UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE: OPENAI, INC., COPYRIGHT INFRINGEMENT LITIGATION This document relates to: Case No. 1:23-cv-08292-SHS-OTW Case No. 1:23-cv-10211-SHS-OTW	25-md-3143 (SHS) (OTW) Hon. Sidney H. Stein Hon. Ona T. Wang REDACTED

CLASS PLAINTIFFS' RESPONSE TO MICROSOFT'S RULE 56.1 STATEMENT OF UNDISPUTED MATERIAL FACTS IN SUPPORT OF MICROSOFT'S MOTION FOR SUMMARY JUDGMENT

Pursuant to Rule 56.1 of the Local Rules of the United States District Court for the Southern District of New York, David Baldacci, Taylor Branch, Michael Connelly, Sylvia Day, Jonathan Franzen, Christopher Golden, Andrew Sean Greer, John Grisham, David Henry Hwang, Matthew Klam, George R.R. Martin, Jodi Picoult, Stacy Schiff, James Shapiro, The Authors Guild, Belfry Holdings, Inc., Columbus Rose, Ltd., Daring Greatly Corporation, Gattopardo, LLC, Hieronymus, Inc., Sylvia Day LLC, and Wo & Shade LLC (collectively, "Class Plaintiffs") respectfully submit this response to Defendant Microsoft Corporation's ("Microsoft") statement of material facts as to which Microsoft contends there are no genuine issues to be tried.

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Class Plaintiffs object to the phrase "vast quantities of text" as vague and ambiguous. Subject to and without waiving this objection, **Undisputed** that a large language model ("LLM") is "a machine learning deep neural network" trained on text with the objective of predicting the next token in a sequence. **Disputed** to the extent Microsoft contends that an LLM must be "trained on vast quantities of text." LLMs can be trained on comparatively small amounts of high-quality data like books. Microsoft researchers trained the phi-1 LLM on 7

billion tokens of pre-training data—roughly 42 times fewer than the ~300 billion OpenAI used to train GPT-3—and found that "[b]y intentionally selecting and generating high-quality data," they could "achieve state-of-the-art results on code-generation tasks with a much smaller model and less compute than existing approaches." Nelson Decl. Ex. 937 at 1, 4. A model trained on less data can likewise perform domain-specific tasks when trained on domain-specific data. Nelson Decl. Ex. 938 at 76:18–23.

OpenAI's own records further confirm that data quality, not volume, drives model performance. [redacted]

Further **Disputed** to the extent Microsoft contends that the object of next-token prediction is to generate novel natural language rather than to approximate the text on which the model was trained. [redacted]

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Class Plaintiffs object to the phrase "among the most complex of such domains" as vague and ambiguous, because Microsoft neither identifies the "domains" against which natural language is compared nor supplies any metric for "complex[ity]." Subject to and without waiving that objection, **Disputed** that "[n]atural language is among the most complex of such domains," and **Disputed** that the complexity of a domain determines the volume or diversity of data required to train a model. Zhao Decl. Ex. B ¶¶ 43, 45.

Further Disputed that all machine learning necessarily requires large amounts of data to train. Admissible evidence contradicts this statement. As set forth in greater detail in response to Microsoft's Statement of Fact No. 1, certain smaller models trained on smaller amounts of curated data can outperform larger models trained on larger amounts of lower-quality data.

Further Disputed because the training objective for LLMs is next-token prediction, which means adjusting the parameters of the LLM so that its outputs match the training data as closely as possible. Zhao Decl. Ex. B ¶¶ 43, 45. Memorization and generalization are not opposites, and whatever generalization occurs depends on the model encoding the expression contained in its training data rather than acquiring an understanding of language independent of it. Zhao Decl. Ex. B ¶¶ 14, 27.

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Undisputed that LLMs are capable of natural language processing tasks, such as summarization, translation, content generation, and question answering. **Disputed** that LLMs are capable of "reasoning." LLMs are statistical prediction machines that mimic human reasoning. Zhao Decl. Ex. B ¶¶ 14, 16. [redacted] **Further Disputed** to the extent Microsoft invokes the phrase "general-purpose technology" in a misleading way, as an undisputed economic definition rather than a description of technology. Dkt. 1875-1 McDonald Rebuttal Rep. at 47 (describing the General Purpose Technology Framework is drawn from the economics of growth and innovation literature and is used to explain patterns of aggregate productivity growth but is not meant to be used as an analytical framework for evaluating effects on markets or industries).

Microsoft Statement of Fact No. 4: The process of training an LLM involves:

Class Plaintiffs' Response: This umbrella clause is **Disputed** to the extent that "training" refers to anything other than pretraining a large language model, and is further **Disputed** to the extent it suggests the ensuing sub-paragraphs contain a complete description of the entire model training process. LLM training includes additional phases that the ensuing sub-paragraphs do not describe, including mid-training and post-training. Nelson Decl. Ex. 352 at -422; [redacted]

Microsoft Statement of Fact No. 4(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. ¶¶ 117-122; *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.

Class Plaintiffs' Response: Disputed. Admissible evidence contradicts Microsoft's statement. First, LLMs can be trained without using "vast quantities of text" as set forth in greater detail in response to Microsoft's Statement of Fact No. 1. Second, there was no preexisting practice that permitted the scraping and crawling of content that would then be used to compete against the creators or owners of that content, or in violation of a website's terms of use. Zhao Decl. Ex. B ¶ 100. Third, neither crawling nor scraping of publicly available training data is a "typical" way to acquire training data. *See* Zhao Decl. Ex. C ¶ 36. Rather, the training corpus for an LLM can and does include sources other than web data obtained from crawling. For example, one can license

content from rightsholders or hire humans to write the content needed for pretraining data. Lasinski Decl. Ex. A ¶¶ 98–110. [redacted]

Microsoft Statement of Fact No. 4(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.

Class Plaintiffs' Response: Undisputed that OpenAI converts language data into tokens. **Disputed** to the extent Microsoft contends that tokenization transforms the underlying expression into something other than the text itself. As Class Plaintiffs' expert Ben Zhao explained: "Tokenizers are the statistical models which convert 'bytes' (the format in which text data is stored on a computer) to a set 'vocabulary' of numbers which the LLM can accept as input." Zhao Decl. Ex. B ¶ 35. [redacted]

Microsoft Statement of Fact No. 4(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.

Class Plaintiffs' Response: Undisputed that tokens are represented by numerical vectors reflecting relationships among tokens. **Disputed** to the extent Microsoft contends that those relationships are properties of the training corpus in the aggregate rather than of the individual works within it, as set forth in greater detail in response to Microsoft's Statements of Fact Nos. 10, 11, and 20.

Microsoft Statement of Fact No. 4(d): Processing those vectors through transformer layers repeatedly makes 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. 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.

Class Plaintiffs' Response: Undisputed that the models at issue in this case encode numerical values in a network of parameters or weights. **Disputed** to the extent Microsoft suggests that "[p]rocessing those vectors through transformer layers repeatedly" transforms the underlying expression into something other than the text itself, as set forth in greater detail in response to

Microsoft's Statement of Fact No. 4(b). Further **Disputed** to the extent Microsoft's phrasing suggests that the model adjusts itself. As Dr. Zhao explained, the "training process is simply a computer program that provides the model with input text from the training data, records whether the model's prediction is correct, measures the degree of error if not, and then applies basic calculus to quantify the contributions to the error and alter the model parameters such that the prediction is accurate." Zhao Decl. Ex. B ¶ 58.

