of biometric identifiers to store, transmit, and protect the identifiers from disclosure (1) using the reasonable standard of care within the entity's industry, and (2) in a manner that is the same as or more protective than the manner in which the entity stores, transmits, and protects other confidential and sensitive information. 740 ILCS 14/15(e). Section 15(e)(2) imposes an asymmetry test: a private entity satisfies it only if its protection of biometric data is at least equal to its protection of its own other confidential and sensitive information. A generally adequate security posture does not satisfy § 15(e)(2) if the entity protects its other confidential information with greater care than it protects biometric data.

168. Microsoft fails the § 15(e)(2) asymmetry test on two independent grounds.

169. *First asymmetry — confidential customer and proprietary data vs. non-user biometric data.* Microsoft contemporaneously protects its own confidential and proprietary information, and the confidential information of its commercial customers, under closed-system controls that are among the most

rigorous in the technology industry. Microsoft maintains the source code for its core commercial products — including Windows, Microsoft Office, Microsoft 365, Microsoft Teams, Azure, GitHub Enterprise, and Microsoft's AI training stack — as closed-source proprietary information under strict access controls, internal classification systems, encryption at rest and in transit, and information-security regimes that are subject to extensive internal audit. Microsoft offers its enterprise customers FedRAMP High and DoD Impact Level 5 authorized Azure environments, SOC 1/2/3 attestations, ISO 27001/27017/27018 certifications, customer-managed encryption keys, Azure Confidential Computing, and contractual data-processing protections through the Microsoft Customer Agreement and the Microsoft Products and Services Data Protection Addendum. Microsoft's published Responsible AI Standard and Responsible AI Transparency Report describe an extensive internal review apparatus for high-risk AI deployments. Microsoft protects customer-uploaded acknowledgment recordings for Custom Neural Voice and Personal Voice under Limited Access registration, speaker verification, and explicit retention statements stating that those recordings are "retained only as long as necessary to perform speaker verification."

170. Microsoft's protection of non-user biometric data — the voiceprints extracted from Plaintiffs' and Class members' voice recordings and encoded in the parameters of Microsoft's commercial voice synthesis models — is the inverse of those protections. Microsoft has not published a retention or destruction policy applicable to non-user training-data biometric identifiers.

Microsoft has not provided a mechanism by which non-users can request access to, correction of, or deletion of their biometric data. Microsoft has not disclosed the sources of the voice recordings used to train its prebuilt Azure neural voices, Azure neural HD voices, MAI-Voice-1, or the foundational models underlying Custom Neural Voice and Personal Voice. The biometric data is encoded permanently in the parameters of models continuously deployed across Azure AI Speech, Microsoft Foundry, Microsoft Copilot, Microsoft Teams, Dynamics 365 Contact Center, the Voice Live API, and acquired Nuance speech offerings. As alleged at ¶¶ 49-57, Microsoft has, by its own published disclosures, refused to identify or destroy this data, and has built none of the safeguards it routinely builds for the confidential information of its commercial customers.

171. The asymmetry violates § 15(e)(2). Microsoft protects its source code, its enterprise customer data, its commercial trade secrets, its employee personnel files, its internal financial records, and its customer-uploaded acknowledgment recordings under materially more protective regimes than it applies to the biometric identifiers of Plaintiffs and Class members. The biometric data is the least protected category of sensitive data Microsoft holds — not because the data is unimportant, but because the Plaintiffs and Class members are non-parties to whom Microsoft chose not to disclose the collection.

172. *Second asymmetry — customer-uploaded biometric data vs. non-user training-audio biometric data.* Microsoft has demonstrated, through the

safeguards alleged at ¶¶ 42-48, that it possesses the operational capability to protect biometric voice data, the contractual infrastructure to define retention and destruction terms, and the technical infrastructure to enforce access and deletion. Microsoft deploys that capability for the speakers with whom it has a commercial relationship and withholds it from the non-user speakers whose recordings Microsoft itself ingested into the training pipeline.

173. The asymmetry between Microsoft's robust protection of customer-uploaded biometric data and the absence of any equivalent protection for non-user biometric data extracted from training audio is itself a § 15(e)(2) violation. A private entity that knows how to protect biometric data, and demonstrably does so for some categories of data subjects, but withholds equivalent protection from another category of data subjects to whom the same statutory duty runs, fails § 15(e)(2)'s requirement that biometric data be protected "in a manner that is the same as or more protective than" the entity's protection of other confidential and sensitive information.

