

Resources for Understanding the Data Sharing Landscape: Rights, Licensing, and Related Initiatives



Sam Grabus, *Doctoral Research Assistant, Drexel*
& Jane Greenberg



Statement of Need:

Existing literature points to inconsistent guidance from Institutional Review Boards (IRBs) regarding how to handle and safely disseminate specific types of sensitive data (Goldenberg et al., 2015). Considering that many agencies and organizations do not share their data because of the risks associated with the mishandling of sensitive and private information, **an understanding of the current rights management and licensing landscape is essential in order to appropriately guide researchers toward data sharing tools and resources.**

Goal:

Compile a first-phase essential resource list of current licensing and rights management efforts that seek to facilitate data sharing including

- initiatives related to standardizing licensing and rights management;
- technological infrastructure;
- ontologies and metadata standards that could be implemented to communicate researcher data sharing needs;
- community-driven efforts;
- and other curated resources for facilitating conversation and progress for mitigating the data sharing challenges across all environments

Methods: Survey of landscape

Outputs:

- Diagram that categorizes initiatives
- Timeline that traces when these initiatives were started
- Spectrum visualization capturing initiatives that focus on data that can be shared in a mostly “open” environment, vs. those that are attempting to facilitate sharing of data types that may be restricted by privacy and proprietary concerns.

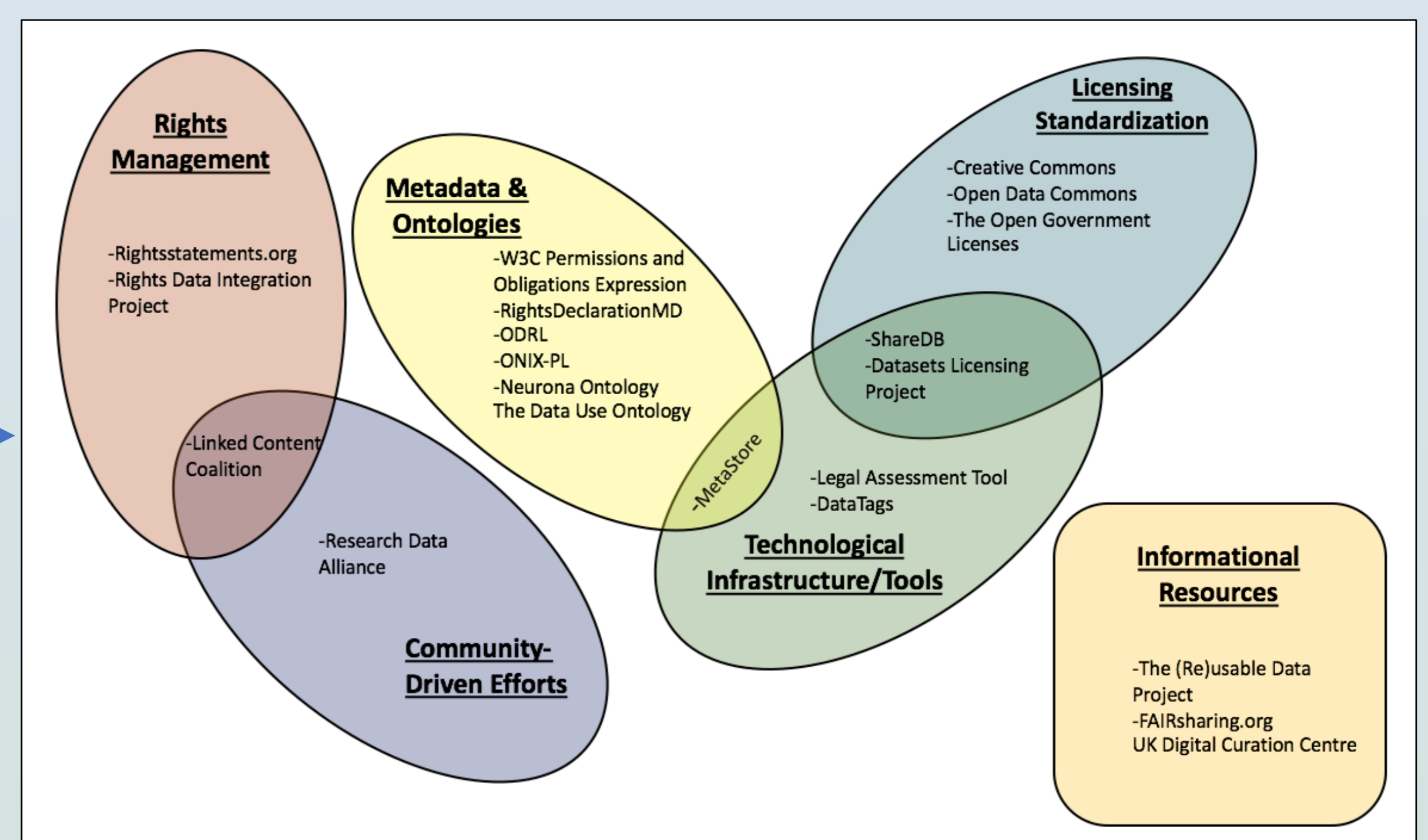
1) Identify Categories of Initiatives

Licensing Standardization
Rights Management
Technological Infrastructure/Tools
Community-Driven Efforts
Metadata & Ontologies
Informational Resources

RE: Metadata, none of the rights or licensing-related standards and schemas were developed specifically for use with research data, and the two ontologies are domain/community-specific.

