UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

In Re: OpenAI, Inc. Copyright Infringement Litigation This Document Relates To: No. 1:23-cv-11195 No. 1:24-cv-03285 No. 1:24-cv-04872	Case No. 1:25-md-3143 (SHS) (OTW) **REDACTED - PUBLIC VERSION**

DEFENDANT MICROSOFT CORPORATION'S RULE 56.1 STATEMENT OF UNDISPUTED MATERIAL FACTS IN SUPPORT OF MOTION FOR SUMMARY JUDGMENT IN NEWS PLAINTIFFS' CONSOLIDATED CASES

Pursuant to Local Civil Rule 56.1, Defendant Microsoft hereby submits this statement of material facts as to which there are no genuine issues to be tried.

COUNT I: COPYRIGHT INFRINGEMENT (17. U.S.C. § 501)

I. COUNT I: COPYRIGHT INFRINGEMENT (17. U.S.C. § 501) – TRAINING CLAIM

A. Fair Use (17 U.S.C. § 107)

1. Purpose and Character of the Use (Factor 1)

1. A large language model ("LLM") is a machine learning deep neural network trained on vast quantities of text with the objective of predicting the next token in a sequence in order to generate natural language.

Lafferty Decl. ¶¶ 16, 33, 77-78; *see also* Hurst Decl. **Ex. 75** (Melton Dep. Tr.) 29:6-30:19, 9:20-10:16, 11:16-17:7; **Ex. 112** (Song Dep. Tr.) 32:13-33:25, 9:8-10:21; **Ex. 54** (Huynh Dep. Tr.) 95:23-96:4, 7:11-8:18, 15:3-7, 15:25-16:25; **Ex. 81** (Nadella Dep. Tr.) 274:24-276:4, 277:3-23, 17:19-18:7, 25:16-22, 29:10-30:3; **Ex. 115**

(Suleyman Dep. Tr.) 35:9-36:8; **Ex. 99** (Ribas Dep. Tr.) 22:2-25, 10:23-11:6, 11:23-12:3, 14:2-23, 17:9-18:10, 18:3-19:19.

2. Machine learning requires large amounts of data, and the more complex and varied the domain, the greater the volume and diversity of data needed for a model to learn patterns that generalize reliably to new inputs. Natural language is among the most complex of such domains.

   Lafferty Decl. ¶¶ 16, 41, 45, 52, 66, 70, 77, 81, 90, 97, 99-100.

3. LLMs are a general-purpose technology capable of a wide variety of natural language processing tasks, such as summarization, translation, reasoning, content generation, and question answering.

   Lafferty Decl. ¶¶ 16, 33, 36, 45, 77-78; *see also* Hurst Decl. **Ex. 112** (Song Dep. Tr.) 32:13-33:25; **Ex. 54** (Huynh Dep. Tr.) 95:23-96 :4, 7:11-8:18, 15:3-7, 15:25-16:25, 17 :14-22; **Ex. 81** (Nadella Dep. Tr.) 274:24-276:4, 277:3-23; **Ex. 115** (Suleyman Dep. Tr.) 35:9-36:8; **Ex. 99** (Ribas Dep. Tr.) 18:3-19:19, 22:2-25; **Ex. 119** (Templeton 30(b)(6) Dep. Tr. (Vol. 2, dated Jan. 21, 2026)) 407:24-408:9; **Ex. 118** (Templeton 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 20, 2026)) 13:14-16, 14:14-20, 15:2-12, 16:18-19:10, 22:1-14, 22:25-23:15, 24:2-25:25, 26:14-27:16, 221:6-222:20.

4. The process of training an LLM involves:

   a) Acquiring a large volume of text to assemble a training corpus—typically by crawling and scraping publicly accessible data on the internet in a manner that has existed before LLMs were invented;

      Lafferty Decl. ¶¶ 69, 79, 89-91, 100; Tucker Decl. ¶ 13(d), 167-173; *see also* Hurst Decl. **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 17:2-18:7, 32:5-33:14, 48:20-24, 66:23-67:10, 95:10-23); **Ex. 130** (Yuan

Dep. Tr.) 12:11-14:13, 28:6-12, 8:6-25, 10:9-23.

b) Processing the textual data into a numerical format that can be used by a neural network—*i.e.*, tokenization, whereby readable text is broken into unique "tokens" comprised of short words, segments of longer words, or groupings of characters, and then integers are assigned to each token;

Lafferty Decl. ¶¶ 105-114; *see also* Hurst Decl. **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 19:12-20:18, 176:22-177:23; **Ex. 117** (Sutskever Dep. Tr.) 141:8-142:2, 19:15-19, 21:12-22:7, 26:17-29:8; **Ex. 117.A** (Sutskever Dep. Ex. 12 - Language Models Are Few-Shot Learners GPT-3 Research Paper) 8-11.

c) Embedding the tokens in high-dimensional, mathematical space, where tokens are represented by numerical vectors that point toward and away from each other based on relationships among the tokens gleaned from numerous contexts across the entirety of the training corpus;

Lafferty Decl. ¶¶ 33, 115-123; *see also* Hurst Decl. **Ex. 117** (Sutskever Dep. Tr.) 27:4-15, 28:5-29:15, 53:3-17, 67:9-15, 141:8-142:2; **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 17:2-18:7, 50:10-19.

d) Processing those vectors through transformer layers and repeatedly making small adjustments to the model "parameters," a set of numerical model weights that the model uses to predict the next token in a sequence of text, shape the model's weights.

Lafferty Decl.¶¶ 137-147; *see also* Hurst Decl. **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 17:2-18:7, 18:16-19:4, 176:24-177:23; **Ex. 101** (Rossett Dep. Tr.) 67:25-68:9, 8:1-9, 8:24-9:9, 17:6-18:4, 18:22-23:2;

Ex. 117 (Sutskever Dep. Tr.) 141:8-142:2; **Ex. 117.A** (Sutskever Dep. Ex. 12 - Language Models Are Few-Shot Learners GPT-3 Research Paper) 2-8.

e) Periodically during training, the model will be presented with data it has not seen during training in order to measure the model's generalization ability. Training continues as long as the model's performance on the unseen data continues to improve. Once the model begins to perform well on training data but begins to lose accuracy on testing data, the training stops.

Lafferty Decl. ¶ 71.

5. Training an LLM requires an enormous volume of text. Publicly disclosed training-set sizes for frontier models rose from approximately 2 trillion tokens in 2023 to approximately 18 trillion tokens by late 2024.

Lafferty Decl. ¶¶ 91-92; Hurst Decl. **Ex. 215** (OPCO_MDL_004424223) at -226; **Ex. 197** (MSFT_AI_MDL_000221039) at -042.

6. The GPT models at issue were pre-trained on corpora of this scale: GPT-3 on approximately 300 billion tokens; GPT-4 and GPT-3.5 Turbo on approximately [redacted] tokens (epoch weighted); and GPT-4o on approximately [redacted] tokens (epoch weighted).

Lafferty Decl. ¶¶ 91-92, 94-97; Hurst Decl. **Ex. 196** (MSFT_AI_MDL_000205724); **Ex. 215** (OPCO_MDL_004424223) at -226; **Ex. 197** (MSFT_AI_MDL_000221039) at -042; **Ex. 96** (Radford Dep. Tr.) 20:25-23:3; **Ex. 227** (Paino Decl.).

7. Training an LLM also requires a wide diversity of text. Model developers draw training data from a broad range of sources, including web data, software code, books,

encyclopedic and scientific sources, forum posts, and transcripts.

Lafferty Decl. ¶¶ 34, 89-91, 93-97; Hurst Decl. **Ex. 117** (Sutskever Dep. Tr.) 26:16-21, 27:4-15, 28:5-29:15, 33:17-34:9, 67:9-15, 255:12-18; **Ex. 196** (MSFT_AI_MDL_000205724); **Ex. 215** (OPCO_MDL_004424223) -226; **Ex. 197** (MSFT_AI_MDL_000221039) -042; **Ex. 198** (MSFT_AI_MDL_000221643); **Ex. 213** (OPCO_MDL_001938121).

8. The internet is the only practicable source of text at a scale sufficiently voluminous and varied to train a general-purpose LLM.

Lafferty Decl. ¶¶ 45, 90-91; Usdan Decl. ¶ 10; *see also* Hurst Decl. **Ex. 99** (Ribas Dep. Tr.) 80:3-81:23; **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 48:20-24, 67:2-6; **Ex. 117** (Sutskever Dep. Tr.) 27:4-15, 28:5-29:15, 33:17-34:9, 67:9-15, 255:12-18; **Ex. 130** (Yuan Dep. Tr.) 28:6-25, 12:11-14:13.

9. LLM training requires an enormous volume and a wide diversity of textual data because the fundamental objective of training is generalization—producing a model capable of responding to new, previously unseen inputs across an effectively unlimited range of contexts and tasks.

Lafferty Decl. ¶¶ 16, 17, 34, 90-91, 98, 154-156.

10. The capacity to generalize is an emergent property of learning across the complete distribution of the training data. No individual work is responsible for conferring that capability.

Lafferty Decl. ¶¶ 16, 33-34, 45, 66, 80-81, 99, 145-146; *see also* Hurst Decl. **Ex. 81** (Nadella Dep. Tr.) 90:9-23, 274:12-275:12.

11. Because the statistical patterns a model learns are distributed redundantly across millions of sources spanning trillions of tokens, the removal of any single work or

collection of works from the training corpus would not materially alter the statistical structure the model learns.

Lafferty Decl. ¶ 33-34, 99, 322.

12. A properly trained LLM is a set of billions or even trillions of numerical parameters (mathematical weights) that encode generalized statistical relationships of language such that the model can simulate conversation and reasoning.