Microsoft Statement of Fact No. 4(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.

Class Plaintiffs' Response: Disputed insofar as the statement asserts that the objective of LLM training is generalization. The training objective is next-token prediction, which means adjusting the parameters of the LLM so that its outputs match the training data as closely as possible, and generalization is at most a hoped-for outcome rather than the objective, as set forth in greater detail in response to Microsoft's Statement of Fact No. 9. Zhao Decl. Ex. B ¶¶ 14, 16, 25-27; Zhao Decl. Ex. C ¶ 21; [redacted]

Further Disputed to the extent Microsoft contends Defendants effectively withheld data "during training in order to measure the model's generalization ability." As Class Plaintiffs' expert Ben Zhao explained, "because OpenAI has not properly deduplicated or cleaned its training data—

and given its admitted issues with contamination for in-scope models such as GPT-3—there is no basis to conclude that the use of such 'test' sets during training is reliably measuring generalization." Zhao Decl. Ex. B ¶ 70.

[redacted] *see also* Nelson Decl. Ex. 945 (explaining most benchmarks for large language models are fully public and their questions are widely mirrored across the internet, so there is "systematic risk that models were trained on the very data used to evaluate them.").

Further Disputed that "the training stops." There are additional phases of LLM training, including mid-training and post-training. Nelson Decl. Ex. 352 at -422; *see also* [redacted]

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Disputed that "[t]raining an LLM requires an enormous volume of text." As set forth in greater detail in response to Microsoft's Statement of Fact No. 1, LLMs can be trained without using "an enormous volume of text." Smaller LLMs typically do not

require trillions of tokens. Nelson Decl. Ex. 948 (describing an LLM that was trained on 300 billion tokens). Otherwise, **Undisputed.**

[redacted]

Microsoft Statement of Fact No. 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) at -226; Ex. 197 (MSFT_AI_MDL_000221039) at -042; Ex. 198 (MSFT_AI_MDL_000221643); Ex. 213 (OPCO_MDL_001938121).

Class Plaintiffs' Response: Disputed that "[t]raining an LLM also requires a wide diversity of text" in all circumstances, as set forth in greater detail in response to Microsoft's Statements of Fact Nos. 1 and 2. [redacted]

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Disputed. Admissible evidence contradicts Microsoft's statement. Microsoft's statement wrongly presupposes a strict dichotomy between the internet and other non-internet sources. In reality, model developers can use both internet and non-internet

sources and thus reduce reliance on the internet. For instance, they can hire humans to write training data or acquire large volumes of books through licensing deals. *See* Lasinski Decl. Ex. A ¶¶ 98–110. In addition, the internet is not "singularly" suited to this purpose. As Dr. Zhao explained, "few online sources provide coherent, curated long form text similar to that of published books," which is "critical to performance on LLM tasks that require longer contexts." Zhao Decl. Ex. A ¶ 66.

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Disputed. First, as set forth in greater detail in response to Microsoft's Statement of Fact No. 1, LLMs can be trained without using "an enormous volume" or "wide diversity of textual data."

Second, the training objective for LLMs is next-token prediction, which means adjusting the parameters of the LLM so that its outputs match the training data as closely as possible. Zhao Decl. Ex. B ¶¶ 14, 25-27; Nelson Decl. Ex. 50(a) at 175:6-176:10 [redacted] *see also* Nelson Decl. Ex. 298 at -468.

Third, generalization can be a hoped-for outcome of model training, but it is not the training objective. As Dr. Zhao explains, the models "predict[] text in the distribution of the training corpus[,] [r]ather than engaging in a generalized process of learning language," Zhao Decl. Ex. B ¶ 16, and the "loss function is the quantitative metric which controls parameter updates during

model training[,] . . . measure[ing] how accurately the model measures its training data, not human language as a whole," Zhao Decl. Ex. C ¶ 21.

[redacted]

Fourth, admissible evidence proves that Defendants trained the at-issue LLMs not for an "unlimited range of contexts and tasks", but rather to become capable of book-like outputs. Zhao Decl. Ex. B ¶ 137.

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Disputed. Microsoft identifies no evidence that the capacity to generalize is an "emergent property" of learning across "the complete distribution of the training data" rather than a product of encoding the individual works within it, and the record is to the contrary: what a model retains scales with its capacity rather than converging on structure shared across the corpus. [redacted] **Further Disputed** insofar as the statement asserts that generalization is the objective of model training. As set forth in greater detail in response to Microsoft's Statement of Fact No. 9, the training objective is to match the training data as closely as possible.

Further Disputed that individual works do not contribute to a model's capacity: each individual book used as training data leaves its own imprint on the LLM and contributes to its ability. As Dr. Shan explained, "memorization and generalization are not opposites," "memorizing training examples is mathematically necessary for near-optimal generalization," and "[t]he capacity to memorize also grows with model scale." Shan Decl. Ex. D ¶ 38. **Undisputed** only that no single work, standing alone, is responsible for a model's capacity to generate text.

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Disputed to the extent Microsoft states that removing a category of works from the training corpus would not materially alter the statistical structures the model encodes with respect to that category. [redacted]

That result is consistent with the distinct role books play in model capability. Books are crucial sources of long-form text that enable LLMs to process coherent long-form text as an input and produce it as an output, and they provide expressively varied text that enables LLMs to produce text across a wide array of genres, subjects, and styles. Zhao Decl. Ex. A ¶ 12(a)(iv); *see also id.* ¶ 21 [redacted]

[redacted]

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Disputed insofar as the statement implies there is only one way to "properly train" an LLM. **Undisputed** that an LLM can be "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[.]" **Further Disputed** as to "conversation" and "reasoning," as set forth in greater detail in response to Microsoft's Statement

of Fact No. 3. [redacted]

Microsoft Statement of Fact No. 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, 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.

Class Plaintiffs' Response: Disputed to the extent Microsoft suggests that when an LLM's generalization is "strong" its output will be novel and will not reproduce or regurgitate the data on which it was trained. To the contrary, optimal generalization depends on the model memorizing its training data. Shan Decl. Ex. D ¶ 14 ("models must memorize long-tail training examples to approach optimal performance on naturally-occurring data distributions"); *id.* ¶ 38 ("memorization and generalization are not opposites," and "memorizing training examples is mathematically necessary for near-optimal generalization"); Zhao Decl. Ex. B ¶ 41 (the view that "models do not memorize but rather 'learn' and 'generalize' as humans do is fundamentally inconsistent with the mathematical process at the heart of LLM training: lossy compression").

[redacted]

Microsoft Statement of Fact No. 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.

Class Plaintiffs' Response: Undisputed that models "can at times generate text that was part of their training corpus, a known phenomenon called regurgitation." **Disputed** that "instances

of regurgitation in a properly trained model are extremely rare" and that regurgitation "reflects an unintended side-effect of the training process rather than the intended operation of the model." At the threshold, the models at issue were not "properly trained" within the meaning of Microsoft's assertion: [redacted]

[redacted]

[redacted]

Microsoft Statement of Fact No. 15: Books Plaintiffs' expert Dr. Shan analyzed logs produced by Microsoft of 8.2 million real-world conversations with Microsoft's Copilot in order to assess whether any of those Copilot outputs contained what he qualified as a "regurgitation" of Books Plaintiffs' Asserted Works. Shan's definition of a "regurgitation" was a contiguous near-verbatim block of at least 30 consecutive words (with certain other exclusions).

