SCIENCE EDUCATION

Crowdsourcing and Curating Online Education Resources

The diverse wealth of crowdsourced online tools can benefit from professional coordination and quality control.

Darrell Porcello* and Sherry Hsi

The Internet is a growing source of open educational resources (OERs) focused on Science, Technology, Engineering, and Math (STEM). These STEM OERs are not only shared openly and free for all to use, but often provide licenses that permit modification and reuse. Educators must have access to tools that pinpoint valuable resources while avoiding substandard ones. We discuss how multiple information sources, user communities, and online platforms might be coordinated to craft effective experiences in digital-rich learning environments.

There is no shortage of STEM OER collections. Public education and government agencies, television stations, state-level after-school organizations, science museums, and other organizations produce high-quality educational content for the Web and form extensive collections of collaborative, project-based, and open-ended STEM educational activities. Individual educators, scientists, and hobbyists use online platforms like Instructables, MakeProjects, and Pinterest to self-publish instructions for creative projects.

But as the number of STEM OERs grows online, how do educators decide what collections to use when searching for digital content? In a large kindergarten-to–high school (K–12) annual U.S. national survey, 41% of principals responded that it was "difficult to evaluate quality of digital content" (*1*), but >50% of teachers responded that the most important factors in evaluating content were "being referred by a colleague," "free," and "created by educators" (*2*), none of which is necessarily a hallmark of quality.

This highlights a difference within the online world of OERs. On one hand, collections using crowdsourcing allow a wide range of online users to contribute, choosing their own descriptions and keywords to catalog, review, and manage OERs. This can produce large and loose collections. Although dedicated users from the resulting online communities might be more willing to detail their instructional experiences and personal connections to OERs, overall quality may be quite variable. On the other hand, having professional staff with disciplinary expertise populate, catalog, and maintain curated collections usually results in fine-tuned, coherent, and smaller collections. Although these collections grow more slowly, often with less vocal user communities, they can quickly respond to new quality or educational standards.

Over the past decade, online platforms for STEM OERs have vacillated between these two poles. New efforts are combining the best aspects of both strategies to create sustainable online platforms to make it easier for educators to discover and use high-quality STEM OERs, confident in their scientific accuracy and pedagogical approach.

Four Components Essential to OER Success

Convergence Toward Common Metadata. A standard set of terms, or metadata, used by a dedicated community of users to tag digital resources is a necessary characteristic of STEM OER collections. Grade level, subject area, cost per group, and resource type are just a few of the metadata fields relevant to educators. For example, metadata could allow an educator using Howtosmile.org, a free digital library of almost 3500 STEM OERs for out-of-school educators, to identify 82 hands-on activities on evolution for 8- to 11-year-olds.

The National Science Digital Library (NSDL), initially funded by the U.S. National Science Foundation, set out to establish a common set of metadata fields and controlled vocabulary—a well-defined list of words to choose from when populating a metadata field, e.g., "biology" or "physics" for subject area—for online STEM educational resources by knitting together disciplinary and audience-specific communities (*3*). This provided the foundation for a longstanding consortium of online platforms for STEM OERs [e.g., (*4*)]. The core infrastructure team of the NSDL recently updated to the Learning Application Readiness (LAR) metadata format in response to the increasing demand for STEM OERs to be "aligned to educational goals, curriculum, or professional development needs of users" (*5*).

Synergistic efforts of the Learning Resource Metadata Initiative (LRMI), funded by the Bill and Melinda Gates Foundation, establish a simplified set of standards to tag all educational online content, making it easier to publish, discover, and deliver quality OERs on the Web (*6*). Careful cataloging through metadata fuels advanced search engines and more effective discovery of STEM OERs by educators. Metadata can also provide information about how a resource aligns to curricular standards for STEM OERs.