174. Independently, Microsoft fails the § 15(e)(1) reasonable-standard-of-care test as to voiceprints extracted from non-user training data. The reasonable standard of care within the AI and machine learning industry for the handling of biometric training data includes documented provenance for training data, access controls, audit logging, encryption, a documented retention and destruction schedule, and a mechanism through which data subjects may request access to, correction of, or deletion of their biometric data. Microsoft has not published documented provenance for the bulk of the

voice training data used to build its commercial voice synthesis models, as alleged at ¶¶ 49-57; has not published any access-control protocol, audit-logging policy, encryption standard, or retention and destruction schedule specifically applicable to non-user voice training data or to the biometric characteristics encoded in its voice models; and has not provided any mechanism through which non-user data subjects may request access to, correction of, or deletion of their biometric data from Microsoft's training datasets, model parameters, or commercial inference pipelines. A private entity cannot adequately protect biometric data belonging to individuals to whom it has never disclosed possession of their data, because those individuals have no ability to monitor, verify, or challenge how their data is stored, transmitted, or protected. The reasonable industry standard of care for biometric data presupposes some mechanism by which the data subject can verify the integrity of the protection; Microsoft has provided none.

175. Microsoft's violations of § 15(e) were intentional or reckless, as alleged at ¶¶ 107-118. In the alternative, Microsoft's violations were negligent.

176. Plaintiffs and the Class seek all relief available under 740 ILCS 14/20 on the terms set forth in Count I.

Count VI

Violation of the Illinois Right of Publicity Act,
765 ILCS 1075/1 et seq.

Brought on behalf of the Class

177. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

178. Plaintiffs bring this Count individually and on behalf of the Class.

179. The Illinois Right of Publicity Act protects an individual's right to control commercial use of their identity. 765 ILCS 1075/30(a) provides that a person may not use an individual's identity for commercial purposes during the individual's lifetime without having obtained previous written consent from the individual. IRPA defines "identity" to include "any attribute of an individual that serves to identify that individual to an ordinary, reasonable viewer or listener," and expressly enumerates "voice" among the protected attributes. 765 ILCS 1075/5.

180. Each Plaintiff and each Class member is an individual whose distinctive voice — including timbre, tone, cadence, phrasing, accent, and stylistic vocal expression — is part of the individual's identity within the meaning of IRPA.

181. Microsoft used Plaintiffs' and Class members' identities for commercial purposes within the meaning of IRPA, by extracting and modeling the distinctive vocal characteristics embodied in their recordings and using those characteristics to develop, train, and operate the commercial voice synthesis products that Microsoft monetizes through the integrated commercial chain alleged at ¶¶ 68(a)-(h). Microsoft holds out its products' ability to generate realistic, expressive, and human-sounding voices — and, in the case of Custom Neural Voice and Personal Voice, their ability to synthesize speech that sounds "nearly identical to" a source human voice — as a core commercial feature of those products, a capability built from the identities of the

individuals whose recordings were used to train the foundational models. Microsoft did not obtain written consent from any Plaintiff or Class member to use their identity, including their voice, for any commercial purpose.

182. Independently, the Illinois Right of Publicity Act, as amended by P.A. 103-836 effective January 1, 2025, prohibits knowingly distributing, transmitting, or making available to the general public a sound recording or audiovisual work containing an "unauthorized digital replica" of an individual without the individual's consent or the consent of an authorized representative. 765 ILCS 1075/30(b). The amended IRPA further imposes liability on any person who "materially contributes to, induces, or facilitates" such distribution. 765 ILCS 1075/30(d). On information and belief, Microsoft's commercial voice products — including the prebuilt Azure neural voice catalog, Azure neural HD voices, Custom Neural Voice, Personal Voice, MAI-Voice-1, and the Voice Live API — generate, distribute, and make available to the public voice outputs that constitute unauthorized digital replicas of Plaintiffs and Class members within the meaning of the amended IRPA. Microsoft materially contributes to and facilitates the distribution of such replicas through its integrated commercial channels, including Azure AI Speech, Microsoft Foundry, Microsoft Copilot, Microsoft Teams, Dynamics 365 Contact Center, and the deployment of MAI-Voice-1 in Microsoft Copilot Daily and Copilot Podcasts.