Class Plaintiffs' Response: Disputed to the extent Microsoft attempts to provide an exhaustive account of Dr. Shan's methodology, including his definition of "regurgitation." Shan Decl. Ex. D § IV. **Further Disputed** to the extent Microsoft suggests that the 8.2 million conversations constitute a representative or complete sample. The logs were selected and produced by Microsoft, and Dr. Shan's findings within that sample are a lower bound on the models' reproduction of the Asserted Works, as set forth in greater detail in response to Microsoft's Statement of Fact No. 15(d). Otherwise, **Undisputed**.

Microsoft Statement of Fact No. 15(a): For 202 of the 212 Asserted Works, Dr. Shan identified no "regurgitations" in the 8.2 million Copilot logs;

Class Plaintiffs' Response: Disputed. Admissible evidence contradicts Microsoft's statement. Microsoft's log production is a sample of Copilot conversations selected by Microsoft, and the absence of a matching output within that sample does not establish that a model is incapable of reproducing a given Asserted Work. Shan Decl. Ex. F ¶ 56 ("Because I understand that the produced logs were not assembled using search terms assembled by the Class Plaintiffs in this case, the set which I searched may itself under-represent regurgitation."). **Undisputed** only that output log samples produced by Microsoft in this case include regurgitations of at least 10 of Plaintiffs' Works, which is a floor rather than a measure of the models' reproduction capacity. Dr. Shan's testing of the models themselves identified 437 distinct verbatim reproductions of Plaintiff-book text drawn from 43 of the 212 Plaintiff books and all 14 Plaintiff authors, and Dr. Shan explained that his log-based search can only understate the true rate, as set forth in greater detail in response to Microsoft's Statement of Fact No. 15(d). Shan Decl. Ex. D ¶ 22; Shan Decl. Ex. F ¶ 56.

Microsoft Statement of Fact No. 15(b): Of the remaining 10 works, out of the 8.2 million conversations, Dr. Shan identified a total of 24 conversations meeting his definition of containing one or more "regurgitations," a rate of .00029%;

Class Plaintiffs' Response: Undisputed that output log samples produced by Microsoft in this case include at least 24 instances of conversations containing regurgitations. **Disputed** that a regurgitation rate can be extrapolated from that sample. **Further Disputed** that .00029% measures the models' propensity to reproduce the Asserted Works, as set forth in greater detail in response to Microsoft's Statement of Fact No. 15(d).

Microsoft Statement of Fact No. 15(c): To arrive at those 24 conversations, Dr. Shan's search code combined the "Text" field, which contains Copilot responses displayed to users, and the "HiddenText" field, which contains web-grounding material not displayed to users; the search did not distinguish between the two;

Class Plaintiffs' Response: Disputed to the extent Microsoft purports to provide an exhaustive account of Dr. Shan's methodology. Shan Decl. Ex. D § IV. **Further Disputed** insofar as Microsoft contends that "web-grounding material not displayed to users" does not represent an instance of the model or inference process regurgitating Class Plaintiffs' Works. *See* Shan Decl. Ex. F ¶ 59. Otherwise, **Undisputed**.

Microsoft Statement of Fact No. 15(d): Shan found no matches of entire books or book chapters when performing his analysis. Lafferty Decl. ¶¶ 366-68, 378; Hurst Decl. Ex. 241 (Shan Opening Report) ¶¶ 40, 156 (24 responses; 10 works; 30-or-more near-verbatim words); Hurst Decl. Ex. 243 (Shan Reply Report) ¶ 59.

Class Plaintiffs' Response: Disputed to the extent Microsoft extrapolates from the sample or uses this statement as a measure of its models' capacity to reproduce the Asserted Works. As Dr. Shan explained: "The search misses genuine reproductions for at least three independent reasons: which conversations I could examine, what I compared them against, and how the logs were sampled. Because of the format of the production and computational limits, I searched only the conversations that named an Asserted Work's author and set the rest aside, so any reproduction in a conversation that did not name the author was never counted. I also matched against OCR-derived copies of the books, whose scanning artifacts can break exact word matches and cause real reproductions to be missed. Because I understand that the produced logs were not assembled using

search terms assembled by the Class Plaintiffs in this case, the set which I searched may itself under-represent regurgitation. Each of these can only lower the count, so the reported rate is a lower bound on how often these models reproduce the books." Shan Decl. Ex. F ¶ 56.

That limitation is inherent in log-based searching rather than particular to Dr. Shan's methodology: researchers studying extraction have found that "[m]uch of the training data that we extract is only discovered when prompting the LM with a particular prefix," so that absent the right prompt, memorized content goes undetected. Nelson Decl. Ex. 950 at 2645. Dr. Shan's testing of the models themselves, as distinct from the produced logs, identified "437 distinct verbatim reproductions of Plaintiff-book text," spanning "more than 44,098 reproduced words in aggregate, drawn from 43 of the 212 Plaintiff books and all 14 Plaintiff authors," with reproductions from 11 of those books exceeding 100 contiguous words. Shan Decl. Ex. D ¶ 22. The longest contiguous passage Dr. Shan extracted was a 1,899-word excerpt from *A Game of Thrones*, which "was reproduced in thousands of independent extraction attempts, ruling out the possibility that the match is a coincidence," and Dr. Shan measured a peak extraction rate of 41% for that work. *Id.* ¶¶ 6, 22. Otherwise, **Undisputed**.

Microsoft Statement of Fact No. 16: 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) 2-3.

Class Plaintiffs' Response: Undisputed that Microsoft provided OpenAI with the Bing Index [redacted]

Microsoft Statement of Fact No. 17: 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.

Class Plaintiffs' Response: Disputed to the extent Microsoft contends that it distributed "portions" of the Bing Index to OpenAI as opposed to the full version, as set forth in greater detail in response to Microsoft's Statement of Fact No. 16.

Further Disputed that the provision of the Bing Index to OpenAI was *solely* a step in the process for the purpose of training LLMs. [redacted]

[redacted]

Microsoft Statement of Fact No. 18: Books Plaintiffs' Asserted Works were created to be read, and are sold and licensed to readers for entertainment, information, and edification. Tucker Decl. ¶¶ 11(b), 48, 52-54; Hurst Decl. Ex. 35 (Franzen Dep. Tr.) 21:18-22:18; Ex. 9 (Baldacci Dep. Tr.) 79:7-14; Ex. 128 (Wolicki Dep. Tr.) 21:1-6, 162:16-163:8.

Class Plaintiffs' Response: Undisputed that the Asserted Works were created and that they are sold and licensed. **Disputed** to the extent Microsoft limits the scope of licenses to "entertainment, information, and edification." Class Plaintiffs who were asked testified that they would consider licensing their works for generative AI training, confirming that the use Defendants made without permission is one for which a market exists and can be licensed. David Baldacci agreed that he would consider licensing his works for "AI training" if someone approached him about it. Nelson Decl. Ex. 920 at 210:19-211:6. Sylvia Day testified that if asked to license her works for training a large language model, she would "absolutely present [the offer] to [her] team and negotiate" and her "primary goal is to monetize [her] work in every way possible." Nelson Decl. Ex. 919 at 157:15-24. John Grisham agreed that he would be "willing to sit down with an artificial intelligence company and license [his] asserted works for training AI models." Nelson Decl. Ex. 925 at 182:16-22. Taylor Branch testified that he would consider an offer to "licens[e] [his] books for use" as AI training data. Nelson Decl. Ex. 921 at 110:4-7. Stacy

Schiff testified that she would be "willing to discuss licensing" her works for AI training. Nelson Decl. Ex. 928 at 195:14-196:2.