Lafferty Decl. ¶¶ 16, 33-35, 45, 52, 71-72, 74-76, 90-91, 94-100, 139-40, 154-55, 157-58, 164; *see also* Hurst Decl. **Ex. 81** (Nadella Dep. Tr.) 90:9-23, 275:4-276:13; **Ex. 103** (Russinovich Dep. Tr.) 106:4-107:15, 9:3-14, 9:18-23, 12:24-13:15, 14:18-15:17, 16:5-10; **Ex. 16** (Brockman Dep. Tr.) 319:20-320:12, 23:19-23, 24:3-8, 26:22-27:9, 37:9-38:8; **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 17:2-18:7, 50:10-51:23, 181:25-182:4; **Ex. 117** (Sutskever Dep. Tr.) 27:4-15, 28:5-29:15, 33:17-34:9, 67:9-15, 255:12-18; **Ex. 96** (Radford Dep. Tr.) 20:25-23:3.

13. When generalization is strong, an LLM generates coherent and contextually appropriate text about unfamiliar subjects because its parameters represent the deeper structure of language rather than the surface content of the training set; the resulting output is novel text distinct from the source material used for training.

Lafferty Decl. ¶¶ 76-77, 81, 98, 141, 145, 155-56; *see also* Hurst Decl. **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 48:20-50:19, 104:4-17, 180:21-182:4, 184:18-185:18; **Ex. 16** (Brockman Dep. Tr.) 265:19-266:5.

14. Models can at times generate text that was part of their training corpus, a known phenomenon called regurgitation. However, instances of regurgitation in a properly trained model are extremely rare. A model that has a tendency to regurgitate its training data reflects an unintended side-effect of the training process rather than the

intended operation of the model.

Lafferty Decl. ¶ 35; *see also* Usdan Decl. ¶¶ 19-20.

15. At OpenAI's request, Microsoft provided OpenAI with portions of the Bing Index, the web index that powers Microsoft's Bing search engine.

Hurst Decl. **Ex. 118** (Templeton 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 20, 2026)) 304:9-16; **Ex. 119** (Templeton 30(b)(6) Dep. Tr. (Vol. 2, dated Jan. 21, 2026)) 390:22-391:1; **Ex. 38** (Gaggara 30(b)(6) Dep. (dated Aug. 20, 2025)) 44:13-15; **Ex. 37** (Gaffney 30(b)(6) Dep. Tr.) 131:20-133:2, 9:2-5, 9:15-10:5, 10:19-25; **Ex. 37.A** (Gaffney Dep. Ex. 16) at 2-3.

16. The provision of portions of the Bing Index to OpenAI was a step in the process for the purpose of training LLMs.

Hurst Decl. **Ex. 107** (Shan Dep. Tr.) 318:11-319:9; **Ex. 37** (Gaffney 30(b)(6) Dep. Tr.) 131:20-133:2; **Ex. 37.A** (Gaffney Ex. 16); **Ex. 16** (Brockman Dep. Tr.) 323:11-16.

17. To create the Deucalion model, Microsoft participated in post-training on GPT 3.5 Turbo using a synthetic dataset derived from Bing Chat user conversations, of which at most .02% included New York Times content.

Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. Vol. 2 (dated July 31, 2026)) 525:3–15; Lafferty Decl. ¶¶ 25, 336, 338-39.

18. The purpose of creating the Deucalion model was to improve the performance of GPT 3.5 Turbo in following instructions and accurately answering user questions as part of the Copilot product.

Hurst Decl. **Ex. 85** (Parakhin Dep. Tr.) 204:5-22, 206:13-207:22, 15:18-21, 16:1-5, 18:7-12, 22:25-23:8, 23:21-25:6, 26:12-27:4; **Ex. 199**

(MSFT_AI_MDL_002644093); **Ex. 113** (Song 30(b)(6) Dep. Tr.) 46:5–47:13, 53:8–19; Lafferty Decl. ¶¶ 339-40.

19. The New York Times's experts performed no causal analysis demonstrating that the use of at most .02% of New York Times content to construct the synthetic Deucalion post-training dataset contributed in any measurable way to the performance of the Deucalion model.

    Lafferty Decl. ¶¶ 337-39.

20. News Plaintiffs' expert Dr. Goldstein states he also found The Denver Post's content in the Deucalion model post-training dataset, but he did not quantify that content or ascribe any value to it in the performance of the model.

    Hurst Decl. Ex. **231** (Goldstein Opening Report) ¶ 181.

21. News Plaintiffs publish the Asserted Works (which they define as consisting of individual articles) to be read by people for information, edification, and entertainment.

    Hurst Decl. **Ex. 59** (Julien 30(b)(6) Dep. Tr.) 23:1-25:3; **Ex. 42** (Gordon Dep. Tr.) 20:12-21:6; **Ex. 13** (Bauerlein Dep. Tr. (dated January 27, 2026)) 46:20-25; **Ex. 31** (Dolnick Dep. Tr.) 21:4-15; **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-14; **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 185:3-8; **Ex. 18** (Buckingham 30(b)(6) Dep. Tr. (Vol. 2, dated Dec. 19, 2025)) 311:1-15; Tucker Decl. ¶¶ 11(b), 95-97, 102.

2. Nature of the Work (Factor 2) and Amount and Substantiality Used (Factor 3)

22. The LLM training process deconstructs works into tokens, then analyzes the patterns among various pieces of tokenized content in order to translate those patterns into

model weights–all to gain access to underlying facts, ideas, linguistic patterns, and statistical relationships.

Lafferty Decl. ¶¶ 77, 80-81, 105-114, 137-39, 141, 144-47, 283; *see also* Hurst Decl. **Ex. 104** (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 17:2-18:7, 18:12-19:4, 19:12-20:18, 48:25-49:24, 50:10-19, 104:15-17, 176:22-177:23; **Ex. 101** (Rossett Dep. Tr.) 18:22-23:2, 67:25-68:9.

23. The patterns extracted through training—including grammar, syntax, semantic relationships, facts, ideas, and concepts—are properties discernable only from the totality of the training data rather than from any individual work.

Lafferty Decl. ¶¶ 16, 33-34, 45, 66, 80-81, 99, 145-146, 283; *see also* SUMF ¶10.

24. News articles published by News Plaintiffs consist predominantly of factual news reporting.

***The New York Times*:** Hurst Decl. **Ex. 207** (NYT_01077210) (I.3 nyt-2021-2022 – Asserted Works Overview); Procas 30(b)(6) 65:12-66:2. ***New York Daily News*:** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 131:1-21; **Ex. 79.B** (Morris Dep. Ex. 52 – Asserted Works Overview). ***Chicago Tribune*:** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 122:2-125:8; **Ex. 79.A** (Morris Dep. Ex. 49 – Asserted Works Overview). ***Orlando Sentinel*:** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 144:19-145:16; **Ex. 79.D** (Morris Dep. Ex. 56 – Asserted Works Overview). ***Sun Sentinel*:** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 148:20-149:13; **Ex. 79.E** (Morris Dep. Ex. 58 – Asserted Works Overview). ***Mercury News*:** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 138:20-139:21**; Ex. 79.C** (Morris Dep. Ex. 54 – Asserted Works Overview). ***The Denver Post*:** Hurst

Decl. Ex. 180 (DNP-00636807 - DP Authors_part2). ***The Orange County Register:*** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 162:18-163:4; **Ex. 79.G** (Morris Dep. Ex. 62 – Asserted Works Overview). ***Pioneer Press:*** Hurst Decl. **Ex. 79** (Morris 30(b)(6) Dep. (dated Mar. 25, 2026)) 154:6-155:22; **Ex. 79.F** (Morris Dep. Ex. 60 – Asserted Works Overview). ***CIR:*** ECF No. 582 (No. 1:24-cv-04872) (Joint Stipulation and [Proposed] Order Regarding Asserted Works); (Ex. 2 to CIR's First Amended Compl. (1:24-cv-04872-SHS-OTW, ECF 88)); **Ex. 138** (Exhibit A to CIR's First Set of Interrogs. to Defendants).

25. The New York Times's *Wirecutter* articles consist predominantly of factual descriptions and comparisons of the attributes and/or performance of various consumer products.

ECF No. 1-69 (No. 1:23-cv-11195) (Exhibit K to NYT Complaint).

26. The New York Times's *NYT Cooking* recipes consist predominantly of one or more lists of ingredients and the steps necessary to transform those ingredients into the specified dish.

Hurst Decl. **Ex. 207** (NYT_01077210) (I.3 nyt-2021-2022 – Asserted Works Overview); Procas 30(b)(6) 65:12-66:2; **Ex. 124** (Velasquez Dep. Tr.) 99:13-17; **Ex. 30** (Diggs Dep. Tr.) 198:7-23.

3. Impact Upon the Market (Factor 4)

The New York Times Company (The New York Times)

27. From 2019 to 2024, the vast majority of The New York Times's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than LLM training or grounding LLM-based general

purpose chatbots (*e.g.*, syndication)).

Hurst Decl. **Ex. 263** (NYT 2019 10-K); **Ex. 264** (NYT 2020 10-K); **Ex. 265** (NYT 2021 10K); **Ex. 63.A** (Levien Dep. Ex. 1 - NYT 2022 10K); **Ex. 266** (Levien Dep. Ex. 2 - NYT 2023 10K); **Ex. 63.B** (Levien Dep. Ex. 3 - NYT 2024 10K); **Ex. 10** (Bardeen 30(b)(6) Dep. Tr.) 17:22-18:21, 18:22-19:3; 46:19-47:6; **Ex. 129** (Yang 30(b)(6) Dep. Tr.) 78:24-79:16; **Ex. 76** (Miller 30(b)(6) Dep. Tr.) 130:20-24; **Ex. 45** (Grossman-Cohen 30(b)(6) Dep. Tr.) 24:22-29:19, 101:18-102:16; **Ex. 77** (Miller Dep. Tr.) 10:3-11, 14:2-20; Tucker Decl. ¶¶ 30-31, Figure 2.