183. Microsoft's violations of IRPA were willful and knowing, as alleged at ¶¶ 107-118. Microsoft's customer-facing safeguards and contractual allocations, alleged at ¶¶ 40, 45, and 47, are Microsoft's own acknowledgments

that the commercial use of an individual's voice to generate synthetic speech requires that individual's consent. Microsoft imposed those obligations on its downstream customers and end users. Microsoft chose not to apply equivalent consent infrastructure to Plaintiffs and Class members, whose voices Microsoft itself extracted to train the foundational models.

184. Plaintiffs and the Class have suffered concrete injury from Microsoft's IRPA violations, including loss of control over the commercial use of their identities and voices, dilution and commodification of their distinctive voices, and economic harms including the diversion of licensing value and the diminished demand for authentic vocal performances in the markets where Plaintiffs and Class members earn their livelihoods, as alleged at ¶¶ 75-82.

185. Plaintiffs and the Class seek all relief available under 765 ILCS 1075/40, including actual damages, profits attributable to Microsoft's unauthorized use of Plaintiffs' and Class members' identities, the statutory minimum of $1,000 per violation under 765 ILCS 1075/40 where actual damages are below that amount, punitive damages, injunctive relief, attorneys' fees pursuant to 765 ILCS 1075/55, and such other relief as the Court deems just and proper.

Count VII

Violation of the Illinois Consumer Fraud and Deceptive Business Practices Act, 815 ILCS 505/1 et seq.

Brought on behalf of the Class

186. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

187. Plaintiffs bring this Count individually and on behalf of the Class.

188. The Illinois Consumer Fraud and Deceptive Business Practices Act prohibits "unfair or deceptive acts or practices … in the conduct of any trade or commerce." 815 ILCS 505/2. ICFA's prohibition of unfair practices is broader than its prohibition of deceptive practices and reaches conduct that, even absent affirmative misrepresentation, offends Illinois public policy, is immoral, unethical, oppressive, or unscrupulous, or causes substantial injury to consumers that consumers could not reasonably have avoided. *Robinson v. Toyota Motor Credit Corp.*, 201 Ill. 2d 403, 417-18 (2002).

189. Microsoft engaged in trade and commerce in Illinois within the meaning of ICFA by marketing, offering, and distributing voice synthesis services to Illinois customers through the integrated commercial chain alleged at ¶¶ 68(a)-(h), and by collecting, extracting, and commercially exploiting the biometric identifiers of Illinois persons in connection with that commerce.

190. Microsoft's conduct is unfair within the meaning of ICFA. Microsoft's collection and commercial exploitation of the biometric identifiers of Illinois persons without their knowledge or consent offends Illinois public policy as expressed in BIPA — the Illinois statute enacted specifically to address the unauthorized collection of biometric identifiers, 740 ILCS 14/5 — and as expressed in IRPA — the Illinois statute enacted specifically to address the unauthorized commercial use of personal identity, 765 ILCS 1075/30. Microsoft's conduct caused substantial injury to Plaintiffs and Class members, including the diversion of licensing income they would otherwise have earned

for the use of their voices and the displacement of their professional voice work in the markets where they earn their livelihoods, as alleged at ¶¶ 75-82. Plaintiffs and Class members could not reasonably have avoided this injury because they had no knowledge that Microsoft was collecting their biometric data.

191. Microsoft's conduct is also independently deceptive within the meaning of ICFA. Microsoft's data, privacy, and security documentation for Azure AI Speech uses hedging language ("may process biometric voice signatures") to obscure the nature and extent of the biometric processing Microsoft performs during training. Microsoft's disclosure-for-voice-and-avatar-talent documentation, alleged at ¶ 47, attempts to push Microsoft's own statutory obligations onto downstream customers who have no operational relationship with Plaintiffs and cannot satisfy those obligations as to Plaintiffs. Microsoft has maintained total opacity about the sources of its prebuilt voice training data, the Azure neural HD voices training data, and the MAI-Voice-1 training data while commercially exploiting that data at scale across its products. The omissions and the obligation-shifting are materially misleading to consumers and to the public, including to voice professionals whose work Microsoft's products commercially substitute for, about whether Microsoft's voice products are built from biometric data collected with appropriate consent.

192. Microsoft's unfair and deceptive conduct proximately caused actual injury to Plaintiffs and Class members, including lost and diminished licensing income, suppressed voiceover and narration rates, diverted

opportunities, and loss of control over their biometric data and professional identities. These injuries flow directly from Microsoft's decision to collect and commercially exploit Plaintiffs' biometric data without authorization, and not merely from the existence of competing AI products.