Michael Connelly, Christopher Golden, Andrew Sean Greer, George R.R. Martin, and Jodi Picoult gave comparable testimony. Nelson Decl. Ex. 922 at 156:17-25; Nelson Decl. Ex. 923 at 89:19-90:7; Nelson Decl. Ex. 924 at 112:17-22; Nelson Decl. Ex. 926 at 140:13-141:3, 141:24-142:5, 151:24-152:10, 153:1-3; Nelson Decl. Ex. 927 at 77:7-22; *see also* Lasinski Decl. Ex. B ¶ 238 (analyzing authors' willingness to participate in the licensing market, including an Authors Guild survey finding that 95 percent of writers think they should be compensated for use of their works in developing generative AI).

Microsoft Statement of Fact No. 19: 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.

Class Plaintiffs' Response: Undisputed that language data is converted into "tokens" and that LLM models encode numerical values in a network of "parameters" or weights. Otherwise, **Disputed.** The purpose of the training process is not to "gain access to underlying facts, ideas, linguistic patterns, and statistical relationships." Rather, the goal is to adjust the parameters of the LLM so that its outputs match the training data as closely as possible. Zhao Decl. Ex. B ¶¶ 14, 25-27. As set forth in greater detail in response to Microsoft's Statement of Fact No. 13, the models

are "not a generalized understanding independent of the training data[;] [t]hey are a compressed encoding of the training data," and "[w]hen the model generates output that resembles its training data in style, structure, or directly reproduces content, it is decompressing data encoded during training." Zhao Decl. Ex. B ¶¶ 26-27. Moreover, the purpose of training includes tuning the LLM's parameters to encode the specific structure of different types of text, including the precise expression used in those texts and the expressive styles of different writers. Zhao Decl. Ex. A ¶ 12(a)(iv). Further, the LLM training process results in verbatim memorization. Shan Decl. Ex. D ¶¶ 22, 31, 45.

Microsoft Statement of Fact No. 20: The patterns extracted through training—including grammar, syntax, semantic relationships, facts, ideas, and concepts—are properties of the aggregate training distribution rather than of any individual work. Lafferty Decl. ¶¶ 16, 33-34, 45, 66, 80-81, 99, 145-146. 283.

Class Plaintiffs' Response: Disputed. As explained by Class Plaintiffs' experts, LLMs memorize individual works from their training data. Shan Decl. Ex. B ¶ 25 (explaining "the model has memorized the content of the work itself rather than copying or repairing its input"); Zhao Decl. Ex. B ¶ 72. Memorization of training data in an LLM means that during pre-training the LLM strongly associate the tokens making up a block of training text with one another as well as with references to that text. Zhao Decl. Ex. B ¶¶ 16, 26–27. [redacted]

[redacted]

Microsoft Statement of Fact No. 21: Books Plaintiffs are authors whose traditional markets are publication by commercial publishers of books to be read in various formats such as ink on paper, e-books, and audiobooks, as well as the licensing of those commercially published books for translation into other languages and production in other entertainment mediums. Tucker Decl. ¶¶ 12(a), 25, 48, 64.

Class Plaintiffs' Response: **Disputed** to the extent Microsoft limits the scope of licensing to licenses for "translation into other languages and production in other entertainment mediums." Admissible evidence contradicts Microsoft's statement. As set forth in greater detail in response to Microsoft's Statement of Fact No. 18, Class Plaintiffs who were asked testified that they would consider licensing their works for generative AI training, and an Authors Guild survey found that

95 percent of writers think they should be compensated for use of their works in developing generative AI. Lasinski Decl. Ex. B ¶ 238.

Microsoft Statement of Fact No. 22: Books Plaintiffs could not identify any lost sales or licensing opportunities for their Asserted Works as a result of the use of their Asserted Works to develop any of the LLMs at issue. Tucker Decl. ¶¶ 11(a), 20-46; *see also* Hurst Decl. Ex. 35 (Franzen Dep. Tr.) 46:5-8, 90:13-24, 91:1-22; Ex. 87 (Picoult Dep. Tr.) 93:5-16, 99:3-13; Ex. 39 (Golden Dep. Tr. (dated Mar. 17, 2025)) 210:16-20, 300:7-14, 301:20-302:2; Ex. 40 (Golden Dep. Tr. (dated Feb. 10, 2026)) 55:23-57:9; Ex. 128 (Wolicki Dep. Tr.) 148:17-151:20, 188:11-190:1; Ex. 108 (Shapiro Dep. Tr.) 91:15-92:4, 136:22-24, 223:6-17, 236:19-239:3; Ex. 9 (Baldacci Dep. Tr.) 224:20-225:10, 233:3-13, 238:1-15; Ex. 43 (Greer Dep. Tr.) 107:16-108:19, 110:2-19, 113:15-18; Ex. 73 (Martin Dep. Tr.) 216:22-218:6, 219:12-21; Ex. 55 (Hwang Dep. Tr.) 149:8-150:24, 151:16-152:25, 153:21-154:12; Ex. 106 (Schiff Dep. Tr.) 46:23-47:19; Ex. 44 (Grisham Dep. Tr.) 176:22-178:11.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. First, several of the Class Plaintiffs testified that they lost sales as a result of Defendants' unauthorized acquisition and use of Plaintiffs' works in training. Jodi Picoult testified that she lost "at least 27 sales" due to Microsoft's infringement. Nelson Decl. Ex. 927 at 93:5–10. Mary Rasenberger, CEO of the Authors Guild, testified that Defendants' piracy of the Authors Guild's works "were sales that should have happened" and "there is a loss" from Defendants' piracy. Nelson Decl. Ex. 967 at 17:10–16.

Second, several of the Class Plaintiffs testified that they lost the opportunity to license their work to Defendants as a result of Defendants' unauthorized acquisition and use of Plaintiffs' works

in training. David Baldacci testified that since Defendants "didn't come to [him] to license [his] work and they went to a pirate website, [he] lost the ability to license [his] work" and "lost those sales." Nelson Decl. Ex. 5 at 221:16–25. John Grisham testified that when Defendants "took [Plaintiffs'] protected property without approval, without compensation, without licensing and used it for their LLMs, [he] lost money, as did every writer." Nelson Decl. Ex. 925 at 180:17–22. Sylvia Day confirmed that she lost a license to Microsoft Copilot because Microsoft used her works without placing a license "there [was] no license in front of [her]." Nelson Decl. Ex. 919 at 174:15–18. Mary Rasenberger, the Authors Guild corporate representative, testified that the Authors Guild "lost licensing income" for its works that defendants used as AI training data. Nelson Decl. Ex. 967 at 19:15–24.

Microsoft Statement of Fact No. 23: For 12 Asserted Works published in or after January 2017 (the most relevant time period), analysis of royalty statements and sales data shows the trend in unit sales and royalty earnings did not decline following the introduction of Copilot and ChatGPT; the observed trends reflect ordinary product-lifecycle dynamics. Tucker Decl. ¶¶ 11(a) 24-37.