28. The New York Times's revenues have increased since the use of its Asserted Works to train the LLMs at issue.

a) The New York Times's total revenue has steadily increased since 2020.

Hurst Decl. **Ex. 263** (NYT 2019 10-K); **Ex. 264** (NYT 2020 10-K); **Ex. 265** (NYT 2021 10K); **Ex. 63.A** (Levien Dep. Ex. 1 - NYT 2022 10K); **Ex. 266** (Levien Dep. Ex. 2 - NYT 2023 10K); **Ex. 63.B** (Levien Dep. Ex. 3 - NYT 2024 10K); Tucker Decl. ¶¶ 11(a), 30-31, Figure 2.

b) [redacted]

Hurst Decl. **Ex. 263** (NYT 2019 10-K); **Ex. 264** (NYT 2020 10-K); **Ex. 265** (NYT 2021 10K); **Ex. 63.A** (Levien Dep. Ex. 1 - NYT 2022 10K); **Ex. 266** (Levien Dep. Ex. 2 - NYT 2023 10K); **Ex. 63.B** (Levien Dep. Ex. 3 - NYT 2024 10K); **Ex. 129** (Yang 30(b)(6) Dep. Tr.) 76:22-77:7; **Ex. 10** (Bardeen 30(b)(6) Dep. Tr.) 95:4-96:4; **Ex. 63** (Levien Dep. Tr.) 21:1-25:10, 36:15-39:6, 42:2-15; **Ex. 63.A** (Levien Dep. Ex. 1) at 36; **Ex. 63.B** (Levien Dep.

Ex. 3) 36; Tucker Decl. ¶¶ 11(a), 30-31, Figure 2.

c) [redacted]

Hurst Decl. **Ex. 263** (NYT 2019 10-K); **Ex. 264** (NYT 2020 10-K); **Ex. 265** (NYT 2021 10K); **Ex. 63.A** (Levien Dep. Ex. 1 - NYT 2022 10K); **Ex. 266** (Levien Dep. Ex. 2 - NYT 2023 10K); **Ex. 63.B** (Levien Dep. Ex. 3 - NYT 2024 10K); **Ex. 63** (Levien Dep. Tr.) 21:1-23:4, 25:11-19, 36:15-39:6, 42:16-18, 21:1-23:4, 36:15-39:6; **Ex. 63.A** (Levien Dep. Ex. 1) at 36; **Ex. 63.B** (Levien Dep. Ex. 3) at 36; **Ex. 100** (Robins Dep. Tr.) 205:17-206:3; Tucker Decl. ¶¶ 11(a), 30-31, Figure 2.

d) [redacted]

Hurst Decl. **Ex. 263** (NYT 2019 10-K); **Ex. 264** (NYT 2020 10-K); **Ex. 265** (NYT 2021 10K); **Ex. 63.A** (Levien Dep. Ex. 1 - NYT 2022 10K); **Ex. 266** (Levien Dep. Ex. 2 - NYT 2023 10K); **Ex. 63.B** (Levien Dep. Ex. 3 - NYT 2024 10K); **Ex. 63** (Levien Dep. Tr.) 25:20-26:14, 36:15-39:6, 43:3-13; **Ex. 63.A** (Levien Dep. Ex. 1) at 45; **Ex. 63.B** (Levien Dep. Ex. 3) at 43; Tucker Decl. ¶¶ 11(a), 30-31, Figure 2.

e) [redacted]

Tucker Decl. ¶¶ 11(a), 32-33, Figure 3, Figure 4; *see also* Hurst Decl. **Ex. 129** (Yang 30(b)(6) Dep. Tr.) 76:22-77:7; **Ex. 10** (Bardeen 30(b)(6) Dep. Tr.) 95:4-96:4; **Ex. 63** (Levien Dep. Tr.) 21:1-23:25:10, 36:15-39:6, 42:2-15;

Ex. 63.A (Levien Dep. Ex. 1) at 40; **Ex. 63.B** (Levien Ex. 3) at 10.

29. To the extent The New York Times asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following: News Plaintiffs' expert Dr. Goldstein analyzed logs of 8.2 million real-world conversations with Microsoft's consumer version of Copilot produced by Microsoft. The 8.2 million conversations were selected from a population of approximately 1,134,799,717 total conversations between users and Microsoft's consumer Copilot occurring between July 2023 to September 2024 based on matches with certain keyword terms in any text-containing field. The keyword terms used were primarily the names of News Plaintiffs' publications and URLs including the domain names of their publications (*e.g.*, nytimes.com), as supplied by News Plaintiffs.

Hurst Decl. ¶¶ 2-3; Hurst Decl. **Ex. 231** (Goldstein Opening Report) ¶¶ 382-84.

30. The conversation logs contained [redacted] fields for outputs displayed to a user, as well as [redacted] fields containing web grounding data retrieved by the Copilot system.

Lafferty Decl. ¶¶ 201; Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37.

31. Any logs that hit on a keyword in any of those fields were produced.

Hurst Decl. ¶3; Hurst Decl. **Ex. 231** (Goldstein Opening Report) ¶ 382-84; Lafferty Decl. ¶ 342.

32. Goldstein performed his analysis of the 8.2 million logs in order to assess whether any of the Copilot outputs contained what he qualified as a "regurgitation" of News Plaintiffs' Asserted Works.

Hurst Decl. Ex. 231 (Goldstein Opening Report) ¶ 297; Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 21.

a) Subject to certain exclusions for user-supplied data or lack of originality (which he

assessed using an LLM), Dr. Goldstein ultimately found a regurgitation when either (1) there was a 16-gram match between text of an article in an Asserted Work and either [redacted] or [redacted] fields in a chat log, or (2) using a 4-gram continuity score there was a 16-gram sequence comprised of 4-gram matches, allowing for a single word gap between each 4-gram match (again, in either [redacted] or other [redacted] fields) (this was designed to address privacy redactions). A gram is approximately equivalent to a word.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶¶ 22–25, 27–28.

b) Goldstein found that of The New York Times's 4,519,337 Asserted Works, 2,601 articles matched in the Copilot [redacted] fields (the grounding data) using the 4x4 gram continuity score and 2,462 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

c) Goldstein also found that of The New York Times's 4,519,337 Asserted Works, 183 New York Times articles matched in the Copilot [redacted] fields (the outputs displayed to the user) using the continuity score and 104 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2 (Sur-reply).

d) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

Daily News, LP (Daily News)

33. From fiscal year 2019 to 2024, the vast majority of the Daily News's revenue from

exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 141** (Daily News Resp. to MSFT ROG No. 24); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; Ex. 66 (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 140** (Daily News 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5; Tucker Decl. ¶¶ 34, 36, Figure 6B.

34. [redacted].

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 36, Figure 6B.

b) [redacted].

Tucker Decl. ¶¶ 11(a), 34, 36, Figure 5A, Figure 6B.

c) [redacted].

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

35. To the extent the Daily News asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of the Daily News's 525,341 Asserted Works, 356 articles matched in the Copilot

[redacted] fields using the 4x4 gram continuity score and 333 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of The Daily News's 525,341 Asserted Works, 15 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 6 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2 (Sur-reply).

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

Chicago Tribune Company, LLC (Chicago Tribune)

36. From fiscal year 2019 to 2024, the vast majority of the Chicago Tribune's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 134** (Chicago Tribune Resp. to MSFT ROG No. 24); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 133** (Chicago Tribune 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-

69:5; Tucker Decl. ¶¶ 34-35, Figure 6A.

37. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 35, Figure 5A, Figure 6A.

b) [redacted]

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

38. To the extent the Chicago Tribune asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of the Chicago Tribune's 1,205,350 Asserted Works, 1,200 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 1,159 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 3; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of the Chicago Tribune's 1,205,350 Asserted Works, 82 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 45 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5;

Lafferty Decl. ¶ 409, Table 2 (Sur-reply).

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

Sun-Sentinel Company, LLC (Sun Sentinel)

39. From fiscal year 2019 to 2024, the vast majority of the Sun Sentinel's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 167** (Sun Sentinel Resp. to MSFT ROG No. 23); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; Tucker Decl. ¶¶ 34, 37, Figure 6C.

40. [redacted].

a) [redacted].

Tucker Decl. ¶¶ 11(a), 34, 37, Figure 5A, Figure 6C.

b) [redacted].

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

41. To the extent the Sun Sentinel asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr.

Goldstein found:

a) Of the Sun Sentinel's 134,518 Asserted Works, 121 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 117 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of the Sun Sentinel's 134,518 Asserted Works, 9 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 5 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2 (Sur-reply).

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

Orlando Sentinel Communications Company, LLC (Orlando Sentinel)

42. From fiscal year 2019 to 2024, the vast majority of the Orlando Sentinel's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than use in LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 160** (Orlando Sentinel Resp. to MSFT ROG No. 23); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 156** (Orlando Sentinel 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67**

(Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5; Tucker Decl. ¶¶ 34, 38, Figure 6D.

43. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 38, Figure 6D.

b) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 38, Figure 5A, Figure 6D.

c) [redacted].

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

44. To the extent the Orlando Sentinel asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of the Orlando Sentinel's 424,822 Asserted Works, 289 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 276 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of the Orlando Sentinel's 424,822 Asserted Works, 17 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and

9 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶¶ 26–28, 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2.

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

San Jose Mercury-News, LLC (The Mercury News)

45. From fiscal year 2019 to 2024, the vast majority of The Mercury News's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than use in LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 149** (Mercury News Resp. to MSFT ROG No. 24); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 148** (Mercury News 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5; Tucker Decl. ¶¶ 34, 39, Figure 6E.

46. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 39, Figure 6E.

b) [redacted].

Tucker Decl. Ex. ¶¶ 11(a), 34, 39, Figure 5B, Figure 6E.

c) [redacted]

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

47. To the extent The Mercury News asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of The Mercury News's 130,331 Asserted Works, 931 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 877 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of The Mercury News's 130,331 Asserted Works, 94 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 57 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2.

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

DP Media Network, LLC (The Denver Post)

48. From fiscal year 2019 to 2024, the vast majority of The Denver Post's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and

licensing (for purposes other than use in LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 145** (Denver Post Resp. to MSFT ROG No. 24); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 140** (Denver Post 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5; Tucker Decl. ¶¶ 34, 40, Figure 6F.

49. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 40, Figure 6F.

b) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 40, Figure 5B, Figure 6F.

c) [redacted]

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

50. To the extent The Denver Post asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of The Denver Post's 89,269 Asserted Works, 1,005 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 969 using the 16-gram

score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of The Denver Post's 89,269 Asserted Works, 123 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 78 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2.

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

ORB Publishing, LLC (The Orange County Register)

51. From fiscal year 2019 to 2024, the vast majority of The Orange County Register's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than use in LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 157** (Orange County Register Resp. to MSFT ROG No. 23); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 156** (OC Register 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5; Tucker Decl. ¶¶ 34, 41, Figure 6G.

52. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 41, Figure 6G.

b) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 41, Figure 5B, Figure 6G.

c) [redacted]

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

53. To the extent The Orange County Register asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of The Orange County Register's 35,603 Asserted Works, 36 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 32 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of the Orange County Register's 35,603 Asserted Works, 1 article matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 1 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2.

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

Northwest Publications, LLC (Pioneer Press)

54. From fiscal year 2019 to 2024, the vast majority of the Pioneer Press's revenue from exploitation of its Asserted Works was earned through subscriptions, advertising, and licensing (for purposes other than use in LLM training or grounding LLM-based general purpose chatbot (e.g., syndication)).

Hurst Decl. **Ex. 164** (Pioneer Press Resp. to MSFT ROG No. 23); **Ex. 56** (Ibrahim 30(b)(6) Dep. Tr.) 26:13-24; **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 14:10-20; **Ex. 163** (Pioneer Press 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5; Tucker Decl. ¶¶ 34, 42, Figure 6H.

55. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 42, Figure 6H; Hurst Decl. **Ex. 20** (Burbach Deposition) 22:7-23:1.

b) [redacted]

Tucker Decl. ¶¶ 11(a), 34, 42, Figure 5B, Figure 6H.

c) [redacted]

Hurst Decl. **Ex. 66** (Loretto 30(b)(6) Dep. Tr.) 95:7-18.

56. To the extent the Pioneer Press asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: Employing the methodology described in Undisputed Facts ¶¶ 29-32(a), News Plaintiffs' Expert Dr. Goldstein found:

a) Of the Pioneer Press's 47,527 Asserted Works, 326 articles matched in the Copilot [redacted] fields using the 4x4 gram continuity score and 300 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 1, 4; Lafferty Decl. ¶ 408, Table 1 (Sur-reply).

b) Of the Pioneer Press's 47,527 Asserted Works, 21 articles matched in the Copilot [redacted] fields (i.e., the user-facing responses) using the continuity score and 12 using the 16-gram score.

Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Tables 2, 5; Lafferty Decl. ¶ 409, Table 2.

c) Goldstein did not report a match to any entire article.

See Hurst Decl. **Ex. 235** (Goldstein Supplemental Report) ¶ 37, Table 2 (highest threshold Goldstein reported was a continuity score of 50 or above).

The Center for Investigative Reporting (CIR)

57. From 2019 to 2024, the vast majority of CIR's revenue from exploitation of its Asserted Works was earned through subscriptions, donations, and licensing (for purposes other than use in LLM training or grounding LLM-based general purpose chatbot (*e.g.*, syndication)).

Hurst Decl. **Ex. 18** (Buckingham 30(b)(6) Dep. Tr. Vol. 2 (dated Dec. 19, 2025))

311:1-15, 405:25-408:22; **Ex. 17** (Buckingham 30(b)(6) Dep. Tr. (Vol. 1, dated Dec. 18, 2025)) 194:5-7; **Ex. 17A** (Buckingham Dep. Ex. 30 – Building the investigative newsroom of the future); **Ex. 12** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 2, dated Jan. 27, 2026)) 378:19-379:3, 385:9-24; **Ex. 11** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 26, 2026)) 293:25-294:22, 297:22-25, 299:13-21; Ex. 13 (Bauerlein Dep. Tr. dated Jan. 27, 2026) 58:16-64:18; **Ex. 12.A** (Bauerlein Dep. Ex. 47); Tucker Decl. ¶ 44.

58. [redacted]

a) [redacted]

Tucker Decl. ¶¶ 11(a), 44, 48, Figure 8; *see also* Hurst Decl. **Ex. 17** (Buckingham 30(b)(6) Dep. Tr. (Vol. 1, dated Dec. 18, 2025)) 194:5-7; **Ex. 12** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 2, dated Jan. 27, 2026)) 385:9-24.

b) [redacted]

Tucker Decl. ¶¶ 11(a), 44-47, Figure 7; *see also* Hurst Decl. **Ex. 11** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 26, 2026)) 299:13-21.

59. To the extent CIR asserts harm from regurgitation as a result of the use of its content for LLM training, Microsoft offers the following facts: CIR's expert Dr. Wenger analyzed the same logs of approximately 8.2 million Copilot conversations described in Undisputed Facts ¶¶ 29-32(a).

Hurst Decl. **Ex. 246** (Wenger Opening Report) at 104-05; **Ex. 247** (Wenger Reply Report) at 113 & n.162; Lafferty Decl. ¶¶27, 28, Table 1 (Rebuttal), 360.

60. Wenger analyzed the Copilot [redacted] and [redacted] fields using a filter built from 16-grams of CIR's Asserted Works.

Hurst Decl. **Ex. 246** (Wenger Opening Report) at 104-05; **Ex. 247** (Wenger Reply Report) at 113 & n.162; Lafferty Decl. ¶¶ 28, Table 1 (Rebuttal), 360.

a) Dr. Wenger treated as a "putative regurgitation" any Copilot [redacted] that contained at least five 16-gram matches to an individual *Mother Jones* Asserted Work.

Hurst Decl. **Ex. 246** (Wenger Opening Report) at 74–75, 105–06; Lafferty Decl. ¶ 360.

b) Dr. Wenger found 51 articles over the five 16-gram match threshold in the Copilot [redacted] fields, after exclusion of public-domain text (e.g., the Constitution or the Bible).

Hurst Decl. **Ex. 246** (Wenger Opening Report) at 38, 40, 105–07; Lafferty Decl. ¶¶ 360–61.

c) In her analysis of the Copilot logs, Wenger did not report a match to any entire article.

Hurst Decl. **Ex. 246** (Wenger Opening Report) at 106–07; Lafferty Decl. ¶ 360.

As to all News Plaintiffs

61. Microsoft's expert Dr. Lafferty calculated a regurgitation rate in the user-facing output [redacted] for all The New York Times and Daily News Plaintiffs' Asserted Works of 0.004% of all logged Copilot conversations, and for all CIR's Asserted Works of 0.00008% of all logged Copilot conversations.

Lafferty Decl. ¶¶ 28, Table 1 (Rebuttal), 374-77, Table 4 (Rebuttal) (New York

Times and Daily News Plaintiffs); *id.* ¶¶ 361–64, 344 Table 2 (Rebuttal) (CIR).

62. Between 2019 and 2024, News Plaintiffs did not enter into any agreement to license or sell their Asserted Works for the purpose of use as training data for general purpose LLMs.

New York Times: Hurst Decl. **Ex. 46** (Grossman-Cohen 30(b)(6) Dep. Tr.) 24:22-29:19, 101:18-102:16; **Ex. 77** (Miller Dep. Tr.) 10:3-11, 14:2-20. ***Daily News:*** Hurst Decl. **Ex. 140** (Daily News 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Chicago Tribune:*** Hurst Decl. **Ex. 133** (Chicago Tribune 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Orlando Sentinel:*** Hurst Decl. **Ex. 159** (Orlando Sentinel 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Sun Sentinel:*** Hurst Decl. **Ex. 166** (Sun Sentinel 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***The Mercury News:*** Hurst Decl. Ex. 148 (Mercury News 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***The Denver Post:*** Hurst Decl. **Ex. 144** (Denver Post 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026))

206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***The Orange County Register:*** Hurst Decl. **Ex. 156** (OC Register 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Pioneer Press:*** Hurst Decl. **Ex. 163** (Pioneer Press 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***CIR:*** Hurst Decl. **Ex. 12** (Bauerlein Dep. Tr. dated Jan. 27, 2026) 58:16-64:18; **Ex. 12.A** (Bauerlein Dep. Ex. 47 – Bauerlein email with Bloomberg reporter (dated July 25, 2024)); **Ex. 18** (Buckingham 30(b)(6) Dep. Tr. (Vol. 2, dated Dec. 19, 2025)) 405:25-408:22.

63. There are substantial impediments to the formation of a market for licensing a training corpus sufficient to train the type of general purpose LLMs at issue in this case. For example:

a) A very large volume and a large variety of textual data is necessary for the training corpus to develop general purpose LLMs like those at issue; as such, no individual work has value by itself as training material nor is it currently technologically possible to isolate the contribution of any particular piece of data to an LLM's output generation.