193. Microsoft's conduct was willful, knowing, and in reckless disregard of the rights and interests of Plaintiffs and Class members, as alleged at ¶¶ 107-118.

194. Plaintiffs and the Class seek all relief available under ICFA, including actual damages, punitive damages under 815 ILCS 505/10a(c), injunctive relief, attorneys' fees and costs, and such other relief as the Court deems just and proper.

Count VIII

Violation of the Illinois Uniform Deceptive Trade Practices Act, 815 ILCS 510/1 et seq.

Brought on behalf of the Class (injunctive relief only)

195. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

196. Plaintiffs bring this Count individually and on behalf of the Class, seeking injunctive relief under 815 ILCS 510/3.

197. The Illinois Uniform Deceptive Trade Practices Act prohibits a person from engaging in conduct that, in the course of business, "causes likelihood of confusion or of misunderstanding as to the source, sponsorship, approval, or certification of goods or services," 815 ILCS 510/2(a)(2), or "causes

likelihood of confusion or of misunderstanding as to affiliation, connection, or association with or certification by another," 815 ILCS 510/2(a)(3).

198. Microsoft's commercial voice products generate voice outputs that sound like real human speakers, including voice outputs that, on information and belief, replicate or closely simulate the distinctive vocal characteristics of Plaintiffs and Class members. Once generated, these voice outputs can be downloaded, shared, embedded in Microsoft Copilot Daily and Copilot Podcasts experiences, integrated into Dynamics 365 Contact Center deployments, and otherwise commercially exploited by Microsoft's customers and end users without consumer-facing disclosure that the voice was AI-generated, that the model that generated the voice was built using voiceprints collected without consent, or that the individual whose vocal characteristics are reproduced has not authorized the use. The absence of disclosure creates likelihood of confusion or misunderstanding about whether real persons created, endorsed, sponsored, approved, or have any affiliation with the AI-generated voice content.

199. Plaintiffs and Class members are persons likely to be damaged by Microsoft's deceptive practices within the meaning of 815 ILCS 510/3. Microsoft's conduct diverts demand from licensed human voice performances and impairs source attribution and authorization-status disclosure in the voice services markets where Plaintiffs and Class members earn their livelihoods. Actual confusion and actual damages need not be shown to obtain injunctive relief.

200. Plaintiffs and the Class seek preliminary and permanent injunctive relief under 815 ILCS 510/3, including injunctive relief requiring Microsoft to provide adequate consumer-facing disclosure that voice outputs are AI-generated and that the individuals whose vocal characteristics are reproduced did not authorize the use. If the Court finds that Microsoft has willfully engaged in deceptive trade practices within the meaning of 815 ILCS 510/3, Plaintiffs also seek reasonable attorneys' fees.

Count IX

Unjust Enrichment (Illinois Common Law)

Brought on behalf of the Class

201. Plaintiffs reallege and incorporate by reference all allegations in this Complaint.

202. Plaintiffs bring this Count individually and on behalf of the Class. This Count is pled in the alternative pursuant to Fed. R. Civ. P. 8(d)(2)–(3). Plaintiffs do not seek duplicative recovery.

203. Under Illinois common law, a defendant is liable for unjust enrichment when the defendant has unjustly retained a benefit to the plaintiff's detriment, and the defendant's retention of that benefit violates fundamental principles of justice, equity, and good conscience.

204. Microsoft obtained substantial benefits from Plaintiffs' and Class members' voice recordings, voiceprints, and identity attributes without permission and without compensation. These benefits include the avoided costs of licensing or obtaining consent for the voice recordings used to train

Microsoft's foundational voice synthesis models; the product capability and competitive advantage Microsoft captured by training its models on a diverse corpus of professional human voices, including the voices of Plaintiffs and Class members; and the revenue Microsoft generates and continues to generate through the commercial exploitation of those models across the integrated commercial chain alleged at ¶¶ 68(a)-(h).

205. Microsoft obtained these benefits at Plaintiffs' and Class members' expense. Plaintiffs and Class members invested time, talent, training, and resources to develop their voices and create professional recordings. Microsoft's unauthorized extraction of voiceprints from those recordings diverted economic value from Plaintiffs and Class members to Microsoft, and Microsoft's commercial deployment of products built on those voiceprints continues to compete with and displace Plaintiffs and Class members in the markets where they earn their livelihoods, as alleged at ¶¶ 75-82.