Class Plaintiffs' Response: **Disputed**. Admissible evidence contradicts Microsoft's statement. First, the royalty statements for several of the 12 Asserted Works Dr. Tucker identified in paragraph 11(a) of her declaration show a decline in royalties after the introduction of ChatGPT in November 2022 and the introduction of Copilot in 2023. Royalties for Andrew Sean Greer's novel *Less*, for example, declined from $99,780.85 in 2022 to $45,065.83 in the first half of 2023. Nelson Decl. Ex. 968 at -334; Nelson Decl. Ex. 969 at -407. The *Less* royalties then decreased from $45,065.83 in the first half of 2023 to $32,213.40 in the second half of 2023. *Id.* at -407; Nelson Decl. Ex. 970 at -457.

[redacted]

[redacted]

Microsoft Statement of Fact No. 24: For 125 Asserted Works published prior to January 1, 2017, the trend in sales also did not decline following the introduction of Copilot and ChatGPT. Tucker Decl. ¶¶ 11(a), 38-40.

Class Plaintiffs' Response: Disputed. [redacted]

[redacted]

Microsoft Statement of Fact No. 25: The only existing and specific economic harm Books Plaintiffs have identified is the loss of payment for the single copy of each work used in LLM training. Hurst Decl. Ex. 44 (Grisham Dep. Tr.) 89:8-21; Ex. 106 (Schiff Dep. Tr.) 44:18-25, 45:11-46:6; Ex. 40 (Golden Dep. Tr. (dated Feb. 10, 2026)) 55:23-56:9, 58:6-13, 64:4-7; Ex. 87 (Picoult Dep. Tr.) 99:14-20; Ex. 128 (Wolicki Dep. Tr.) 149:22-150:5, 174:8-18; Ex. 9 (Baldacci Dep. Tr.) 223:3-14.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. Plaintiffs have identified other forms of economic harm, including market substitution and displacement of book sales. Plaintiffs' expert Michael Lasinski explained that authors are harmed by "the entry of a large quantity of new, competing content directly into the marketplace where Plaintiffs' works are sold," that the harm is amplified as "the introduction of AI-generated content meets or exceeds consumer acceptance levels," "thereby introducing a level of competition that did not exist before," and that "works can be created using AI quickly at lower cost and have lower prices than human-created works, on average," increasing "the likelihood of pricing pressure relative to copyrighted books." Lasinski Decl. Ex. B ¶ 411; *see also* Lasinski Decl. Ex. C ¶ 84 ("[Q]uality-driven harms and volume-driven harms – operate simultaneously in the marketplace for books, including copyrighted books, thereby providing adverse effects on authors of copyrighted books."); Seeley Decl. Ex. B ¶¶ 50-51 ("[G]enerative AI lowers barriers to entry in the market for copyrighted books and leads to a proliferation of AI content," and noting that Dr. Shan "generated more than 13,000 multi-chapter books in the span of two weeks using OpenAI's model").

Plaintiffs also identified harm in the form of lost licensing revenues that they otherwise would have received if Defendants had lawfully licensed their works for training and other purposes. Mr. Lasinski detailed "an emerging licensing market for acquiring access and use rights to data and content, including books, for training and/or output [which] has grown in scale and complexity." Lasinski Decl. Ex. B ¶ 53; *see also* Lasinski Decl. Ex. A ¶ 14 ("Defendants' alleged wrongful conduct and infringement of the Copyrighted Works has negatively impacted the Plaintiffs in the form of lost revenues from . . . foregone licensing opportunities".). The viability of the licensing market for AI training and related purposes is "supported by Microsoft's own data

initiatives and executed strategies, past practices and contract amendments, analyses related to content/data priorities and pricing, establishment of a licensing framework for negotiations, and real-world execution of data access and use agreements[.]" Lasinski Decl. Ex. B ¶¶ 104–13.

Plaintiffs identified further harm in the form of obstacles to authors' ability to monetize their works. Class Plaintiffs' expert Mark Seeley noted that researchers at the University of Cambridge found that "[o]ver a third (39%) of novelists reported that their income has already been negatively impacted by GenAI," which "was mostly attributed to competition from AI-generated books and loss of jobs which provide supplementary streams of income"; that "85% of published novelists said they think their future income will be negatively impacted by GenAI"; and that "over half (51%) of published novelists agreed that AI is likely to displace their work entirely." Seeley Decl. Ex. B ¶ 37 (citing Dr. Clementine Collett, *The Impact of Generative AI on the Novel*, at 8 (Nov. 2025), available at: https://www.mctd.ac.uk/wp-content/uploads/2025/11/MCTD-AIAndTheNovel-Report-Web.pdf). Likewise, Class Plaintiffs testified that AI-generated books and summaries have reduced demand for their works. Nelson Decl. Ex. 54 at 90:18–91:5.

Plaintiffs also identified reputational harm stemming from Defendants misconduct. Ms. Picoult testified, for example, that AI-generated outputs that are based on her style or use her characters, or which are inaccurate derivatives of her original works, can damage her reputation as an author. Nelson Decl. Ex. 927 (Picoult Tr.) at 207:5–208:12.

Microsoft Statement of Fact No. 26: Even assuming "market dilution" is relevant to the fourth factor:

Class Plaintiffs' Response: Class Plaintiffs object to the phrase "market dilution" as vague and ambiguous, and further object to this statement to the extent it purports to state a legal conclusion regarding the scope of the fourth fair use factor. Subject to and without waiving those objections, **Disputed.** Class Plaintiffs incorporate by reference their responses to Microsoft's Statements of Fact Nos. 25 and 26(a) through 26(d). In addition, admissible evidence contradicts Microsoft's assertion. As Mr. Lasinski explains, the harm to authors operates through two concurrent market mechanisms—volume-driven entry of competing content and quality-driven consumer acceptance of that content—and both are measurable on this record. Lasinski Decl. Ex. B ¶¶ 410–11. The 2026 Reimers and Waldfogel study [redacted] demonstrate "that there has been a significant increase in the supply of books in the market, driven specifically by the proliferation of AI-generated content," *id.* ¶ 410, [redacted]

Microsoft Statement of Fact No. 26(a): Books Plaintiffs have identified no instance in which an AI-generated book displaced a sale of any of Books Plaintiffs' commercially published Asserted Works; Hurst Decl. Ex. 35 (Franzen Dep. Tr.) 90:13-24; Ex. 108 (Shapiro Dep. Tr.) 90:15-92:4; Ex. 43 (Greer Dep. Tr.) 107:16-108:19; Ex. 40 (Golden Dep. Tr. (dated Feb. 10, 2026)) 47:21-48:6; Ex. 55 (Hwang Dep. Tr.) 150:10–24; Ex. 4 (Amer Dep. Tr.) 250:25–251:8; Ex. 9 (Baldacci Dep. Tr.) 231:5–12; Ex. 44 (Grisham Dep. Tr.) 187:8–13; Ex. 87 (Picoult Dep. Tr.) 94:16–23; Ex. 14 (Branch Dep. Tr.) 73:11–74:1; Ex. 73 (Martin Dep. Tr.) 219:13–21.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. Class Plaintiff David Baldacci testified that on the same day he released his book *Nash Falls*, AI-generated summaries of that book were sold "for a lower price" than his book. Nelson Decl. Ex. 5 at 227:3–228:6. Likewise, James Shapiro testified that the "availability of these [AI-generated] summaries makes it easy for students who do not want to buy a recommended text, to shortcut that and not purchase my material, my book, my copyrighted material, but to allow you, through theft of my work, to profit from this. I think that does me significant harm." Nelson Decl. Ex. 54 at 90:18–91:5.

