May 8, 2025

The Honorable Coke M. Stewart
Acting Director and Under Secretary of Commerce for Intellectual Property
United States Patent and Trademark Office
600 Dulany Street
Alexandria, VA 22314

The Honorable Michael J. K. Kratsios
Director of the Office of Science and Technology
Executive Office of the President
1650 Pennsylvania Ave NW
Washington, DC 20504

BSA Submission on Artificial Intelligence and Copyright

Dear Acting Director Stewart and Director Kratsios,

The Business Software Alliance ("BSA") shares the attached submission with the US Patent & Trademark Office ("USPTO") and the Office of Science and Technology Policy ("OSTP") on the subject of Artificial Intelligence ("AI") and copyright. The purpose of this written submission is to provide additional context on this important subject at a time of heightened interest in these issues.

I. About BSA

BSA is the leading advocate for the global software industry before governments and in the international marketplace.[1] Representing the United States' leading patent and copyright holders in an industry that supports over $100 billion in Research & Development ("R&D") (i.e., 27% of all private sector R&D), BSA supports both strong copyright protection and continued US leadership in AI. Holding hundreds of thousands of patents and valuable copyrights, BSA members invest heavily in the US intellectual property ("IP") system and recognize that US resilience and strength in innovation, science, and technology are critical to our national and economic security.

As pioneers in the development of enterprise AI, BSA members have unique insights into this technology's tremendous potential to spur digital transformation[2] and support IP and innovation.[3] As enterprise software publishers, BSA members occupy a unique position as both copyright licensors and copyright licensees; as both IP holders and innovators; and as both AI developers and deployers. Our comments follow.

II. US AI Leadership is Critical to the US National Interest

US leadership in AI is critical to US national defense, economic security, and technological leadership. Statements by US political, military, economic, and academic leaders reflect this consensus view. For example, as stated by the White House:

- "Continued American leadership in Artificial Intelligence is of paramount importance to maintaining the economic and national security of the United States."[4]

- "The United States of America is the leader in AI, and our administration plans to keep it that way";[5] and

- Recent AI advances "from Chinese compan[ies]… should be a wakeup call for our industries that we need to be laser-focused on competing to win."[6]

III. US AI Leadership Requires Access to Large and Diverse Data Sets

Access to data is a critical factor in safeguarding US AI leadership, in line with the White House AI Action Plan. Access to a sufficiently large *corpus* of factual and mathematical variables ensures that computational analysis of training data[7] can reliably identify correlations, patterns, or other metadata to develop a model that can make predictions or recommendations with minimal risk of inaccuracy or error.[8]

Data used to train AI may include everything from machine-to-machine data (e.g., satellite transmission data) and international trade statistics to published materials, blog posts, and website comments. "Raw data" is often not in a form that is suitable for AI training, such that it must be transformed through a resource-intensive procedure to create a semantic structure that the machine can process. Preparing data for structured computational analysis helps address quality challenges, such as missing values, duplicates, outliers, and inconsistent formatting across the entire data set. Training data will often ultimately appear as a sequence of tokens that stretches across the entire AI corpus. Please see Annex II for a more detailed discussion of Data Preparation and the other stages of the AI Development process.

IV. Copyright Issues Raised in the Context of AI Inputs

Any copyright analysis must distinguish between AI inputs and AI outputs. We address each in turn below.

A. AI Inputs: Copyrighted Content May be Incorporated into Large and Diverse AI Data Sets

Computational analysis is typically applied to a large training data corpus that may comprise millions or billions of tokenized data elements – i.e., AI inputs. Depending on how the model is trained, data accessible over the Internet may be collected as part of the raw data set that is subsequently transformed into an AI training corpus of "tokenized" data for purposes of computational analysis.

Raw data from the Internet may be part of copyrighted works. US copyright law sets a low threshold to establish "originality" and provides that copyright arises automatically upon a work's creation, even if it the work is not registered.[9]

However, copyright protection does not extend to facts, ideas, or mathematical concepts, which are the subject of the computational and statistical analytical processes undertaken in AI training.[10] As explained in Annex III, computational analysis is a "Fair Use" under US copyright law.

