Los Alamos NATIONAL LABORATORY

Facilitating Use of Unencumbered Publications as Data for AI

Nicholas Taylor

Research Library

CNI Spring Membership Meeting

13 April 2026

LA-UR-26-22604

AI-Accelerated Science

LANL scientists are looking to develop **scientific AI foundation models**

Aside from AI-ready research data, a logical way to **imbue models** with **scientific domain knowledge** is to train them on **relevant publication corpora**

Then the models, plus physical instruments, can be **orchestrated via agents** for **semi-autonomous science and R&D**

"GM-Social Post_DOE insta.jpg" by U.S. Department of Energy under Public Domain

LOS ALAMOS NATIONAL LABORATORY Research Library

Content Licenses + Copyright

Our content licenses largely **don't allow for use of publications with AI** nor do they permit systematic downloading.

Licensing pathways are **emerging, slowly**.

Court rulings on use of copyrighted works for AI training under Fair Use are **few and not necessarily precedential**.

Our contracts mostly **override Fair Use**, anyway.

LOS ALAMOS NATIONAL LABORATORY Research Library

"Antelope Valley California Poppy Reserve 2019" by waltarrrrr under CC BY-NC-ND 2.0

Unencumbered Works

While pursuing licensing pathways, we're also helping our researchers find works that are **unencumbered for AI use cases**.

That is, works **without significant licensing or copyright restrictions** that would prevent their use with AI.

These can be either **unencumbered versions** of specific publications or **alternate, unencumbered publications**.

"Hot air balloons in the sky" by Manny Becerra under Unsplash License

Tools and Sources

To find unencumbered publications, we canvass **open content indexes**, **open repositories**, and **open access journals**.

Key tools and sources include arXiv, CORE, OSTI.GOV, and Unpaywall.

API access and accurate metadata, particularly as related to licenses, are useful features supporting these searches.

"Details of old keys in a box" by Ivan Radic under CC BY 2.0

LOS ALAMOS NATIONAL LABORATORY Research Library

Frameworks for Legal Use

We’re looking for publications under an **applicable use framework**.

Some **Creative Commons licenses** (i.e., CCØ, CC BY) allow for AI use cases.

The **Government Use Right** permits Government use of Government-funded or -authored works.

Other open access licenses may or may not allow for AI use cases.

“Supreme Court detail” by Dave Newman under CC BY 2.0

LOS ALAMOS NATIONAL LABORATORY
Research Library