Further, Mr. Lasinski opines that market displacement occurs by way of at least two concurrent market mechanisms. Lasinski Decl. Ex. B ¶ 411.

Microsoft Statement of Fact No. 26(b): Few if any people are using Copilot for the purpose of generating full books; Amir Decl. ¶¶ 9(b), 47-51; Bail Decl. ¶ 113; Tucker Decl. ¶ 12(d), 83-88; *see also* Hurst Decl. Ex. 40 (Golden Dep. Tr. (dated Feb. 10, 2026)) 62:17-63:2; Ex. 60 (Kazi Dep. Tr.) 143:19-144:1; Ex. 2 (Albu Dep. Tr.) 206:23-25; Ex. 43 (Greer Dep. Tr.) 145:12-15; Ex. 97 (Rasenberger Dep. Tr. (dated Feb. 12, 2026)) 282:18–23; Ex. 40 (Golden Dep. Tr. (dated Feb. 10, 2026)) 62:17-63:10; Ex. 4 (Amer Dep. Tr.) 252:17-25.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. [redacted]

[redacted]

Further **Disputed** because the analyses Microsoft cites were not designed to measure book-length generation and cannot support the conclusion Microsoft draws from them. Dr. Bail classified Copilot conversations against a fixed taxonomy of fifteen user-intent categories, and Microsoft identifies no category within that taxonomy corresponding to the generation of a full book. Bail Decl. ¶ 113. A classification scheme that defines no category for book-length output cannot establish that "[f]ew if any people are using Copilot for the purpose of generating full books." Nor can Dr. Amir's survey bear that weight: it measured only the task descriptions put to respondents, and even within those constraints respondents reported using Copilot to "[w]rite a story" and "create stories." Amir Decl. ¶¶ 47-51; Amir Decl. App'x H.1.

Microsoft Statement of Fact No. 26(c): AI-generated books are viewed with disfavor by consumers; Tucker Decl. ¶¶ 12(b), 70-76, 87; *see also* Hurst Decl. Ex. 4 (Amer Dep. Tr.) 250:25-251:8; Ex. 87 (Picoult Dep. Tr.) 98:24-99:13; Ex. 35 (Franzen Dep. Tr.) 90:21-24; Ex. 55 (Hwang Dep. Tr.) 152:5-25; Ex. 44 (Grisham Dep. Tr.) 176:22-177:12; Ex. 43 (Greer Dep. Tr.) 108:16-19; Ex. 40 (Golden Dep. Tr. (dated Feb. 10, 2026)) 47:21-49:24; Ex. 60 (Kazi Dep. Tr.) 287:4-24; Ex. 97 (Rasenberger Dep. Tr. (dated Feb. 12, 2026)) 119:24-121:17.

Class Plaintiffs' Response: **Disputed.** Class Plaintiffs object to the phrase "viewed with disfavor by consumers" as vague and ambiguous, and further object that the cited deposition testimony reflects individual authors' opinions about consumer preferences rather than evidence of consumer behavior. Subject to and without waiving those objections, **Disputed.** Admissible evidence contradicts Microsoft's statement. Consumers cannot reliably identify AI-generated

writing: a 2023 study cited by Dr. Shan "showed that humans failed to identify AI writing approximately 50% of the time, and this study reflected text generated by LLMs weaker than, e.g., GPT-4o." Shan Decl. Ex. D ¶ 63. Consumer behavior likewise contradicts any claimed disfavor:

[redacted]

Mr. Lasinski explains, harm operates through a quality-driven mechanism precisely because "the introduction of AI-generated content meets or exceeds consumer acceptance levels," such that those works "can more directly compete with Plaintiffs' works, thereby introducing a level of competition that did not exist before." Lasinski Decl. Ex. B ¶ 411; *see also* Seeley Decl. Ex. B ¶¶ 50-51.

Microsoft Statement of Fact No. 26(d): Books Plaintiffs concede they do not expect AI to replace them. Tucker Decl. ¶¶ 45-46, 68-69; Hurst Decl. Ex. 25 (Connelly Dep. Tr.) 90:24-91:6; 160:3-13; Ex. 35 (Franzen Dep. Tr.) 90:21-24; Ex. 55 (Hwang Dep. Tr.) 152:5-25; Ex. 44 (Grisham Dep. Tr.) 176:22-177:12; Ex. 43 (Greer Dep. Tr.) 108:16-19; Ex. 87 (Picoult Dep. Tr.) 22:10-15; 98:24-99:13; Ex. 55 (Hwang Dep. Tr.) 150:10–24; Ex. 9 (Baldacci Dep. Tr.) 231:5–12.

Class Plaintiffs' Response: **Undisputed** only that no Class Plaintiff testified that he or she expects artificial intelligence to replace authors entirely. **Disputed** to the extent Microsoft contends that the absence of such testimony bears on market harm, which does not require wholesale replacement. Admissible evidence contradicts any suggestion that Class Plaintiffs do not expect AI to displace sales of and demand for their works. For example, Class Plaintiff John Grisham testified that "there is very real risk that we could be eliminated or seriously damaged by the flood of cheap books, legal thrillers." Nelson Decl. Ex. 24 at 175:1-17; *see also* Nelson Decl. Ex. 931 at

42:15–43:15 (testifying that "AI is producing, potentially long-term, millions of competitive works"); Nelson Decl. Ex. 927 at 53:19–54:14 (testifying that defendants' products allow generation of "more stuff that benefits them commercially but not me"); Nelson Decl. Ex. 54 at 67:11-20 (testifying that ChatGPT poses a threat to his future livelihood); Nelson Decl. Ex. 51 at 44:4-17 (testifying that she believed her future books were less likely to succeed).

Microsoft Statement of Fact No. 27: No agreements for licensing books as LLM training data existed at the time that the LLMs at issue were trained. Tucker Decl. ¶¶ 13(a), 90-91; *see also* Hurst Decl. Ex. 254 (2025.11.21 Hwang Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 255 (2025.11.21 Hieronymous Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 257 (2025.11.21 Golden Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 258 (2025.11.21 Gattopardo Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 259 (2025.11.21 Daring Greatly Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 293 (2025.11.21 Columbus Rose Supp. Resps. To Microsoft's 1st Set of Interrogs.) 7-8; Ex. 260 (2025.11.21 Belfrey Holdings Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 261 (2025.11.21 Wo & Shade LLC Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 253 (2025.11.21 Sylvia Day LLC Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 256 (2025.11.21 Greer Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. [redacted]

[redacted]

[redacted] ¶ 199 (referencing "[a] 2017 paper titled 'Artificial Intelligence's Fair Use Crisis' not[ing] that 'an established market for training data already exists.'").