B. AI Inputs: Computational Analysis Involves the Non-Consumptive Statistical Analysis of Large and Diverse Data Sets

Computational analysis of AI data set inputs does not involve the consumption of any copyrighted works for their expressive content. Rather, such analysis involves mathematical calculations of probabilities, correlations, trends, and other patterns across the entire tokenized data set. Such analysis seeks to understand only the statistical patterns (e.g., the relationships of specific tokens in relation to other tokens) distributed across the entire data set. These statistical patterns are themselves not expressive content protected by copyright law.

Computational analysis may involve two sets of reproductions that potentially implicate the Copyright Act: (1) reproductions necessary to create a corpus of "training data," and (2) transient reproductions that are incidental to the computational process of training the AI model. In each case, the reproductions are "intermediate" in the sense that they are not visible or otherwise made available to the public. Instead, the reproductions are the necessary byproduct of a technical process that is aimed at identifying non-copyrightable information *about* the underlying corpus of works – i.e., the correlations and patterns that inform the creation of the AI model and enable it to make predictions based on future data inputs. Such intermediate, non-expressive reproductions do not compete with and have no impact on the economic interests that copyright is intended to protect. As further discussed in **Annex III**, these reproductions constitute a Fair Use.

C. AI Inputs: Copyright Holders Have Many Enforcement Tools at their Disposal to Prevent their Works from Being Included in AI Training Data Sets

Copyright holders have enforcement tools to prevent material from being incorporated in data sets. Copyright holders should exercise available statutory mechanisms as part of a robust US IP protection and enforcement framework. As discussed below, copyright holders and AI development companies should also collaborate with respect to other mechanisms, such as licenses and terms of service and Do Not Crawl or Do Not Train tags (e.g., robots.txt), to respect the preferences of copyright holders on AI training issues.

D. AI Inputs: Interested Stakeholders Should Continue Discussions on Voluntary "Opt Out" Mechanisms for Copyright Holders

While large and diverse data sets are critical to US AI leadership, we recognize that some web publishers and content owners want more say in how their content is used in or with AI. One way to reflect and respect those wishes is through existing and new voluntary "opt out" tools. For example, publishers can indicate through the use of certain robots.txt signals a preference that the AI crawler or Bot avoid their site, a specific page or other content. AI developers publish the specification web publisher can use to tell the developers crawlers not to crawl the site for purposes of AI training.[11]

At the present time, there are active cross-industry consultations relating to automated tools to indicate a rights-holder does not want a website used for training purposes, similar to the current "do not crawl" tools that apply to search engines. These discussions are occurring in organizations like the IETF (Internet Engineering Task Force), W3C (World Wide Web Consortium), and C2PA (Coalition for Content Provenance and Authenticity). Participants in these and similar discussions include representatives of the creative and technology industries, as well as civil society and academia.

V. Copyright Issues Raised in the Context of AI Outputs

Copyright issues raised in the context of AI outputs are legally and practically different from issues raised in the context of AI inputs, as discussed below.

A. AI Outputs: Generative AI Can Produce Outputs that May Infringe Copyright

Copyright holders should have full and effective remedies when their rights are infringed. This principle applies equally to outputs generated using AI systems and outputs generated in other ways. Whenever such infringement is found, it is critical to fully compensate artists and creators for any damages caused. In most AI use cases, such as those described above, the output of an AI system will not implicate copyright at all. However, AI, like other technologies, could be used to create infringing material. In those cases, existing copyright law should prove adequate to address questions of infringement.

B. AI Outputs: In Generative AI, Sufficiently Large and Diverse Data Sets (on the Input Side) Can Help Minimize the Risk of Infringement (on the Output Side)

While it is important not to conflate training data with the output of an AI system, it is worth noting that the more data available for training, the less risk there is of any possible infringement. In any well-designed AI model trained on a sufficiently large data set, computational analysis should never (or only in the rarest of circumstances) produce outputs that are "substantially similar," let alone "virtually identical," to any specific copyrighted work. As Professor Matthew Sag of Emory University has explained,

> At the moment, memorization [i.e., the accidental replication of a particular work from a training set] is an edge case. For the most part, the link between the training data and the output of generative AI is attenuated by a process of decomposition, abstraction, and remix. Generally, pseudo-expression generated by large language models does not infringe copyright because these models "learn" latent features and associations within the training data, they do not memorize snippets of original expression from individual works.[12]

Former General Counsel of the US Copyright Office Sy Damle has also explained that:

> [A]ttempts to build AI models using smaller sets of licensed or public domain material will lead to models that are less effective and, ironically, more likely to (inadvertently) create outputs that simply regurgitate their training data.[13]

Additionally, as noted above, some AI developers and deployers are already taking steps to engage with artists and creators on how to support their work in a changing digital environment and taking steps to limit misuse of an AI system, such as limiting what prompts can be used.