Lafferty Decl. ¶¶ 16, 34, 94-97, 99, 168, 283, 290, 322, 433; Usdan Decl. ¶¶ 10-11; Hurst Decl. **Ex. 53** (Hoeschele Dep. Tr. (Vol. 1, dated Apr. 2, 2025)) 16:17-22, 17:1-23, 19:12-20:1, 23:24-24:25 25:15-23, 26:6-22, 35:18-36:11; **Ex. 117** (Sutskever Dep. Tr.) 27:4-15, 28:5-29:15, 67:9-15, 255:12-256:2.

b) The Asserted Works were a minuscule share of the training corpus used to train

the GPT models.

Lafferty Decl. ¶¶ 94-97, 105-109; Tucker Decl. ¶¶ 13(a), 134, 137-138.

c) For a traditional copyright license, rightsholders monetize the inherent value of specific creative expression contained within their individual works. By contrast, when licensing data for LLM training, value is not attributable to any discrete piece of interchangeable data, but rather emerges from statistical distributions observable only across large-scale datasets. There is no settled or workable means for pricing any given piece of training data.

Tucker Decl. ¶¶ 13(b), 149-153.

d) A developer seeking to license a training corpus would need to obtain rights from millions of dispersed authors and content creators.

Tucker Decl. ¶¶ 13(a)-(b), 130-133, 139-144; Usdan Decl. ¶¶ 10-11, 22, 24.

e) Rightsholders for individual works can be difficult and costly to identify and locate, increasing the transaction costs of negotiating any proposed training data licenses.

Tucker Decl. ¶¶ 13(b), 139-144; Usdan Decl. ¶ 24

f) Negotiations for the aggregation of the necessary volume and variety of data are susceptible to holdout problems.

Tucker Decl. ¶¶ 13(b), 145-148.

64. After the release of the LLMs at issue in this case, News Plaintiffs have continued to write and publish newspapers and magazines.

Hurst Decl. **Ex. 116** (Sulzberger Dep. Tr.) 177:24-178:6; 246:5-17; **Ex. 12** (Bauerlein 30(b)(6) Dep. Tr.) (Vol. 2, dated Jan. 27, 2026) 423:6-16; **Ex. 59** (Julien Dep. Tr.) 33:12-17; Tucker Decl. ¶¶ 14, 30-48, 237-238.

65. To the extent that News Plaintiffs claim that using the Asserted Works to train the LLMs at issue resulted in loss of conversions or traffic, *see infra* ¶¶ 86-91, 93-97, 99-103, 105-109, 111-115, 117-121, 123-127, 129-133, 135-139, 141-146 which are hereby incorporated by reference.

II. COUNT I: COPYRIGHT INFRINGEMENT (17. U.S.C. § 501) – GROUNDING AND OUTPUT CLAIMS RELATED TO WEB GROUNDING

A. Fair Use (17 U.S.C. § 107)

1. Purpose and Character of the Use (Factor 1)

66. An LLM is inherently limited by the data included in the training corpus used to create its statistical parameters; accordingly, a standalone, LLM-based chatbot cannot reliably—or usefully—respond to those queries that require information omitted from or first discovered after the data of the training corpus.

Abib Decl. ¶¶ 9-10; Lafferty Decl. ¶¶ 18, 178, 189-91; Hurst Decl. **Ex. 99** (Ribas Dep. Tr.) 155:25-156:20.

67. Microsoft added a functionality called "web grounding" to its LLM-based chatbot, Copilot. Web grounding enables Copilot to retrieve content that was obtained from the public internet, and to use such content in responding to user queries.

Abib Decl.¶¶ 8-22; Hurst Decl. **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9, **Ex. 1.A** (Abib 30(b)(6) Ex. 3); **Ex. 71** (Manivannan 30(b)(6) Dep. Tr.) 50:15-51:6, 9:6-9, 11:20-22, 12:5-16, 13:6-9; **Ex. 85** (Parakhin Dep. Tr.) 216:9-217:17.

68. By providing Copilot with an ever-evolving source of information, web grounding greatly expands the usefulness of Copilot for users because Copilot can accurately and more completely respond to a significantly broader range of user queries than a standalone LLM-based chatbot limited by its training data.

Abib Decl. ¶ 10; Lafferty Decl. ¶ 178; Hurst Decl. **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9; **Ex. 1.A** (Abib 30(b)(6) Ex. 3); **Ex. 71** (Manivannan 30(b)(6) Dep. Tr.) 50:15-51:6.

69. Copilot web grounding also improves LLM performance by reducing the likelihood of

so-called "hallucinations"—seemingly plausible but inaccurate outputs—because the chatbot relies on not just next token prediction but also retrieved content in order to provide more reliable responses supported by material synthesized from publicly accessible web sources.

Abib Decl. ¶ 10; Bird Decl. ¶¶ 30, 39; *see also* Hurst Decl.; **Ex. 26** (Cook Dep. Tr.) 111:19-112:9, 116:1-11, 117:1-118:4, 10:12-20, 11:1-6, 16:14-17:12, 20:24-21:1, 22:1-24:18; **Ex. 71** (Manivannan 30(b)(6) Dep. Tr.) 50:18-22; Lafferty Decl. ¶¶ 178, 312.

70. Web grounding enhances a user's experience; for example, by accessing information from Microsoft's Bing search infrastructure, Copilot with web grounding improves knowledge discovery, synthesis, and location by enabling direct and informed responses to user queries in the context of a conversational chatbot.

Abib Decl. ¶¶ 5-7, 23-26; Hurst Decl. **Ex. 85** (Parakhin Dep. Tr.) 212:17-217:17; **Ex. 85.A** (Parakhin Dep. Ex. 25 - The new Bing: Our approach to Responsible AI documentation); **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9, **Ex. 1.A** (Abib 30(b)(6) Ex. 3).

71. Copilot's web grounding works by connecting the LLM with Microsoft Bing via Retrieval Augmented Generation, or RAG. When the Copilot system recognizes that responding to a user query may benefit from web grounding, the query proceeds down a "grounding pipeline" as follows:

a) First, the user's query is tokenized in order to be interpreted and acted upon by the LLM;

Cook Decl. ¶ 5; Hurst Decl. **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9; **Ex. 1.A** (Abib 30(b)(6) Ex. 3); **Ex. 101** (Rossett Dep. Tr.) 69:20-72:23.

b) Second, Copilot sends a search-optimized query to the Bing Index, obtaining the

top results: six potentially responsive webpages;

Cook Decl. ¶ 5; Hurst Decl. **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9; **Ex. 1.A** (Abib 30(b)(6) Ex. 3); **Ex. 101** (Rossett Dep. Tr.) 69:20-72:23.

c) Third, each webpage—or a truncated portion thereof—is tokenized;

Cook Decl.¶ 5; Hurst Decl. **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9; **Ex. 1.A** (Abib 30(b)(6) Ex. 3); **Ex. 101** (Rossett Dep. Tr.) 69:20-72:23.

d) Fourth, all of the tokenized grounding content is placed alongside the tokenized user query into a temporary "context window" and passed along to Copilot's LLM;

Cook Decl. ¶ 5; Hurst Decl. **Ex. 1** (Abib 30(b)(6) Dep. Tr.) 14:22-16:9; **Ex. 1.A** (Abib 30(b)(6) Ex. 3); **Ex. 101** (Rossett Dep. Tr.) 69:20-72:23.

e) Fifth, the LLM analyzes this "context window" in order to generate a response using its underlying statistical parameters.

Cook Decl. ¶ 5; Hurst Decl. **Ex. 101** (Rossett Dep. Tr.) 69:20-72:23.

72. Copilot's web-grounded responses are influenced by all tokens in the model's context window; the grounded content tokens alter the LLM's probability distribution (its statistical prediction of the next sequence of tokens), which influences the token sequence generated in response to the user query. Copilot is thereby able to synthesize different pieces of grounding data and the user prompt in generating a response, while also drawing upon the background understanding, insight, and reasoning capacity encoded into its underlying LLM parameters.

Cook Decl. ¶ 6; Lafferty Decl. ¶¶ 189-91; Hurst Decl. **Ex. 101** (Rossett Dep. Tr.) 69:20-72:23.

73. Copilot provides attribution and links to every webpage used as grounding content,

helping users navigate directly to those webpages to verify information or learn more.

Abib Decl. ¶¶ 6, 10; Hurst Decl. **Ex. 26** (Cook Dep. Tr.) 111:14-112:17, 147:20-148:1; **Ex. 99** (Ribas Dep. Tr.) 65:7-12.

74. Microsoft deployed a multi-layered mitigation system to prevent Copilot from outputting copyrighted content in response to a user query. Microsoft's mitigations include, for example: metaprompt and system prompt instructions; user input classifiers and filters; and an output detection system.

Bird Decl. ¶¶ 34-43; *see also* Hurst Decl. **Ex. 26** (Cook Dep. Tr.) 37:5-10, 52:17-53:20, 61:16-62:25, 65:18-66:3, 67:2-24, 69:5-11, 77:22-78:3, 107:3-6, 137:22-138:14, 144:7-145:13; **Ex. 85** (Parakhin Dep. Tr.) 217:18-218:20; **Ex. 99** (Ribas Dep. Tr.) 132:8-135:12.

2. Nature of the Work (Factor 2)

75. Microsoft incorporates Undisputed Facts ¶¶ 22-26 as though fully restated herein.

3. Amount and Substantiality Used (Factor 3)

76. Copilot web grounding is not designed to display grounding content to a user as part of the LLM's grounded response; that is, the system has no deterministic functionality for passing grounding content through to a user.

Cook Decl. ¶ 7.

77. As described in Undisputed Fact ¶ 72, Copilot uses grounding data within the LLM's context window to provide a newly generated string of tokens through next-token prediction.

Cook Decl. ¶ 7; Lafferty Decl. ¶¶ 189-91; *see also* Hurst Decl. **Ex. 26** (Cook Dep. Tr.) 73:22-11.