Microsoft Statement of Fact No. 27(a): Books Plaintiffs have not offered any evidence that they were approached by any LLM developers to license their works for training; Hurst Decl. Ex. 43 (Greer Dep. Tr.) 112:24-113:3; Ex. 128 (Wolicki Dep. Tr.) 157:16-18, 211:13-16; Ex. 106 (Schiff Dep. Tr.) 195:19-23; Ex. 14 (Branch Dep. Tr.) 108:21-109:6; Ex. 9 (Baldacci Dep. Tr.) 214:16-215:2; Ex. 87 (Picoult Dep. Tr.) 165:10-15; Ex. 44 (Grisham Dep. Tr.) 184:17-24; Ex. 254 (2025.11.21 Hwang Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 255 (2025.11.21 Hieronymous Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 257 (2025.11.21 Golden Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 258 (2025.11.21 Gattopardo Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 259 (2025.11.21 Daring Greatly Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 293 (2025.11.21 Columbus Rose Supp. Resps. To Microsoft's 1st Set of Interrogs.) 7-8; Ex. 260 (2025.11.21 Belfrey Holdings Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 261 (2025.11.21 Wo & Shade LLC Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 253 (2025.11.21 Sylvia Day LLC Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 256 (2025.11.21 Greer Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8.

Class Plaintiffs' Response: Undisputed that Plaintiffs did not testify that they were directly approached by any LLM developers to license their works for training. **Disputed** to the extent Microsoft's statement is intended to suggest that Defendants did not desire or value

Plaintiffs' books for purposes of developing LLMs. [redacted]

[redacted] and Dario Amodei described acquiring books and papers through proper channels as "a huge legal/practice/business slog." *Id.* ¶¶ 107 n.523, 221 (citing Nelson Decl. Ex. 999 at -113).

Microsoft Statement of Fact No. 27(b): Books Plaintiffs have not offered any evidence that they sought to license their works for LLM training. Hurst Decl. Ex. 43 (Greer Dep. Tr.) 112:24-113:3; Ex. 106 (Schiff Dep. Tr.) 195:19-23; Ex. 98 (Rasenberger 30(b)(6) Dep. Tr.) 169:2-17; Ex. 14 (Branch Dep. Tr.) 108:25-109:6; Ex. 254 (2025.11.21 Hwang Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 255 (2025.11.21 Hieronymous Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 257 (2025.11.21 Golden Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 258 (2025.11.21 Gattopardo Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 259 (2025.11.21 Daring Greatly Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 293 (2025.11.21 Columbus Rose Supp. Resps. To Microsoft's 1st Set of Interrogs.) 7-8; Ex. 260 (2025.11.21 Belfrey Holdings Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 261 (2025.11.21 Wo & Shade LLC Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8; Ex. 253

(2025.11.21 Sylvia Day LLC Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 6-8; Ex. 256 (2025.11.21 Greer Suppl. Resps. to Microsoft's 1st Set of Interrogs.) 7-8.

Class Plaintiffs' Response: Undisputed that Plaintiffs did not testify that they directly sought to license their works for AI training. **Disputed** to the extent Microsoft's statement is intended to suggest Plaintiffs' unwillingness to do so or a lack of a licensing marketplace. As set forth in greater detail in response to Microsoft's Statement of Fact No. 18, Plaintiffs testified that they would be willing to enter into a license for generative AI model training, and an Authors Guild survey found that 95 percent of writers think they should be compensated for use of their works in developing generative AI. Lasinski Decl. Ex. B ¶ 238 & n.861. Plaintiffs would have entered into such a license agreement had Defendants sought to negotiate and pay a fair price rather than take their works without permission. Numerous Class Plaintiffs have also participated in the settlement in *Bartz et al. v. Anthropic PBC*, No. 4:24-cv-05417-AMO (N.D. Cal.), further demonstrating their willingness to be compensated for the use of their works as AI training data. *See, e.g.*, Nelson Decl. Ex. 925 at 103:12–104:5; Nelson Decl. Ex. 931 at 53:20–54:15; Nelson Decl. Ex. 920 at 215:16–218:12.

Microsoft Statement of Fact No. 28: 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.

Class Plaintiffs' Response: **Disputed.** Class Plaintiffs incorporate by reference their responses to Microsoft's Statements of Fact Nos. 28(a) through 28(g) below. In addition, admissible evidence contradicts Microsoft's assertion. *See* Lasinski Decl. Ex. B ¶¶ 32, 159–167, 199 (explaining "the book marketplace, like marketplaces for other copyrighted works, possesses

the institutional and contractual infrastructure that has enabled and continues to enable, adaptation to technological change");

Microsoft Statement of Fact No. 28(a): For example: 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)) 26:6-22, 35:18-36:11, 16:17-22, 17:1-23, 19:12-20:1, 23:24-24:25 25:15-23; Ex. 117 (Sutskever Dep. Tr.) 28:5-29:15, 67:9-15, 255:12-256:2.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement.

[redacted]

Further, Class Plaintiffs' expert, Dr. Ben Zhao, explained methods by which the value of LLM training data could be isolated. *See, e.g.*, Zhao Decl. Ex. A ¶ 64 (describing "a methodology known as machine unlearning," which has been used to identify "books as high-impact data with disproportionate influence on model capabilities relative to their share of training tokens"). [redacted]

Microsoft Statement of Fact No. 28(b): The Asserted Works were a minuscule share of the training corpus used to train the GPT models; Tucker Decl. ¶¶ 13(b), 95, 99-100; Lafferty Decl. ¶¶ 94-97, 105-109.

Class Plaintiffs' Response: **Disputed.** Class Plaintiffs object to the phrase "minuscule share" as vague and ambiguous. Subject to and without waiving that objection, **Disputed** because Microsoft offers no evidence quantifying the share of any training corpus attributable to the Asserted Works, and because share of tokens does not measure contribution: copyrighted books, including the Asserted Works, were an especially important part of the training corpus used to train the GPT models.

[redacted]

Microsoft Statement of Fact No. 28(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(c), 111-114.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. Defendants' own experts opined that pricing training data is workable. [redacted]

Microsoft Statement of Fact No. 28(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(c), 89, 101-106; Usdan Decl. ¶¶ 10-11, 22, 24.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. The purported difficulties are not especially hard to surmount. [redacted]

Microsoft Statement of Fact No. 28(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(c), 102-106; Usdan Decl. ¶ 24.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. As explained by Class Plaintiffs' expert Michael Lasinski, "the traditional book market has repeatedly expanded its licensing frameworks to accommodate innovations such as e-books and audiobooks." Lasinski Decl. Ex. B ¶¶ 12, 28–29, 159–167. [redacted]

Plaintiffs in *Bartz et al. v. Anthropic PBC*, No. 4:24-cv-05417-AMO (N.D. Cal.) were able to locate hundreds of thousands of rightsholders during the settlement administration process, resulting in a claims rate of nearly 93 percent as of April 16, 2026. Lasinski Decl. Ex. B ¶ 157 & n.580. Public registries, like Amlet, also exist "allowing users to recognize a work [and] verify the rights holders," and "when for-profit enterprises have attempted to create a new revenue stream for book copyright holders, all parties to the licensing negotiation are able to efficiently identify themselves and reach agreement." *Id.* ¶¶ 68, 167.

Microsoft Statement of Fact No. 28(f): Negotiations for the aggregation of the necessary volume and variety of data are susceptible to holdout problems; Tucker Decl. ¶¶ 13(c), 107-110.

Class Plaintiffs' Response: **Disputed.** Admissible evidence contradicts Microsoft's statement. Class Plaintiffs' expert Michael Lasinski explains that "the Tucker Report's discussion of holdout problems is speculative, as she has not identified any alleged instances of holdout related to the licensing market for copyrighted books for LLM training data. Indeed, the facts of

this case demonstrate that licensors are willing participants in the market for licensing or accessing data for LLM training." Lasinski Decl. Ex. B ¶ 197; *see also* Lasinski Decl. Ex. A Attachments 6.0, 6.1 (listing Microsoft's and OpenAI's executed agreements).