Annex I – Applications of AI

We offer below a few (widely recognized) examples of insights, predictions, and other outputs derived from computational analysis in the machine learning context:

- Automated flight management and air traffic control based on computational analysis of meteorological conditions, real-time fuel consumption, aircraft operational data, nearby air traffic conditions, airport congestion, and numerous other data elements.[14]

- Identification of chemical and cellular anomalies for early diagnosis, prevention, and treatment in the fields of oncology, autoimmune disorders, and Parkinsons and Alzheimers disease.[15]

- By integrating generative AI into security operations, organizations can effectively identify and address security anomalies, as well as detect and mitigate potential threats.

- Predictive climate modeling based on computational analysis of satellite data, weather station data, topographical information, and various IoT and sensor data.[16]

- Improved carbon tracking and mitigation based on computational analysis of transportation logs, meter readings, fuel purchase records, atmospheric pollution tracking, and visual monitoring of power plants and other facilities, and other data sources.[17]

- Computational analysis to map vulnerable seaside areas to produce cyclone risk maps and guide investment plans for cyclone shelters, schools, health facilities, and other infrastructure for disaster planning and survivability.[18]

- Predictive typing and other office productivity solutions (e.g., an "auto-complete" function that suggests the letters "…cerely yours" after the typist inputs the letters, "sin"),[19] or the creation of sound effects or special effects to assist creators and film producers in developing new artworks.[20]

While the use cases are diverse, the elements of training each are very similar and discussed below.

Annex II – The AI Development Life Cycle

The AI development life cycle typically includes the following steps:

- Project Conception: Often, AI models are designed as general-purpose tools, such as a natural language processing model, that may later be customized for multiple particular tasks. Sometimes, an AI model is developed with a specific task in mind. In that situation, the AI development team must formulate the "problem" that a system is intended to address and map the structure and target variables that the system is intended to predict. For instance, a fitness app that analyzes a consumer's heart rate to monitor for irregularities that might predict whether that person is at risk of a stroke or heart disease (i.e., the target variable).

- Raw Data Identification: Often, AI models are trained on massive amounts of unlabeled data through self-supervised learning to serve as general purpose tools, such as a natural language processing model, that may later be customized for multiple particular tasks. In other circumstances where the model is being developed for a specific task, the AI development team must identify a relevant universe of "raw data" that will subsequently be transformed and structured. Data sources are as diverse as the potential applications of machine learning AI and may include everything from machine-to-machine data (e.g., satellite transmission data) and international trade statistics to published materials, blog posts, website comments, and chat room logs. "Raw data" is frequently "messy," requiring significant work to transform the data into a usable form, as outlined below.

- Preparing the Data Set: The AI development team must modify the "raw data" so that it can be semantically understood by the machine and used to train the model. During this process, the team will revise, clean, and normalize the data as necessary. Data is typically semantically and structurally transformed through "tokenization," which involves breaking down a piece of text or data into smaller units (or "tokens") for purposes of computational analysis. Tokenization is necessary to improve the reliability, quality, and suitability of the data for analysis, helping to address quality challenges, such as missing values, duplicates, outliers, and inconsistent formatting across the entire data set. Training data will often ultimately appear as a sequence of tokens that stretches across the entire AI corpus.

- Model definition: After input data has been suitably processed, the AI development team must establish the system's underlying architecture. This includes identifying the variables (i.e., "features") in the training data that the algorithm will evaluate as it looks for patterns and relationships as the basis of a rule for making future predictions. It also includes selecting the type of algorithmic model that will power the system (e.g., linear regression, logistic regression, deep neural network.)[21] Once the data is ready and the algorithm is selected, the team will train the system to produce a functional model that can make predictions about future data inputs.

