STM TDM Reservation Protocol (TDMRep) – Publisher Implementation Testimonial

Matteo Ripamonti - Head of Third-Party Relationships - Elsevier

As one of the first publishers to implement the TDM Reservation Protocol (TDMRep), Elsevier had the opportunity to learn while charting its implementation approach. This case study describes the experience and lessons learned during implementation, which may be helpful to others as they evaluate how to implement TDMRep.

Quick recap:
- Text and data mining (or TDM) is the automated process of analyzing large collections of text or data to uncover patterns, insights, and relationships.
- The TDM Reservation Protocol (TDMRep) is a technical standard that enables publishers to clearly signal, in a machine-readable way, whether and how their content can be mined. This supports legal compliance, transparency, and operational efficiency for both publishers and researchers, and is especially relevant under Articles 3 and 4 of the 2019 EU DSM Directive, with the latter requiring publishers to signal when rights are reserved for text and data mining.
- Implementing the TDMrep protocol is one of the recommendations from the STM Article 4 Task & Finish Group.

Why We Did It – and Why It Matters for Publishers

Elsevier chose to be an early adopter of TDMRep for several strategic reasons, which go beyond compliance, also encompassing efficiency, transparency, and future-preparedness. Other publishers might also view these as potential benefits to adopting TDMRep:

- **Legal Compliance and Risk Reduction:**
  TDMRep provides a standardized, machine-readable way to meet the EU DSM Directive requirement to signal when rights are reserved for TDM, reducing ambiguity and legal exposure.
- **Transparency and Control:**
  TDMRep enables publishers to communicate their TDM policies clearly, consistently, and without ambiguity. Because all machine-readable signals point users and automated tools to the same, publisher-managed information page, publishers can ensure that every user receives accurate, up-to-date TDM policy details. This improves transparency while giving publishers full control over the policy page to which all signals direct, ensuring alignment, clarity, and easy updates over time.
- **Operational Efficiency:**
  Automated signalling minimizes manual inquiries and licensing overhead, saving time for both publishers, researchers and other TDM actors.
- **Improved Relationships with Researchers and Platforms:**
  Clear, consistent policies reduce friction and uncertainty, fostering trust and collaboration with the research community and technology partners.

Advancing human progress together

- **Future-Proofing:**
  Aligning with World Wide Web Consortium (W3C) standards positions publishers for compliance with emerging AI governance frameworks and automated rights signalling requirements.

Collaboration and Standards

Elsevier's implementation of TDMRep has been fully aligned and in participation with all other members of STM community in a [dedicated working group](#) that helped shape its current status. The industry collaboration closely followed the recommendations and technical standards developed in collaboration with the W3C TDMRep Community Group. The W3C specification (https://www.w3.org/community/reports/tdmrep/CG-FINAL-tdmrep-20240510/) defines the protocol for signalling TDM reservation and providing policy information in a machine-readable format. Elsevier actively participated in this community, ensuring the approach aligns with the latest standards and best practices.

Getting Started

The TDMRep protocol establishes standardized signals for communicating TDM reservation status and policy details. As a publisher, the first step was to analyze content types and determine the most effective way to communicate Elsevier's TDM policies. This required close coordination with internal stakeholders and technical teams to ensure clarity and consistency.

Technical Implementation

From a technology standpoint, Elsevier's deployment of TDMRep adopts several implementation options, allowing flexibility and scalability for different content types and platforms:

- **Website-Level Signals:**
- The starting point is one single location for all the TDM policy information to point to:
  - **Well-Known Directory:**
    - A standardized /.well-known/tdmrep.json file placed at the root of the domain serves as a single, authoritative source for TDM policy information. This approach enables automated tools to quickly discover and retrieve consistent policy data without the need to crawl multiple pages, improving efficiency and reducing complexity.
- The simplest approach to refer to this location is to add machine-readable signals directly to Elsevier's websites. This includes:
  - **HTTP Headers:**
    - Each web page can send TDMRep reservation status and policy links via HTTP headers, enabling automated detection by mining tools and platforms.
  - **HTML Meta Tags:**
    - For more sophisticated signals, meta tags embedded in the HTML provide additional context and redundancy for web crawlers and aggregators.
- **File-Level Metadata:**

Advancing human progress together

- For distributed content such as PDFs, EPUBs, and XML files, TDMRep metadata is embedded directly into the file (e.g., in the XMP area for PDFs). This ensures that TDM policy signals travel with the content, regardless of distribution channel.
- **Crossref Metadata Integration:**
  - As an additional option, Elsevier deposits TDMRep indicators into Crossref metadata records. This ensures that platforms and services relying on Crossref can access up-to-date TDM policy information for all Elsevier articles.

Elsevier is actively deploying these options across both new and legacy content, ensuring comprehensive coverage and future-proofing its TDM signalling.

The official Elsevier TDMRep license and policy details are published at: https://www.elsevier.com/legal/tdmrep-license

Our implementation closely follows the W3C TDMRep specification and benefits from ongoing collaboration with the W3C community: https://www.w3.org/community/reports/tdmrep/CG-FINAL-tdmrep-20240510/

Early Signals from the Community:

Since deploying TDMRep, Elsevier has begun receiving an uptick in interactions from various TDM actors who are actively detecting and interpreting the new signals. These inquiries confirm that TDMRep is being read and used as intended.

Summary

Here are some tips about implementing TDMRep that might be useful for other publishers:

- **Metadata early!** If you aren't already adding TDMRep signals to your web domains and article metadata, do that as soon as possible. This is the "longest pole" in the TDMRep implementation tent. The sooner this part is in place, the sooner you will have a critical mass of "TDMRep-compatible" content online.
- **Who's who.** Allow a generous amount of time to get internal agreement on TDM policies. Identify your approvers and stakeholders early.
- **Don't overthink implementation.** TDM policies might be viewed as complicated and difficult to work through, but the actual TDMRep implementation is straightforward and possible to achieve with free or low-cost tools and services.
- **Test.** Of course!
- **Contact:** If you have further questions, feel free to reach out to Matteo Ripamonti (m.ripamonti@elsevier.com)

References:

- https://www.elsevier.com/legal/tdmrep-license
- https://www.w3.org/community/reports/tdmrep/CG-FINAL-tdmrep-20240510/
- https://stm-assoc.org/what-we-do/strategic-areas/standards-technology/tdmrep/

Advancing human progress together