78. Because an LLM's next-token prediction is non-deterministic, an LLM can

incidentally predict token strings that match grounding data. It is also possible that the LLM, to satisfy a particular user prompt, could attempt to provide a portion of tokens from grounding data. These behaviors are intermediated by Copilot's multi-layered mitigation system, described at Undisputed Fact ¶ 74, to prevent Copilot responses from containing significant output matching copyrighted content.

Cook Decl. ¶ 8; Lafferty ¶¶ 187, 191, 398.

79. News Plaintiffs' experts Goldstein and Wenger identified certain outputs containing matching text in the pool of 8.2 million conversations that hit on keywords associated with News Plaintiffs. Goldstein's analysis finds 59,545 of these conversations contain matching text. Wenger's analysis searched for conversations with at least 5 16-gram matches; she found 51 total out of the 8.2 million conversation sample.

See Lafferty Decl. ¶¶ 360-61; 396-401.

80. Meta tag controls are machine readable code on a webpage that operate at the individual HTML page level and communicate to a web crawling robot, such as Microsoft's Bingbot, how to process that webpage. When a website permits crawling by Bingbot, Bingbot retrieves the content found on a particular webpage and stores this content in Microsoft's Bing Index.

Canel Decl. ¶¶ 11-12; Gaggara Decl. ¶ 15; *see also* Hurst Decl. **Ex. 85** (Parakhin Dep. Tr.) 226:23–227:24; **Ex. 99** (Ribas Dep. Tr.) 178:9–17; **Ex. 64** (Liang Dep. Tr.) 169:7-170:10.

81. After the release of Bing Chat, Microsoft implemented a change to its system whereby the existing NOARCHIVE and NOCACHE meta tags were designed to be used by publishers to signify the degree, if any, to which their content should be excluded from

web grounding.

Fabrice Decl. ¶¶ 23-29.

82. On September 22, 2023, Microsoft announced—via a post published on Microsoft Bing Blogs—that it had repurposed the NOARCHIVE and NOCACHE meta tags for publishers to use as AI-related controls. The blog post stated that content tagged with the repurposed NOARCHIVE meta tag "will not be included in Bing Chat answers, not be linked to in the answers." A webmaster's inclusion of the NOARCHIVE meta tag in the HTML on a webpage stored in the Bing Index signals to the Copilot system that it is supposed to exclude that webpage entirely from the data provided to Copilot for web grounding.

Canel Decl. ¶ 27; *see also* Hurst Decl. **Ex. 21** (Canel 30(b)(6) Dep. Tr.) 201:2-20, 202:9-11; **Ex. 21.A** (Canel Dep. Ex. 35 – Microsoft Bing Blog Announcing Meta Tags); **Ex. 93** (Procas 30(b)(6) Dep. Tr.) 179:2-181:8, 182:17-183:9; **Ex. 64** (Liang Dep. Tr.) 181:8–182:21; **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 104:10-105:12.

83. The blog post also stated that portions of content from a web page tagged with the repurposed NOCACHE meta tag "may be included in Bing Chat answers." A webmaster's inclusion of the NOCACHE meta tag in the HTML on a webpage stored in the Bing Index signals to the Copilot system that it is supposed to exclude from grounding data everything on the page other than the page title, URL and a snippet.

Canel Decl. ¶ 27; *see also* Hurst Decl. **Ex. 21** (Canel 30(b)(6) Dep. Tr.) 201:2-20, 202:9-11; **Ex. 21.A** (Canel Dep. Ex. 35 – Microsoft Bing Blog Announcing Meta Tags); **Ex. 85** (Parakhin Dep. Tr.) 217:18–218:20, 226:23–227:24; **Ex. 99** (Ribas Dep. Tr.) 132:8–135:12; **Ex. 93** (Procas 30(b)(6) Dep. Tr.) 179:2-181:8; **Ex. 57**

(Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 104:10-105:12.

84. The blog post also explained that if a webmaster uses neither meta tag, then the content "may be included in Bing Chat answers." Without a tag, when stored in the Bing Index the web page is used as grounding data by the Copilot system if it is selected as one of the relevant sources of data.

Canel Decl. ¶ 27; *see also* Hurst Decl. **Ex. 21** (Canel 30(b)(6) Dep. Tr.) 201:2-20, 202:9-11; **Ex. 21.A** (Canel Dep. Ex. 35); **Ex. 85** (Parakhin Dep. Tr.) 15:18-21, 16:1-5, 18:7-12, 22:25-23:8, 23:21-25:6, 26:12-27:4, 217:18–218:20; **Ex. 93** (Procas 30(b)(6) Dep. Tr.) 179:4-181:8, 182:17-183:9; **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 104:9-105:12.

4. Impact Upon the Market (Factor 4)

The New York Times Company (The New York Times)

85. Microsoft incorporates Undisputed Facts ¶¶ 27-32, 61 as though fully restated herein.

86. Although The New York Times was capable of tracking and in fact collected at least some conversion data (*i.e.,* data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 129** (Yang 30(b)(6) Dep. Tr.) 20:5-22:23, 64:4-65:15, 66:21-67:16; **Ex. 114** (Stricchiola Dep. Tr.) 97:22-98:4, 98:14-99:2, 99:22-24; **Ex. 281** (Bell Dep. Tr.) 144:16-145:3.

87. [redacted]

Tucker Decl. ¶¶ 11(a), 49-53, Figure 9; Hurst Decl. **Ex. 91** (Poferl 30(b)(6) Dep.

Tr.) 15:2-15.

88. Although The New York Times collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026) 109:9-110:5.

89. [redacted]

Tucker Decl. ¶¶ 11(a), 63-64, Figure 12.

90. [redacted]

Tucker Decl. ¶¶ 11(a), 65-66, Figure 13A, Figure 13B, Figure 14A, Figure 14B.

91. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 23.

Daily News, LP (Daily News)

92. Microsoft incorporates Undisputed Facts ¶¶ 33-35, 61 as though fully restated herein.

93. Although the Daily News was capable of tracking and in fact collected at least some conversion data (*i.e.,* data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

94. [redacted]

[redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

95. Although the Daily News collects data regarding the sources of its traffic, its experts did not analyze whether Copilot affected actual traffic volumes to its website.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

96. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 71, 77-78, 80, Figure 15, Figure 16.

97. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

Chicago Tribune Company, LLC (Chicago Tribune)

98. Microsoft incorporates Undisputed Facts ¶¶ 36-38, 61 as though fully restated herein.

99. Although the Chicago Tribune was capable of tracking and in fact collected at least some conversion data (*i.e.,* data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

100. [redacted]

Tucker Decl. ¶¶ 11(a) 49-51, 54-55, Figure 10.

101. Although the Chicago Tribune collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

102. [redacted]

Tucker Decl. ¶¶ 11(a), 67-70, 77-79, Figure 15, Figure 16.

103. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

Sun-Sentinel Company, LLC (Sun Sentinel)

104. Microsoft incorporates Undisputed Facts ¶¶ 39-41, 61 as though fully restated herein.

105. Although the Sun Sentinel was capable of tracking and in fact collected at least some conversion data (i.e., data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

106. [redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

107. Although the Sun Sentinel collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

108. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 72, 77-78, 81, Figure 15, Figure 16.

109. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

Orlando Sentinel Communications Company, LLC (Orlando Sentinel)

110. Microsoft incorporates Undisputed Facts ¶¶ 42-44, 61 as though fully restated herein.

111. Although the Orlando Sentinel was capable of tracking and in fact collected at least some conversion data (i.e., data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

112. [redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

113. Although the Orlando Sentinel collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

114. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 73, 77-78, 82, Figure 15, Figure 16.

115. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

San Jose Mercury-News, LLC (The Mercury News)

116. Microsoft incorporates Undisputed Facts ¶¶ 45-47, 61 as though fully restated herein.

117. Although The Mercury News was capable of tracking and in fact collected at least some conversion data (i.e., data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

118. [redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

119. Although The Mercury News collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

120. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 73, 77-78, 82, Figure 15, Figure 16.

121. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

DP Media Network, LLC (The Denver Post)

122. Microsoft incorporates Undisputed Facts ¶¶ 48-50, 61 as though fully restated herein.

123. Although The Denver Post was capable of tracking and in fact collected at least some conversion data (i.e., data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

124. [redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

125. Although The Denver Post collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

126. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 74, 77-78, 83, Figure 15, Figure 16.

127. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

ORB Publishing, LLC (The Orange County Register)

128. Microsoft incorporates Undisputed Facts ¶¶ 51-53, 61 as though fully restated herein.

129. Although The Orange County Register was capable of tracking and in fact collected at least some conversion data (*i.e.,* data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

130. [redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

131. Although The Orange County Register collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

132. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 75, 77-78, 84, Figure 15, Figure 16.

133. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

Northwest Publications, LLC (Pioneer Press)

134. Microsoft incorporates Undisputed Facts 54-56, 61 as though fully restated herein.

135. Although the Pioneer Press was capable of tracking and in fact collected at least some conversion data (*i.e.,* data indicating the rate and sources from which online traffic referrals resulted in a reader signing up in order to read paywalled content), its experts performed no analysis of that conversion data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 97:21-98:4, 98:14-99:8, 99:22-24; **Ex. 82** (Owen Dep. Tr.) 107:13-23, 117:4-14.

136. [redacted]

[redacted]

Tucker Decl. ¶¶ 11(a), 49-51, 54-55, Figure 10.

137. Although the Pioneer Press collects data regarding the sources of its traffic, its experts performed no analysis of that traffic data.

Hurst Decl. **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24; **Ex. 69** (Malackowksi Dep. Tr. (Vol. 1, dated July 15, 2026)) 109:9-110:5; **Ex. 82** (Owen Dep. Tr.) 117:4-6.

138. [redacted]

Tucker Decl. ¶¶ 11(a), 67-69, 76, 77-78, 85, Figure 15, Figure 16.