206. Microsoft's retention of these benefits is unjust. Microsoft's conduct violated the Illinois statutory protections for biometric data (BIPA) and for personal identity (IRPA), each of which expresses the public policy of Illinois that the unauthorized commercial extraction of biometric and identity-related personal attributes is unlawful. Microsoft's institutional practice confirms the inequity of its retention: Microsoft built and deployed customer-facing rights-acknowledgment infrastructure (Custom Neural Voice, Personal Voice, Enterprise AI Services Code of Conduct, talent-disclosure documentation, the VALL-E 2 and ELaTE research-only retention) for the speakers with whom it

has a commercial relationship, but did not apply equivalent infrastructure to the voices of Plaintiffs and Class members from whom Microsoft itself took voiceprints without consent. Microsoft could have pursued a lawful licensing path with respect to Plaintiffs and Class members; it chose not to.

207. Plaintiffs and the Class seek restitution of the benefits Microsoft has unjustly retained, including disgorgement of profits Microsoft earned from the unauthorized exploitation of Plaintiffs' and Class members' voiceprints and identities, an accounting of those benefits, and such other equitable relief as the Court deems just and proper.

PRAYER FOR RELIEF

WHEREFORE, Plaintiffs, individually and on behalf of all others similarly situated, respectfully request that this Court enter judgment against Defendant Microsoft Corporation and award the following relief:

A. *Injunctive Relief*. Permanent and, where appropriate, preliminary injunctive relief requiring Microsoft to:

(1) Cease the collection, capture, purchase, receipt through trade, or other obtaining of biometric identifiers from voice recordings produced or recorded in Illinois without first providing the written notice, disclosure of specific purpose and duration of collection, and written release that 740 ILCS 14/15(b) requires;

(2) Cease the commercial use, sale, lease, trade, profiting from, or dissemination of voiceprints and biometric information of Plaintiffs and

Class members that Microsoft has already collected without BIPA-compliant consent;

(3) Identify and disclose, by name or other identifying information, the sources of all voice training data used to develop, train, fine-tune, evaluate, or operate Microsoft's foundational voice synthesis models, including without limitation the prebuilt Azure neural voice catalog, the Azure neural HD voices, Custom Neural Voice, Personal Voice, the Voice Live API, MAI-Voice-1, the Speaker Recognition API, the VALL-E, VALL-E X, VALL-E 2, NaturalSpeech, NaturalSpeech 2, NaturalSpeech 3, SpeechX, LiveSpeech, ELaTE, and DeepSinger research model families, and the conversational and speech-recognition models distributed through Microsoft's wholly-owned subsidiary Nuance Communications, Inc.;

(4) Develop and make publicly available a written retention and destruction policy applicable to voiceprints and biometric information of non-users, in compliance with 740 ILCS 14/15(a), and provide a mechanism through which non-user data subjects may request access to, correction of, or deletion of their biometric data from Microsoft's training datasets, model parameters, and commercial inference pipelines;

(5) Destroy all voiceprints and biometric information of Plaintiffs and Class members that Microsoft obtained without BIPA-compliant consent, and certify the destruction;

(6) Destroy or retrain, without the unlawfully obtained voiceprints and biometric information, the foundational voice synthesis models

— and the downstream commercial products built on those models — in which the unlawfully obtained voiceprints and biometric information are encoded, including without limitation the prebuilt Azure neural voice catalog, the Azure neural HD voices, Custom Neural Voice, Personal Voice, the Voice Live API, MAI-Voice-1, the Speaker Recognition API, the VALL-E, NaturalSpeech, and DeepSinger research models, the conversational and speech-recognition models distributed through Nuance Communications, Inc., and any iterative or successor models trained from those models or derived from voiceprints encoded in them. To the extent Microsoft Research has published, on GitHub or any comparable platform, any model artifacts, reference implementations, weights, or other releases that include or enable the reconstruction of model parameters trained on the unlawfully obtained voiceprints, take all reasonable measures to recall and remove the unlawfully encoded model parameters from public availability, including by issuing takedown requests, ceasing further distribution, and ceasing further support of derivative releases. To the extent foundational voice synthesis models containing the unlawfully obtained voiceprints have been transferred to or are deployed through Nuance Communications, Inc., Microsoft International Holdings B.V., Microsoft Ireland Operations Limited, or any other Microsoft affiliated entity, this remedy extends to all such instances;

(7) Cease the use of Plaintiffs' and Class members' identities, including their voices, for commercial purposes without prior written consent, in violation of 765 ILCS 1075/30(a), and cease the distribution or making

available of sound recordings or audiovisual works containing unauthorized digital replicas within the meaning of 765 ILCS 1075/30(b); and

(8) Provide adequate consumer-facing disclosure that voice outputs generated by Microsoft's commercial voice products — including without limitation MAI-Voice-1 outputs deployed in Microsoft Copilot Daily and Copilot Podcasts, Custom Neural Voice and Personal Voice outputs, Voice Live API outputs, Dynamics 365 Contact Center voice outputs, and prebuilt Azure neural voice and Azure neural HD voice outputs — are AI-generated, are derived from foundational voice synthesis models built using voice recordings of unidentified speakers, and have not been authorized by the individuals whose vocal characteristics are reproduced.

