

**AAP Comments on Draft 2019-2020
Federal Data Strategy Action Plan**

The Association of American Publishers (AAP) represents the leading book, journal, and education publishers in the United States on matters of law and policy, advocating for outcomes that incentivize the publication of creative expression, professional content, and learning solutions. As essential participants in local markets and the global economy, our members invest in and inspire the exchange of ideas, transforming the world we live in one word at a time. The U.S. publishing industry supports an extensive network of American businesses and hundreds of thousands of American jobs. Publishers are also an integral part of America’s copyright industries, which collectively add \$1.3 trillion in annual value to U.S. gross domestic product.

AAP appreciates the opportunity to provide comments on the Draft 2019-2020 Federal Data Strategy Action Plan (the “Draft Action Plan”). In addition to investing in the creation of copyrighted works of original expression, AAP’s members also produce other forms of high-quality content, including databases. AAP’s members invest in various technologies, distribution platforms, data analytics methodologies, and other innovative tools that enable teachers, doctors, scientists, researchers, and the public to get the greatest possible benefit from our members’ diverse content. We therefore strongly support the Administration’s Federal Data Strategy Mission to “leverage the full value of Federal data for mission, service, and the public good by guiding the Federal Government in practicing ethical governance, conscious design, and a learning culture.”

AAP members recognize that copyrighted works of original expression may serve as a body of “data” for purposes of “text and data mining” or research in which computational analysis is performed on one or more copyrighted works *in a manner that does not involve reading or displaying substantial portions of such works to understand their expressive content*. However, the need to respect the intellectual property rights embodied in copyrighted works requires care to ensure that, as a general proposition, such works are not inadvertently swept up in plans or strategies to maximize public exploitation of “data” collected, obtained, or maintained by federal agencies simply because the term “data” is broadly characterized to mean “recorded information, regardless of form or the media on which the data is recorded.” Furthermore, datasets may be subject to a wide range of contractual or other proprietary business interests. Accordingly, we submit these comments to highlight one key point:

In implementing the Federal Data Strategy Action Plan, it is essential that federal agencies recognize and safeguard the intellectual property rights and proprietary interest of publishers (and other private parties) whose copyrighted content or proprietary assets find their way into the Federal Government’s hands.

The importance of this point is echoed in the Administration’s Federal Data Strategy Principles and Practices, but unfortunately does not appear anywhere in the Draft Action Plan. The Principles lay the foundation by explaining that agencies should “Exercise Responsibility” (Principle 2) in their stewardship of data by, among other things, ensuring “appropriate access and use” with respect to data. The Practices then note that agencies should “Explicitly Communicate Allowable Use” (Practice 31), including by employing descriptive metadata that “recognizes and safeguards applicable intellectual property rights.”

The Foundations for Evidence-Based Policymaking Act of 2018 (the “Evidence Act”) provides further support for the importance of recognizing and safeguarding intellectual property rights and proprietary interests that may be affected by Federal Government data policy. Title II of the Evidence Act (the Open Government Data Act) establishes a “default of openness” for data management by federal agencies. While the “default” includes requirements for “open Government data assets” to be publicly available in open formats and subject to open licenses regarding their use, the legislation—as explained in the accompanying report of the House Committee on Oversight and Government Reform—addresses “instances where it could be inappropriate for the government to impose open license requirements” (H.Rpt.115-411, p.12). For example, the report makes clear that it would be inappropriate to impose open license requirements on “data that the government uses and maintains but does not own.” In addition to provisions that address such instances involving contractual, privacy, security, business and other confidentiality considerations, the bill contains provisions to ensure that agencies evaluate whether a data asset is “subject to intellectual property rights, including rights under titles 17 and 35, United States Code” to determine whether it can be made publicly available subject to an open license.

A “Rule of Construction” in Section 401 of the Evidence Act makes clear that “[n]othing in this Act, or the amendments made by this Act, may be construed to override, limit, or otherwise affect intellectual property rights, including rights under titles 17 and 35, United States Code.” The bill’s definition of “open Government data asset” further clarifies that such assets will remain subject to restrictions that arise from preexisting “intellectual property rights.” A similar qualification regarding data assets “subject to intellectual property rights” appears in the Act’s provisions regarding OMB Guidance for agency heads “on mak[ing] data open by default” and creating a “comprehensive data inventory” that “provides a clear and comprehensive understanding of the data assets in the possession of the agency.” And with respect to proprietary business assets, the Act’s “Rule of Construction” in Section 401 further makes clear that nothing in the Act may be construed to “affect the authority of a Federal agency regarding the use, disclosure, or licensing of...confidential business information...or data assets restricted from disclosure under a contract or other binding, written agreement.”

The Draft Action Plan should clarify that agencies implementing the proposed actions should do so in a manner that recognizes and safeguards intellectual property rights and proprietary interest in accordance with Administration policy and applicable law, including but not limited to the Evidence Act and the Copyright Act. In order to do so effectively, it will be crucial for agencies to do three things:

1. Agencies must develop a clear understanding of the difference between non-copyrightable data that is unencumbered by intellectual property rights and copyrighted expression that constitutes privately-owned intellectual property. Agencies should take special care to avoid defining (or interpreting legislative definitions of) “data” so broadly as to encompass copyrighted works such as books, journal articles, or other creative expression.
2. Agencies must take care to distinguish between Federal Government data—i.e. data that is property of the Federal Government—and privately-owned proprietary data or other materials. In particular, it is critical that agencies understand that Federal Government *possession, maintenance, or dissemination* of certain data or materials does not mean

that such data or materials are not privately owned or not still subject to restrictions on use based on such ownership. With respect to expressive written materials, unless they were created by or exclusively for the Federal Government, agencies should take special care to determine whether a private party owns copyright in such materials.

3. Before providing public access to data or other materials, agencies should recognize and account for any private ownership interests. Agencies should not provide access to data or materials in a manner that would interfere with a private party's intellectual property rights or proprietary interests. To give just one example, under no circumstances should an agency affix (or be required to affix) an open license, without permission from the copyright owner, to copyrighted journal articles that the agency possesses, maintains, or disseminates.

By doing the three things above, agencies can take an important step to ensure that the Administration's Federal Data Strategy supports creativity, innovation, and economic growth, and does not unintentionally disrupt the foundational property rights that form the basis of our modern creative economy.

AAP is grateful for the Administration's focus on "demonstrating responsiveness" to its data policies through "ongoing input from users and stakeholders" (Federal Data Strategy Principle 7). We look forward to working together with the Administration to ensure that the Federal Data Strategy, as implemented through its yearly action plans, supports continuing private-sector investment in the creation and dissemination of published works.

June 28, 2019

Respectfully submitted,



Matthew Barblan
Vice President, Public Policy
Association of American Publishers