139. [redacted]

Tucker Decl. ¶¶ 12(b), 112-113, Figure 24.

The Center for Investigative Reporting (CIR)

140. Microsoft incorporates Undisputed Facts ¶¶ 57-61 as though fully restated herein.

141. CIR has no paywall on its *Mother Jones* website.

Hurst Decl. **Ex. 18** (Buckingham 30(b)(6) Dep. Tr. (Vol. 2, dated Dec 19, 2025)) 295:9-24.

142. CIR uses a window of approximately two months between publication of its *Mother Jones* magazine and inclusion of that content on the *Mother Jones* website.

Hurst Decl. **Ex. 11** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 1, dated Jan 26, 2026)) 144:21-145:11.

143. CIR's expert did not identify any causal relationship between the introduction of

Copilot and the traffic to the *Mother Jones* website.

Hurst Decl. **Ex. 27** (Cross Dep. Tr.) 142:15-21; **Ex. 74** (McDonald Dep. Tr.) 92:24-100:5.

144. [redacted]

Tucker Decl. ¶¶ 11(a), 86-89, Figure 17, Figure 18, Figure 19.

As to All News Plaintiffs

145. Using both a survey and an analysis of Copilot conversation logs, Microsoft's experts found that consumers rarely use Copilot to obtain information about current events.

Bail Decl. ¶¶ 97, 124-25; Amir Decl. ¶¶ 9(c), 42, 55, 57, 59-60.

146. News Plaintiffs' experts produced no analysis quantifying any loss of traffic, subscriptions, or revenue as a result of readers choosing to read a Copilot output with "regurgitated" content (as measured by Goldstein) as a substitute for reading any of their Asserted Works.

New York Times: Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 281** (Bell Dep. Tr.) 99:16-22, 144:16-145:3, 168:20-169:15; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***Daily News:*** Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.)

63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***Chicago Tribune:*** Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***Orlando Sentinel:*** Hurst Decl. Ex. 41 (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***Sun Sentinel:*** Hurst Decl. **Ex. 41** (Goldstein Dep. Tr.) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***The Mercury News:*** Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***The Denver Post:*** Hurst Decl. **Ex. 41** (Goldstein Dep.

Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); Ex. 82 (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***The Orange County Register:*** Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 109:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***Pioneer Press:*** Hurst Decl. **Ex. 41** (Goldstein Dep. Tr. (Vol 2, dated July 31, 2026)) 508:1–509:10; **Ex. 69** (Malackowski Dep. Tr. (Vol 1, dated July 15, 2026)) 105:24-108:22, 10 9:9-110:5; **Ex. 70** (Malackowski Dep. Tr. (Vol 2, dated July 16, 2026)) 379:4–20, 381:17–384:4, 386:16-388:13); **Ex. 82** (Owen Dep. Tr.) 63:9-12, 70:8-71:18, 104:12-18, 107:13-23, 117:4-14, 217:10-218:10; **Ex. 114** (Stricchiola Dep. Tr.) 82:16-83:10, 97:22-99:24. ***CIR:*** Hurst Decl. **Ex. 125** (Wenger Dep. Tr.) 249:1–13; **Ex. 74** (McDonald Dep. Tr.) 70:6-15, 92:24-100:5, 228:2-10; **Ex. 27** (Cross Dep. Tr.) 142:15-21.

147. Between 2019 and 2024, News Plaintiffs did not enter into any agreement to license their Asserted Works for the purpose of use as grounding data for LLM-based general purpose chatbots.

New York Times: Hurst Decl. **Ex. 46** (Grossman-Cohen 30(b)(6) Dep. Tr.) 24:22-29:18, 101:18-102:16; **Ex. 77 (**Miller Dep. Tr.) 10:3-11, 14:2-20. ***Daily***

News: Hurst Decl. **Ex. 140** (Daily News 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Chicago Tribune:*** Hurst Decl. **Ex. 133** (Chicago Tribune 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Orlando Sentinel:*** Hurst Decl. **Ex. 159** (Orlando Sentinel 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Sun Sentinel:*** Hurst Decl. **Ex. 166** (Sun Sentinel 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 67** (Lown 30(b)(6) Dep. Tr.) 48:24-49:7; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***The Mercury News:*** Hurst Decl. **Ex. 148** (Mercury News 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***The Denver Post:*** Hurst Decl. **Ex. 144** (Denver Post 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***The Orange County Register:*** Hurst Decl. **Ex. 156** (OC Register 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***Pioneer Press:*** Hurst Decl. **Ex. 163** (Pioneer Press 4th Suppl. Resp. to OAI Interrog. No. 3); **Ex. 88** (Pine 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 4, 2026)) 206:6-12, 213:3-

18; **Ex. 68** (Lown Dep. Tr.) 66:10-13, 66:22-69:5. ***CIR:*** Hurst Decl. **Ex. 12** (Bauerlein Dep. Tr. Dated Jan. 27, 2026) 58:16-64:18; **Ex. 12.A** (Bauerlein Dep. Ex. 47); **Ex. 18** (Buckingham 30(b)(6) Dep. Tr. (Vol. 2, dated Dec. 19, 2025)) 405:25-408:22.

148. There are substantial impediments to the formation of a market for licensing data for web grounding LLMs. For example:

a) Because developers cannot predict in advance which specific content and sources will be relevant to any given prompt, a large volume and a large variety of textual data is necessary for grounding general purpose LLMs like those at issue.

Tucker Decl. ¶¶ 13(c), 154-156.

b) No singular piece of data is needed to ground an LLM; any one piece of data used to ground a response can be easily substituted for another.

Tucker Decl. ¶¶ 13(c), 155-157.

c) LLMs generate outputs probabilistically in response to user prompts (including during web grounding); it is difficult to isolate the marginal contribution of an individual piece of data in an LLM's output generation.

Lafferty Decl. ¶¶ 16, 99, 140, 175, 180, 189, 191, 284, 435; Hurst Decl. **Ex. 80** (Murati Dep. Tr.) 108:12-18, 11:13-16, 13:24-14:4, 18:3-14, 114:11-115:6; Tucker Decl. ¶¶ 13(a), 13(c), 155-156, 160-161.

d) Rightsholders for individual works can be difficult and costly to identify and locate, increasing the transaction costs of negotiating any proposed web grounding data licenses.

Tucker Decl. ¶¶ 13(c), 161-165.

e) Negotiations for the necessary volume of data needed for web grounding an LLM

are susceptible to holdout problems.

Tucker Decl. ¶¶ 13(c), 161-165.

149. News Plaintiffs' incentives to create new works have not been reduced following the use of the Asserted Works for web grounding, as demonstrated by their continued publication.

Tucker Decl. ¶¶ 14, 30-48, 237-238; Hurst Decl. **Ex. 116** (Sulzberger Dep. Tr.) 177:24-178:6; **Ex. 12** (Bauerlein 30(b)(6) Dep. Tr.) (Vol. 2, dated Jan. 27, 2026) 423:6-16; **Ex. 59** (Julien Dep. Tr.) 33:12-17.

B. Implied License

150. Microsoft incorporates Undisputed Facts ¶¶ 80-84 as though fully restated herein.

151. Microsoft's blog post announcing the repurposed NOARCHIVE and NOCACHE meta tags further stated that "content with the NOCACHE tag or NOARCHIVE tag will still appear in our search results." Publishers can remain in the Bing Index for use in traditional Bing search, while excluding their content from web grounding.

Canel Decl. ¶¶ 27-28; *see also* Hurst Decl. **Ex. 21** (Canel 30(b)(6) Dep. Tr.) 201:2-20, 202:9-11; **Ex. 21.A** (Canel Dep. Ex. 35 – Microsoft Bing Blog Announcing Meta Tags); **Ex. 85** (Parakhin Dep. Tr.) 226:23–227:24; **Ex. 93** (Procas 30(b)(6) Dep. Tr.) 179:2-181:8, 182:17-183:9; **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 104:9-105:12; **Ex. 34** (Felix Dep. Tr.) 121:23–122:22; **Ex. 64** (Liang Dep. Tr.) 187:10–23; **Ex. 11** (Bauerlein 30(b)(6) Dep. (Vol. 1, dated Jan. 26, 2026)) 171:5–13, 176:25–177:19.

152. Each News Plaintiff employs a webmaster, or equivalent personnel, who holds a Bing Webmaster Tools Account for managing content control for digital publications.

Canel Decl. ¶ 15.

153. Microsoft's announcement via blog post of its repurposing of the NOARCHIVE meta tag for use as an AI-related control was communicated to News Plaintiffs, via for example:

a) Microsoft transmitted a mass notification to all Bing Webmaster Tools account holders when Microsoft's blog post announcing the repurposed NOARCHIVE meta tag was published.

Canel Decl. ¶ 30.

b) Microsoft publicized how it had repurposed the NOARCHIVE meta tag via its post on Microsoft Bing Blogs as well as through industry outreach.

Canel Decl. ¶ 30; *see also* Hurst Decl. **Ex. 85** (Parakhin Dep. Tr.) 279:1-22.

c) Microsoft publicized its blog post featuring the repurposed NOARCHIVE meta tag at webmaster and search engine optimization conferences.

Canel Decl. ¶ 30; *see also* Hurst Decl. **Ex. 99** (Ribas Dep. Tr.) 130:1–131:24; Ex. 85 (Parakhin Dep. Tr.) 279:1-22.

d) Microsoft created a chart summarizing the repurposed meta tags, which was presented directly to The New York Times and to other publishers at an event in Mexico.

Hurst Decl. **Ex. 99** (Ribas Dep. Tr.) 130:1–131:24.

154. After publication of the September 22, 2023 blog post, News Plaintiffs knew about Microsoft's repurposed NOARCHIVE meta tag control, and how to use it on their webpages to signal exclusion from web grounding on News Plaintiffs' content in the Bing Index.