B. *BIPA Damages*. Award each Class member, on each statutory subsection for which Defendant is found liable, the greater of liquidated damages or actual damages on a per-person, per-subsection basis consistent with 740 ILCS 14/20 as amended by P.A. 103-769 (effective August 2, 2024) and as construed in *Clay v. Union Pacific Railroad Co.*, No. 25-2185 (7th Cir. Apr. 1, 2026), in the amount of $1,000 per negligent violation or $5,000 per intentional or reckless violation.

C. *IRPA Damages*. Award the actual damages Plaintiffs and Class members sustained as a result of Microsoft's unauthorized commercial use of their identities and voices; the profits Microsoft earned from that unauthorized use to the extent not taken into account in computing actual damages;

statutory minimum damages of $1,000 per violation where applicable under 765 ILCS 1075/40; and punitive damages under 765 ILCS 1075/40.

D. *ICFA Damages*. Award the actual economic damages Plaintiffs and Class members sustained as a result of Microsoft's unfair and deceptive practices, and punitive damages under 815 ILCS 505/10a(c).

E. *Restitution and Disgorgement*. Order restitution of the benefits Microsoft has unjustly obtained from the unauthorized collection and commercial exploitation of Plaintiffs' and Class members' voiceprints and identities; disgorgement of the profits Microsoft has earned through each of the commercial channels alleged at ¶ 68(a)-(h); and an accounting of all such revenues and benefits.

F. *Non-Duplication*. Plaintiffs do not seek duplicative recovery across Counts. To the extent a particular harm is compensated under one Count, Plaintiffs do not seek to recover the same harm under another Count. Plaintiffs' claims for the same conduct under multiple statutes and theories are pleaded in the alternative pursuant to Federal Rule of Civil Procedure 8(d)(2)–(3); the Court may award the relief that, in its judgment, makes Plaintiffs and Class members whole and addresses the unlawful conduct.

G. *Class Certification*. Certify the Class as defined in this Complaint pursuant to Federal Rules of Civil Procedure 23(b)(2) and 23(b)(3); appoint Plaintiffs as Class Representatives; and appoint Plaintiffs' counsel as Class Counsel.

H. *Judgment*. Enter judgment in favor of Plaintiffs and all Class members and against Defendant Microsoft Corporation on all Counts on which Defendant is found liable.

I. *Attorneys' Fees and Costs*. Award reasonable attorneys' fees, costs, and litigation expenses under 740 ILCS 14/20 (BIPA), 765 ILCS 1075/55 (IRPA), 815 ILCS 505/10a (ICFA), 815 ILCS 510/3 (IUDTPA, upon a finding of willfulness), and any other applicable fee-shifting provision.

J. *Interest*. Award pre-judgment and post-judgment interest at the maximum rate permitted by law, including the rate available under 815 ILCS 205/2 for pre-judgment interest where applicable.

K. *Further Relief*. Award such other and further relief as the Court deems just, equitable, and proper.

JURY TRIAL REQUESTED

Plaintiffs, individually and on behalf of all other Class members, request a trial by jury on all claims so triable.

Dated: May 12, 2026

LOEVY & LOEVY

/s/ Ross Kimbarovsky

Ross Kimbarovsky (6229590)
ross@loevy.com
Jon Loevy (6218524)
jon@loevy.com
Michael Kanovitz (6275233)
mike@loevy.com
Matthew Topic (6290923)
matt@loevy.com
Aaron Tucek (98624)
aaron@loevy.com

LOEVY & LOEVY
311 North Aberdeen, 3rd Floor
Chicago, IL 60607
312.243.5900 (phone)
312.243.5902 (fax)

Attorneys for Plaintiffs Alison Flowers, Carol Marin, Philip Rogers, Robin Amer, Lindsey Dorcus, Yohance Lacour, and Victoria Nassif.