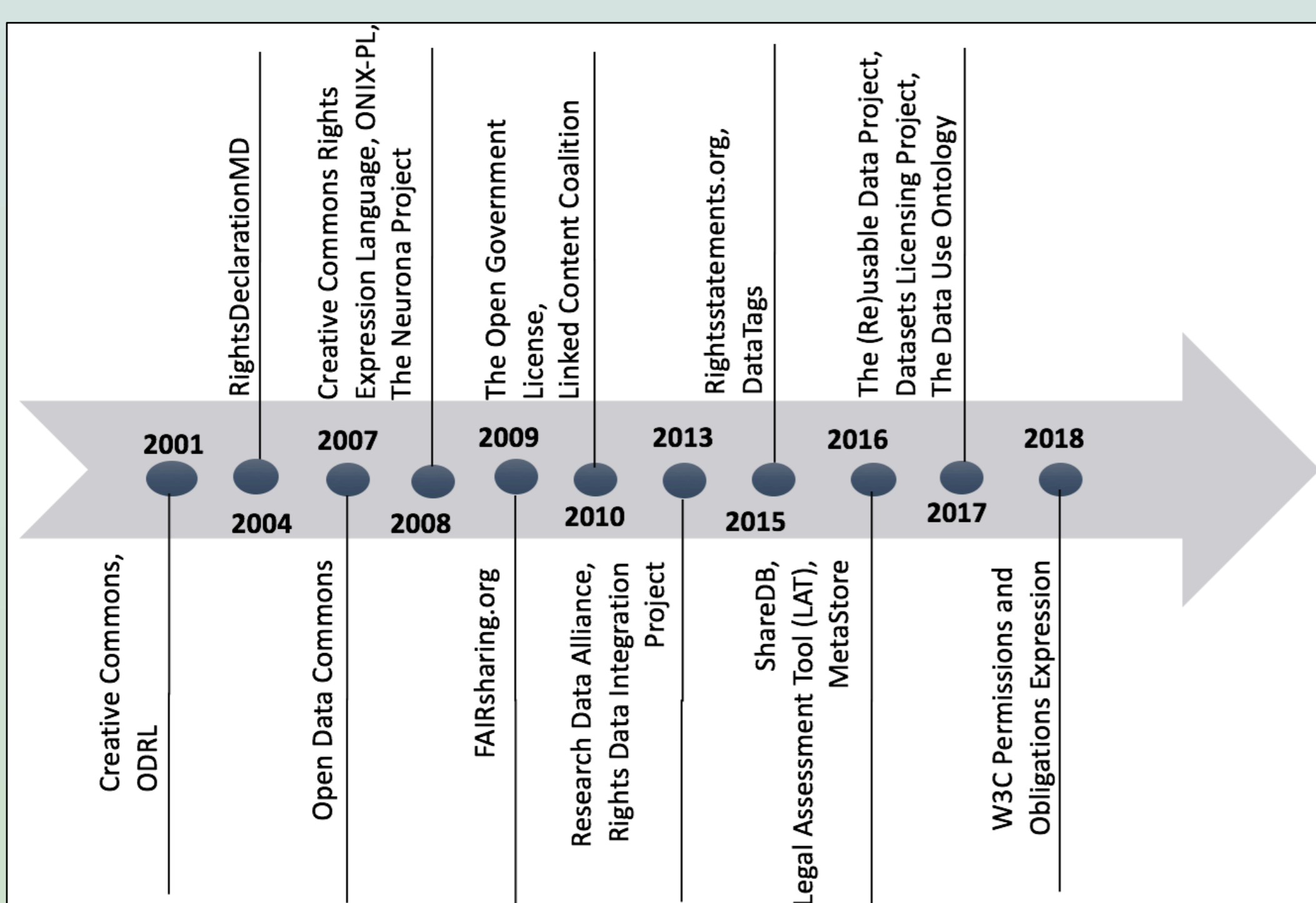
Identifies need for universal/cross-disciplinary data sharing ontology