The New York Times: Hurst Decl. **Ex. 34** (Felix Dep. Tr.) 112:17-114:7, 115:6-14, 117:10-119:2, 119:24–120:19, 121:2–19; **Ex. 34.A** (Times Ex. 105); **Ex. 64**

(Liang Dep. Tr.) 172:25–173:12, 176:2–177:16, 181:8–182:21; **Ex. 93** (Procas, 30(b)(6) Dep. Tr.) 179:2–181:8; Canel Decl. ¶ 30. ***Daily News:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***Chicago Tribune:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***Orlando Sentinel:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***Sun Sentinel:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***The Mercury News:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***The Denver Post:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***The Orange County Register:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***Pioneer Press:*** Hurst Decl. **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13-18, 95:23–97:2, 104:3–105:12, 106:3-108:2; Canel Decl. ¶ 30. ***CIR:*** Canel Decl. ¶ 30.

The New York Times Company (The New York Times)

155. With knowledge of Microsoft's repurposed NOARCHIVE meta tag, The New York Times did not uniformly apply Microsoft's repurposed NOARCHIVE meta tag on a sitewide basis, despite internal recommendations to do so.

Hurst Decl. **Ex. 64** (Liang Dep. Tr.) 175:23–177:16, 183:3–17, 190:15–

195:18; **Ex. 64.A** (Liang Dep. Ex. 14); **Ex. 64.B** (Liang Dep. Ex. 16).

156. Until early 2026, The New York Times selectively implemented the NOARCHIVE meta tag control, blocking some webpages but not others (for example the home page, as well as Wirecutter articles).

Hurst Decl. **Ex. 64** (Liang Dep. Tr.) 174:6–175:22; **Ex. 93** (Procas 30(b)(6) Dep. Tr.) 181:20–182:16.

All Daily News Plaintiffs

157. With knowledge of Microsoft's repurposed NOARCHIVE meta tag, the Daily News Plaintiffs did not uniformly apply that meta tag to their news articles until March 7, 2025.

Daily News*:** Hurst Decl. **Ex. 139** (Daily News 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; **Ex. 57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***Chicago Tribune: Hurst Decl. **Ex. 132** (Chicago Tribune 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; **Ex. 57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***Orlando Sentinel***: Hurst Decl. **Ex. 158** (Orlando Sentinel 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; **Ex. 57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***Sun Sentinel***: Hurst Decl. **Ex. 165** (Sun Sentinel 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18,

122:19–129:19, 137:21-138:17, 139:24–140:1; Ex. **57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***The Mercury News***: Hurst Decl. **Ex. 147** (Mercury News 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; **Ex. 57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***The Denver Post***: Hurst Decl. **Ex. 143** (Denver Post 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; **Ex. 57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***The Orange County Register***: Hurst Decl. **Ex. 155** (OC Register 2d Suppl. Resp. to OAI Interrog. No. 6); Ex. **57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; Ex. 57.B (Jacobs Dep. Ex. 31); Ex. **57.B** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18. ***Pioneer Press***: Hurst Decl. **Ex. 162** (Pioneer Press 2d Suppl. Resp. to OAI Interrog. No. 6); **Ex. 57** (Jacobs 30(b)(6) Dep. Tr. (Vol. 1, dated Feb. 12, 2026)) 95:13–18, 122:19–129:19, 137:21-138:17, 139:24–140:1; **Ex. 57.B** (Jacobs Dep. Ex. 31); **Ex. 58** (Jacobs 30(b)(6) Dep. Tr. (Vol. 2, dated Feb. 13, 2026)) 425:7–18.

The Center for Investigative Reporting (CIR)

158. With knowledge of Microsoft's repurposed NOARCHIVE meta tag, CIR decided not to implement that meta tag at all on any of CIR's webpages.

Hurst Decl. **Ex. 11** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 26, 2026)) 176:9–11, 176:14–16.

COUNT II: VICARIOUS COPYRIGHT INFRINGEMENT

III. COUNT II: VICARIOUS COPYRIGHT INFRINGEMENT

159. Microsoft incorporates Undisputed Facts ¶¶ 1-65 as though fully restated herein. Because training an LLM is fair use, there is no direct infringement. Without a direct infringer, there is no vicarious liability.

COUNTS I-II

IV. ALL COUNTS: COMPILATIONS CONSTITUTE ONE WORK FOR PURPOSES OF STATUTORY DAMAGES

160. News Plaintiffs have long registered their newspapers and magazines, including the Asserted Works, as compilations by issue and not as individual articles.

The New York Times: Hurst Decl. **Ex. 93** (Procas 30(b)(6) Dep. Tr.) 32:13-33:10, 38:21-39:10, 45:2-17; **Ex. 277** (NYT_00146884). ***Daily News:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 273** (DNP-00001769). ***Chicago Tribune:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 271** (DNP-00000945). ***Orlando Sentinel:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 276** (DNP-00348298). ***Sun Sentinel:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 272** (DNP-00001523). ***The Mercury News:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 270** (DNP-00000615). ***The Denver Post:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 269** (DNP-00000279). ***The Orange County Register:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78** (Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 275** (DNP-00036164). ***Pioneer Press:*** Hurst Decl. **Ex. 15** (Brewer 30(b)(6) Dep. Tr.) 25:5-12, 57:8-59:21; **Ex. 78**

(Morris 30(b)(6) Dep. Tr. (dated Dec. 17, 2025)) 14:21-15:4; **Ex. 274** (DNP-00002088). ***CIR:*** Hurst Decl. **Ex. 11** (Bauerlein 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 26, 2026)) 70:23-71:14, 73:9-18; **Ex. 279** (Buckingham 30(b)(6) Dep. Tr. (Vol. 3, dated Feb. 27, 2026)) 23:2-25:5; **Ex. 267** (CIR0002550).

Dated: September 4, 2026

Respectfully submitted,

/s/ Annette L. Hurst

ORRICK, HERRINGTON & SUTCLIFFE LLP
Annette L. Hurst (admitted pro hac vice)
Daniel Justice (admitted pro hac vice)
Robert Scott Shwarts
Catherine Lui (admitted pro hac vice)
Amanda H. Schwartz (admitted pro hac vice)
Ana Mendez-Villamil (admitted pro-hac vice)
The Orrick Building
405 Howard Street
San Francisco, CA 94105-2669
Telephone: (415) 773-5700
Facsimile: (415) 773-5759
ahurst@orrick.com
djustice@orrick.com
aschwartz@orrick.com
amendez-villamil@orrick.com

Christopher J. Cariello
Lisa T. Simpson
Marc Shapiro
Vanessa Sorrentino
Hilda Obeng
51 West 52nd Street
New York, NY 10019
Telephone: (212) 506-3778
Facsimile: (212) 506-5151
ccariello@orrick.com
lsimpson@orrick.com
mrshapiro@orrick.com
vsorrentino@orrick.com
hobeng@orrick.com

Sheryl Koval Garko (admitted pro hac vice)
Laura Brooks Najemy (admitted pro hac vice)
222 Berkley Street
Boston, MA 02116
Telephone: (617) 880-1800
Facsimile: (617) 880-1801
sgarko@orrick.com
lnajemy@orrick.com

Isaac Sawlih Behnawa (admitted pro hac vice)
Johnathan Vaknin (admitted pro hac vice)
355 S. Grand Ave. Suite 2700
Los Angeles, CA 90071
Telephone: (213) 612-2179
Facsimile: (213) 612-2499
ibehnawa@orrick.com
jvaknin@orrick.com

Monica Svetoslavov (admitted pro hac vice)
Louis Andrew Schreiber
2100 Pennsylvania Avenue NW
Washington, DC 20037
Telephone: (202) 339-8400
Facsimile: (202) 339-8500
msvetoslavov@orrick.com
aschreiber@orrick.com

FAEGRE DRINKER BIDDLE & REATH LLP
Jeffrey S. Jacobson
1177 Avenue of the Americas
New York, New York 10036
Telephone: (212) 248-3191
jeffrey.jacobson@faegredrinker.com

Jared B. Briant (admitted pro hac vice)
Kirstin L. Stoll-DeBell (admitted pro hac vice)
Timothy Scull (admitted pro hac vice)
Shelby Pickar-Dennis (admitted pro hac vice)
1144 15th Street, Suite 3400
Denver, Colorado 80202
Telephone: (303) 607-3588
jared.briant@faegredrinker.com
kirstin.stolldebell@faegredrinker.com
timothy.scull@faegredrinker.com
shelby.pickar-dennis@faegredrinker.com

Lora A. Brzezynski (admitted pro hac vice)
Brianna Lynn Silverstein (admitted pro hac vice)
Faith Maggard (admitted pro hac vice)
1500 K Street NW, Suite 1100
Washington, DC 20005
Telephone: (202) 230-5000
Facsimile: (202) 842-8465
lora.brzezynski@faegredrinker.com

brianna.silverstein@faegredrinker.com
faith.maggard@faegredrinker.com

Carrie A. Beyer (admitted pro hac vice)
320 South Canal Street, Suite 3300
Chicago, IL 60606-5707
Telephone: (312) 569-1000
Facsimile: (312) 569-3000
carrie.beyer@faegredrinker.com

Elizabeth M.C. Scheibel (admitted pro hac vice)
2200 Wells Fargo Center, 90 S. 7th Street
Minneapolis, MN 55402
Telephone: (612) 766-7000
Facsimile: (612) 766-1600
elizabeth.scheibel@faegredrinker.com

Andrew Lawrence Van Houter
600 Campus Drive
Florham Park, NJ 07932
Telephone: (973) 549-7018
Facsimile: (973) 360-9831
Andrew.vanhouter@faegredrinker.com

Attorneys for Defendant Microsoft Corporation