- Model Validation, Testing, and Revision. After the model has been trained, it must be validated to determine if it is operating as intended and tested to demonstrate that the system's outputs fall within expected parameters and do not contain unexpected errors or unintended bias. Based on outcome of validation and testing, the model may need to be revised and refined to mitigate these risks.

Annex III – Computational Analysis of AI Training Data Is a Fair Use

The Constitution authorizes Congress to create a framework for protecting intellectual property in order to "promote the progress of science and useful arts."[22] The fair use doctrine is a critical safety valve that ensures the Copyright Act remains consistent with its Constitutional purpose.[23] Codified in Section 107 of the Copyright Act, the fair use doctrine allows for unauthorized uses of copyrighted works in certain circumstances, providing courts with a flexible mechanism for balancing "the interest of authors and inventors in the control and exploitation of their writings and discoveries on the one hand, and society's competing interest in the free flow of ideas, information, and commerce on the other hand."[24]

It is impossible, and unnecessary in light of fair use law's focus on facts of a specific case, to draw a generalized conclusion that *all* applications of AI involving the reproduction of copyrighted works will qualify as fair uses. But the case law suggests strongly that the fair use doctrine will accommodate the creation and use of AI training databases for the purposes of analyzing a large collection of individual works to identify patterns, correlations, and other metadata to develop an AI model that makes predictions about future data inputs.

By way of illustration, this means that an AI developer seeking to create a natural language processing model – such as an AI-driven predictive typing model – can rely on publicly available textual material (ranging from the letters of Flannery O'Connor to anonymous website commentary) to create the training database. In such a scenario, the AI developer would not be reproducing this text for its expressive purpose. Rather, the reproductions would be made solely for the purpose of extracting unprotected information *about* the English language – i.e., the correlations, patterns, and relationships among the 26 letters of the alphabet and the 1 million English language words, as they appear in thousands of stock phrases, figures of speech, similes, metaphors, grammar patterns, and common linguistic formulations and expressions. Neither these letters, words, and phrases, nor the mathematical patterns among them across thousands or millions of writings, are copyright protectable subject matter.

Furthermore, the ultimate use of the computational analysis applied to the data set is a transformative use of the original content. Auto-complete functionality in predictive typing software comprises a new creation in the form of software code that is transformatively distinct both from the entire AI training corpus and from any single work within that corpus. Software code that can suggest the endings of commonly used phrases is many steps removed from any part of a copyrighted work. Finally, such functionality simply does not compete with any copyrighted works in any manner that copyright is intended to protect. Accordingly, the AI developer can safely rely on the fair use doctrine to construct an AI training database using photographs to which she has lawful access.

In addition to the reproductions that may be needed to create an AI training database, reproductions may also be made when the training data undergoes the computational analysis that occurs during the machine learning process. These fleeting reproductions are an inevitable byproduct of working with digital media, which must be loaded into the random access memory (RAM) of a computing device to be accessed, analyzed, manipulated or even deleted.[25] In some circumstances, the creation of RAM copies may constitute a reproduction that is cognizable under the Copyright Act. However, the fleeting nature of the RAM copies involved in the machine learning process are unlikely to meet the Copyright Act's threshold for "fixation" and are therefore unlikely to be considered "copies" that implicate the reproduction right.[26]

The Court's recent decision in *Andy Warhol Foundation v. Goldsmith* does not change the analysis. In *Warhol*, the Court looked at the specific use alleged to infringe and assessed the justification for the use to determine whether it supported a finding of fair use. The Court essentially held that fair use does not protect a licensee who has a license to use a work once, but uses the work additional times and then licenses those subsequent works. As described above, in the case of training an AI system, the use of data is for training purposes and the trained model neither includes a copy of the underlying work nor is it a competing work.

[1] BSA's members include: Adobe, Alteryx, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, Cohere, Dassault Systemes, Databricks, Docusign, Dropbox, Elastic, EY, Graphisoft, HubSpot, IBM, Informatica, Kyndryl, MathWorks, Microsoft, Notion, Okta, OpenAI, Oracle, PagerDuty, Palo Alto Networks, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Trend Micro, TriNet, Workday, Zendesk, and Zoom Communications Inc.