6 Overlapping Categories



2) Timeline: when these initiatives were started

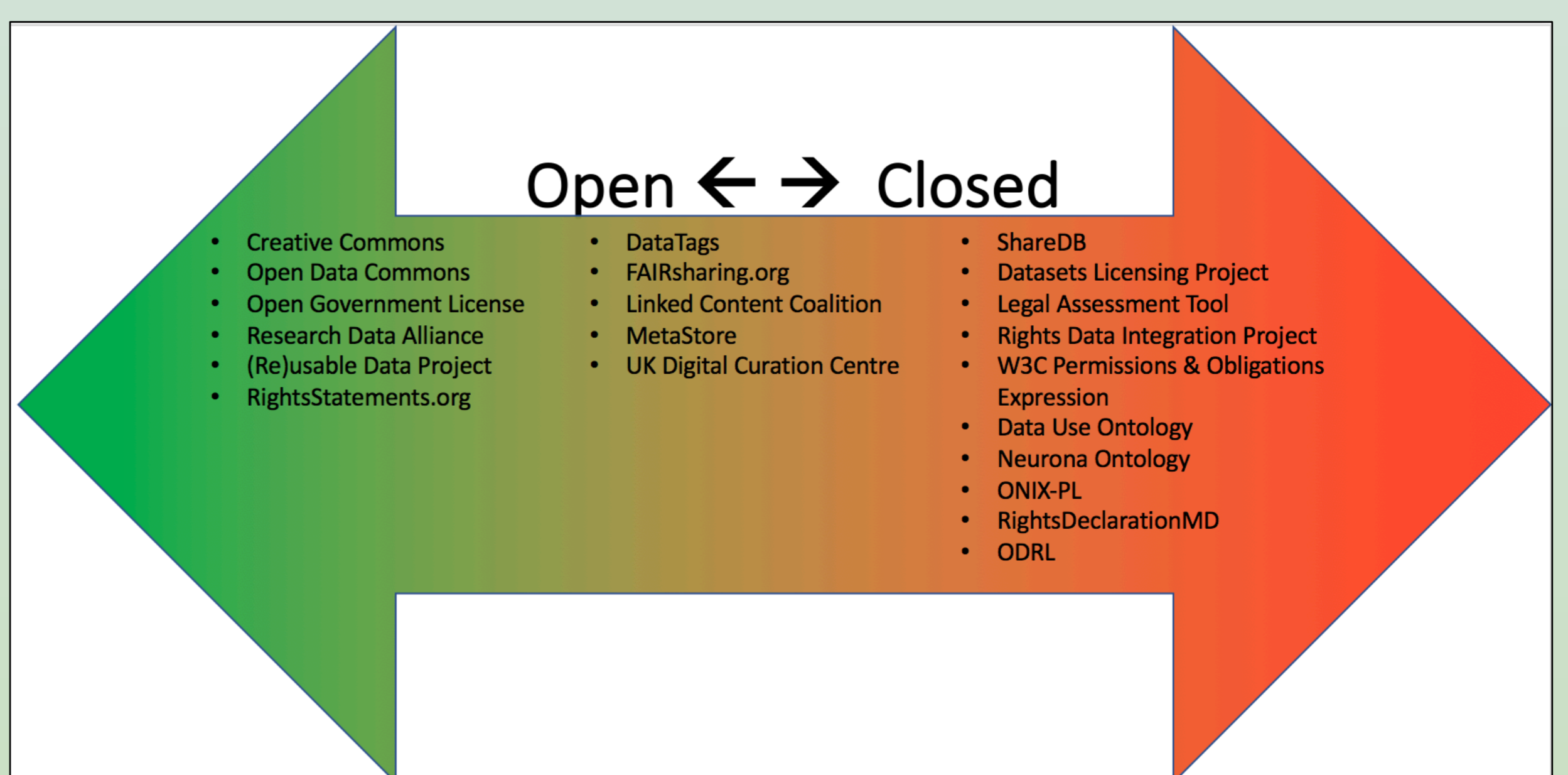
Shift in licensing initiative priorities from “open” to more nuanced and technologically robust, to ensure that sensitive data types can also be responsibly shared.



3) Spectrum: “open” vs. “closed” initiatives

“Open”: Unfettered access and re-use with one or two specifications on how it can be handled in terms of attribution, remixing, and no commercial use.

“Closed”: Looking at the **more complex legal aspects**, in terms of protecting personality identifiable information and proprietary information, and how to express permissions and obligations in a more specific way.



Initiatives in the middle of the spectrum either seek to mitigate both open and closed data types or do not fall neatly into either category

Resources:

Sponsored in part through grants from the Alfred P. Sloan Foundation #G-2014-13746 and from the National Science Foundation NSF ACI #1349002. This research is also supported and funded, in part, by National Science Foundation Award #1636788.

Acknowledgements to the RDA/CODATA Legal Interoperability Interest Group, the RDA/NISO Privacy Implications of Research Data Sets Interest Group, Paul Ulhir, Bob Chen, Simon Hodson, Todd Carpenter, and others, for their support.

1. Goldenberg, A. J., Maschke, K. J., Joffe, S., Botkin, J. R., Rothwell, E., Murray, T. H., . . . Rivera, S. M. (2015). IRB practices and policies regarding the secondary research use of biospecimens. *BMC Medical Ethics*, 16(1), 32. doi:10.1186/s12910-015-0020-1