[redacted]

Admissible evidence contradicts Microsoft's statement. [redacted]

[redacted]

[redacted]

[redacted]

Microsoft Statement of Fact No. 29: After the release of the LLMs at issue in this case, Books Plaintiffs have continued to write and publish books. Tucker Decl. ¶¶ 14, 151; Hurst Decl. Ex. 128 (Wolicki Dep. Tr.) 26:20-27:3, 30:14-20; Ex. 87 (Picoult Dep. Tr.) 166:14-167:4; Ex. 108 (Shapiro Dep. Tr.) 46:1-12; Ex. 106 (Schiff Dep. Tr.) 20:16-21:6; Ex. 35 (Franzen Dep. Tr.) 61:14-18; Ex. 44 (Grisham Dep. Tr.) 178:21-25.

Class Plaintiffs' Response: Undisputed that Class Plaintiffs have continued to write and publish books. **Disputed** to the extent Microsoft contends that continued publication demonstrates the absence of harm. As set forth in greater detail in response to Microsoft's Statement of Fact No. 25, Class Plaintiffs have identified market substitution and displacement of book sales, lost licensing revenues, obstacles to their ability to monetize their works, and reputational harm, none of which is negated by continued authorship.

Microsoft Statement of Fact No. 30: Microsoft did not acquire the Books1 or Books2 datasets used by OpenAI as part of the training corpus in connection with the development of GPT 3 and 3.5. Hurst Decl. Ex. 120 (Tinter 30(b)(6) Dep. Tr.) 166:23-169:17, 13:6-22:18, 29:24-30:22; Ex. 120.A (Tinter Dep. Ex. 3 - Joint Development and Collaboration Agreement) 4-5 (§§1(f), 1(q)); Ex. 85 (Parakhin Dep. Tr.) 79:3-18, 15:18-21, 16:1-5, 18:7-12, 22:25-23:8, 23:21-25:6, 26:12-27:4; Ex. 118 (Templeton 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 20, 2026)) 62:4-25, 296:12-18, 13:14-16, 14:14-20, 15:2-12, 16:18-19:10, 22:1-14, 22:25-23:15, 24:2-27:16; Ex. 96 (Radford

Dep. Tr.) 37:17-41:11, 46:3-13, 117:4-20, 11:15-12:10, 13:22-14:6, 16:3-8, 17:13-19, 20:25-24:22; Ex. 295 (Mann Dep. Tr. (dated Mar. 3, 2025)) 58:14-60:4, 65:22-66:11, 13:19-20:5, 20:21-21:21; Ex. 48 (Hallacy Dep. Tr. (Vol. 1, dated Sept. 30, 2025)) 91:10-12, 92:5-18, 9:3-13, 9:23-10:9, 11:9-22, 29:16-25, 33:22-34:18; Ex. 104 (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 95:10-23, 168:1-11, 169:15-170:17, 9:18-21, 15:2-25, 16:7-11, 17:2-18:7; Ex. 84 (Paino 30(b)(6) Dep. Tr.) 15:12-21:24, 22:12-26:6, 216:11-217:22, 12:18-25, 13:3-6, 13:22-14:7.

Class Plaintiffs' Response: **Disputed.** [redacted]

Microsoft Statement of Fact No. 31: Microsoft did not direct OpenAI to download the Books1 or Books2 datasets from Library Genesis ("LibGen") or any other so-called "shadow library." Hurst Decl. Ex. 120 (Tinter 30(b)(6) Dep. Tr.) 166:23-169:17; Ex. 120.A (Tinter Dep. Ex. 3 – Joint Development and Collaboration Agreement) 4-5 (§§1(f), 1(q)); Ex. 85 (Parakhin Dep. Tr.) 26:12-27:4, 79:3-18; Ex. 118 (Templeton 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 20, 2026)) 62:4-25, 296:12-18; Ex. 96 (Radford Dep. Tr.) 37:17-41:11, 46:3-13, 117:4-20; Ex. 295 (Mann Dep. Tr. (dated Mar. 3, 2025)) 58:14-60:4, 65:22-66:11, 13:19-20:5, 20:21-21:21; Ex. 48 (Hallacy Dep. Tr. (Vol. 1, dated Sept. 30, 2025)) 91:10-12, 92:5-18; Ex. 104 (Ryder 30(b)(6) Dep. Tr. (Vol. 1, dated Jan. 29, 2026)) 95:10-23, 168:1-11; 169:15-170:17; Ex. 84 (Paino 30(b)(6) Dep. Tr.) 15:12-21:24, 22:12-26:6, 216:11-217:22.

Class Plaintiffs' Response: **Disputed** to the extent Microsoft contends that it was unaware that OpenAI's training data was sourced from Library Genesis or that it did not facilitate the acquisition of LibGen. [redacted]

[redacted]

Microsoft Statement of Fact No. 32: To the extent Microsoft acquired data to be used by OpenAI in connection with the development of LLMs, Microsoft did so by using ordinary software robots that crawled content from URLs across the public internet; Microsoft's web robots were not capable of utilizing the Torrent protocol implemented by so-called "shadow libraries." Gaggara Decl. ¶¶ 20-21; Hurst Decl. Ex. 38 (Gaggara 30(b)(6) Dep. Tr. (dated Aug. 20, 2025)) 67:1-20, 143:13-144:6, 8:9-19, 9:25-13:17.

Class Plaintiffs' Response: Class Plaintiffs object to the phrase "ordinary software robots" as vague and ambiguous. Subject to and without waiving this objection, **Undisputed** that Microsoft's web crawlers did not themselves implement the Torrent protocol. **Disputed** to the extent Microsoft contends that its crawling was confined to ordinary public-internet content or that the limits of its crawlers establish that Microsoft did not acquire or host data sourced from shadow libraries.

[redacted]

[redacted]

Microsoft Statement of Fact No. 33: Microsoft incorporates Undisputed Facts 1-32 as though fully restated herein.

Class Plaintiffs' Response: **Disputed.** This is not an appropriate undisputed fact. To the

extent that the Court considers this "fact," Class Plaintiffs incorporate by reference their responses to Microsoft's Undisputed Facts Nos. 1-32 as though fully restated herein.

Microsoft Statement of Fact No. 34: Microsoft incorporates Undisputed Facts 1-33 as though fully restated herein.

Class Plaintiffs' Response: **Disputed.** This is not an appropriate undisputed fact. To the extent that the Court considers this "fact," Class Plaintiffs incorporate by reference their responses to Microsoft's Undisputed Facts Nos. 1-33 as though fully restated herein.

Dated: October 23, 2026

/s/ Justin A. Nelson

Justin A. Nelson
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, TX 77002
Telephone: 713.651.9366
jnelson@susmangodfrey.com
Interim Lead Counsel for Class Plaintiffs

CERTIFICATE OF SERVICE

I hereby certify that on October 23, 2026, I caused the foregoing document to be electronically filed with the Clerk of the United States District Court for the Southern District of New York using the CM/ECF system, which shall send electronic notification to all counsel of record.

/s/ Tyler Alabanza-Behard
Tyler Alabanza-Behard