[2] See BSA | The Software Alliance, Artificial Intelligence in Every Sector, available at https://www.bsa.org/files/policy-filings/06132022bsaaieverysector.pdf.

[3] Business Software Alliance, Comments to US Copyright Office regarding Artificial Intelligence and Copyright, at: https://www.bsa.org/policy-filings/us-comments-to-us-copyright-office-regarding-artificial-intelligence-and-copyright; Business Software Alliance, Artificial Intelligence and Copyright Policy, at: https://www.bsa.org/policy-filings/us-artificial-intelligence-and-copyright-policy

[4] https://trumpwhitehouse.archives.gov/ai/

[5] https://www.cnn.com/2025/02/28/tech/trump-ai-apple-investment-stargate-deepseek/index.html

[6] https://www.reuters.com/world/us/trump-deepseeks-ai-should-be-wakeup-call-us-industry-2025-01-27/

[7] The USCO RFC defines training datasets as, "[a] collection of training material (as defined below) that is compiled and curated for use in machine learning. Examples of training datasets include BookCorpus, ImageNet, and LAION." The USCO further defines training material as, "individual units of material that are used for purposes of training an AI model. They may include a combination of text, images, audio, or other categories of expressive material, as well as annotations describing the material. An example of training material would be an individual image and an associated text ''label'' that describes the image."

[8] *See* BSA | The Software Alliance, *Confronting Bias: BSA's Framework to Build Trust in AI* (2022), at: https://ai.bsa.org/confronting-bias-bsas-framework-to-build-trust-in-ai; *See also,* National Institute of Standards and Technology, *NIST Risk Management Framework* (Jan. 2023), at: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (describing an AI system "as an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments.")

[9] *See e.g.,* 17 USC § 102(a) ("Copyright subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.")

[10] *See* Baker v. Selden, 101 U.S. 99 (1879); *see also* "Article 9.2 of the TRIPS Agreement provides that copyright protection shall extend to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such. In other words, copyright protection does not cover any information or ideas contained in a work; it only protects the original way in which such information or ideas have been expressed. Thus everyone is free to use the information contained in a work, including for the purpose of creating new works. For example, the idea behind a detective novel is not protected as such, but an unauthorized reproduction of that novel that is an expression of the idea is prohibited. This principle, commonly referred to as the 'idea/expression dichotomy', has always been present in copyright doctrine, although it had not been explicitly set out in the provisions of the Berne Convention. Article 9.2 of the TRIPS Agreement is therefore the first explicit confirmation of the principle in multilateral IP law." WTO, Overview of the Agreement on Trade-Related Intellectual Property Rights (2023), at: https://www.wto.org/english/tratop_e/trips_e/ta_docs_e/modules2_e.pdf

[11] *See e.g.,* https://openai.com/index/approach-to-data-and-ai/

[12] *Matthew Sag, Copyright Safety for Generative AI*, 61 Hous. L. Rev. 101 (Aug. 10, 2023), at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4438593

[13] *Statement of Sy Damle re "Artificial Intelligence and Intellectual Property Law - Interoperability of AI and Copyright Law,"* Testimony before the US House of Representatives Committee on the Judiciary, Subcommittee on the Courts, Intellectual Property, and the Internet (May 17, 2023). "Critically, the ability of AI models to duplicate training data is a bug, not a feature." *Id.*

[14] *See e.g.,* M. Durgut, *Artificial Intelligence and Air Traffic Control*, Aviationfile.com website (Jan. 2023), at: https://www.aviationfile.com/artificial-intelligence-and-air-traffic-control/#:~:text=One%20of%20the%20primary%20applications%20is%20to%20help,make%20informed%20decisions%20on%20routing%20and%20scheduling%20flights; Degas et al., *A Survey on Artificial Intelligence and explainable AI in Air Traffic Management*, 12 Applied Sciences 1295 (2022), at: https://www.mdpi.com/2076-3417/12/3/1295; Hanneke Weitering, *How Artificial Intelligence is Transforming Aviation*, Futureflight.aero website (2023), https://www.futureflight.aero/news-article/2023-07-13/beyond-automation-how-artificial-intelligence-transforming-aviation

[15] *See e.g.,* Hunter et al., *The Role of Artificial Intelligence in Early Cancer Diagnosis*, 14(6) Cancers 1524 (2022), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8946688/; Stafford et al., *A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases*, 3 NPJ - Digital Medicine 30 (2020), at:

https://www.nature.com/articles/s41746-020-0229-3; Diogo et al., *Early diagnosis of Alzheimer's disease using machine learning*, 14 Alzheimers Research and Theory 107 (2022), at: https://alzres.biomedcentral.com/articles/10.1186/s13195-022-01047-y; Ahsan et al., *Machine-Learning-Based Disease Diagnosis: A Comprehensive Review*, 10(3) Healthcare Basel 541 (2022), at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8950225/; Luchini et at., *Artificial Intelligence in Oncology*, 126 British J. of Cancer 1 (2022), at: https://www.nature.com/articles/s41416-021-01633-1; Yumar et al., *Artificial intelligence in Disease Diagnosis*, 14(7) J. Ambient Intell Humaniz Comput. 8459 (2023), at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8754556/

[16] Schneider et al., *Harnessing AI and computing to advance climate modelling and prediction*, 13 Nature Climate Change 887 (2023), at: https://www.nature.com/articles/s41558-023-01769-3; World Economic Forum, *The role of machine learning in helping to save the planet* (2021), at: https://www.weforum.org/agenda/2021/08/how-is-machine-learning-helping-us-to-create-more-sophisticated-climate-change-models/; Kaak et al., *Aligning artificial intelligence with climate change mitigation*, 12 Nature Climate Change 518 (2022), at https://www.nature.com/articles/s41558-022-01377-7; Xin et al., *Artificial Intelligence for Climate Change Risk Prediction, Adaptation, & Mitigation*, Ecological Processes (2021), at: https://www.springeropen.com/collections/AICC; Chantry et al., *Opportunities and challenges for machine learning in weather and climate modelling*, 379 Phil. Trans. R. Soc. 83 (2020), at: https://doi.org/10.1098/rsta.2020.0083 (2020).

[17] *See e.g.,* Global Data Alliance, *Cross-Border Data Transfers & Environmental Sustainability* (2023) (internal citations omitted), at: https://globaldataalliance.org/wp-content/uploads/2023/04/04192023gdacbdtsustainability.pdf

[18] *See id.*

[19] *See e.g.,* S. Ashraf, *Desmystifying Autocomplete,* Towards Data Science (2020), at: https://towardsdatascience.com/index-48563e4c1572; A. Wickramarachchi, *Machine Learning: Word Embedding and Predicting,* Towards Data Science (2020), at: https://towardsdatascience.com/machine-learning-word-embedding-and-predicting-e603254e4d7b

[20] *See e.g.,* D. Nelson, *AI Researchers Design Program To Generate Sound Effects For Movies and Other Media*, Unite AI Website (2022), https://www.unite.ai/ai-researchers-design-program-to-generate-sound-effects-for-movies-and-other-media/

[21] *See* BSA AI Bias Framework, *infra* note 7.

[22] U.S. Const. Art I, § 8, Cl. 8.

[23] *See* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 575 (1994).

[24] Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 539, 580 (1985) (quoting Sony Corp. of Am. V. Universal City Studios, Inc., 464 U.S. 417, 429 (1984)).

[25] *See* Aaron Perzanowski, *Fixing RAM Copies*, 104 NW. U. L. REV. 1067 (2010); Jule L. Sigal, *Copyright Infringement Was Never This Easy: RAM Copes and Their Impact on the Scope of Copyright Protection for Computer Programs*, 45 CATH. U. L. REV. 181 (1995).

[26] The reproduction right affords copyright owners the exclusive right to make "copies" of their works. 17 USC § 106. For a copy to trigger the reproduction right, it must be "fixed" in a manner that allows the work to be "perceived, reproduced, or otherwise communicated." 17 USC § 101. A copy is considered "fixed" when its "embodiment" is "sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for more than a period of transitory duration." *Id.* A copy therefore implicates the reproduction right only when it is (1) embodied in a medium and (2) the embodiment subsists for more than a "transitory" duration.