Balancing Expert and Community Definitions of Quality. Learning resource and metadata quality are required for STEM OER online platforms to be accepted by parents, educators, and administrators, but who decides what is quality information? In the Wikipedia model, the written articles are largely crowdsourced, although core community members monitor and control the quality of information. As one of the most-visited sites on the Internet, with more than 23 million articles, the open digital encyclopedia has almost 80,000 volunteer editors who regularly contribute and edit content (*7*). Online STEM OER communities do not have this critical mass to use crowdsourcing as the sole source of quality control. It can play a part, but a professional or trained volunteer

Lawrence Hall of Science, University of California, Berkeley, Berkeley, CA 94720USA.

*Author for correspondence: porcello@berkeley.edu

CREDIT: HAYES THRONTON/LAWRENCE HALL OF SCIENCE staff is required to curate a high-quality collection of STEM OERs.

In addition, as standards alignment and student achievement metrics become more critical parameters for teachers and administrators choosing STEM OERs, professional services are required for increased reliability and accuracy in metadata creation. These points are demonstrated in NASAWavelength.org, a new STEM OER collection from NASA where resources are reviewed by educators using a defined set of criteria (*8*), aligned to the AAAS Project 2061 Benchmarks for Science Literacy (*9*) when appropriate, and offered through a powerful search engine and a visual browse organized by relevant educational concepts. The peer educator and subject-area expert review process for the CPALMS online platform (*10*), aimed at Florida K–12 teachers, can only accept STEM OERs based on rigorous and standards-driven instructional materials.

As digital libraries evolve to identify effective teaching strategies and just-in-time information for educators, a new specialized profession is emerging for a dedicated cyberlibrarian. This person would review incoming resources against a set of quality criteria to ensure only the best materials with strong STEM content, pedagogy, and standards alignments are included in STEM OER collections. The "cybrarian" also can serve as the key decision-maker in weighing administrator goals, adding regional-specific content, and determining use with traditionally underrepresented populations in STEM or for students with disabilities.

Community Input. User voices can shape the public face and enhance the usability of online platforms for STEM OERs. Although users can author and submit new content to grow a collection, the most practical community input lies within "digital footprints," e.g., page views, comments, ratings, or newly formed subcollections, left by users' online interactions and behaviors across an online platform. Users often leave footprints during activities they do regardless of their benefit to the community. For example, users on Howtosmile.org are >10 times more likely to create and save a subcollection (*11*) of STEM hands-on activities than leave a comment or rating for an individual resource. A user's subcollection might inform other users exploring a set of STEM OERs. Web sites like Gooru (*12*) are combining STEM OERs into lesson plans to attract educators.

Another strategy for community input is to have users participate in cataloging duties or earn rewards for ranking content. InBloom (*13*), formally the Shared Learning Collaborative, is introducing an easy to use LRMI-based tagging tool for educators to catalog online resources, whereas a forthcoming student-centered platform called Sparticl from Twin Cities Public Television (*14*) uses "folksonomies"—tags offered by community folks—and user ranking to better deliver STEM OERs to teens. No matter how it is eventually used, genuine community input for an online platform can only begin after there is first a critical mass of cataloged, high-quality STEM OERs established.

Interoperability. When an educator is searching for a resource, like an interactive animation illustrating the steps of mitosis, it would be more efficient if a single query could search across multiple online collections, rather than having the educator visit multiple Web sites. Interoperability enables information flow across multiple online platforms to ensure broad-based access to STEM OERs. Imagine searching for STEM OERs similar to using a single site to search for tickets from multiple airlines.

One necessary step toward interoperability is the adoption of an openly licensed, standardized metadata format for STEM OERs. Pushing beyond a basic set of library record fields from the long-standing Dublin Core metadata standard, this new, expanded metadata standard, with an updated controlled vocabulary, can address specific characteristics of STEM OERs and their users. Additions might include reading level, learning style preferences, assessment data, material costs, standards alignment, or user accessibility. The effort toward common metadata has led to several aggregator Web sites that search over large sets of STEM OERs including NSDL.org, Informalscience.org from the Center for Advancement of Informal Science Education (CAISE) (*15*), the cross-U.S. federal agency Learning Registry (*16*), and the EU Open Science Resources Project (*17*).

Although this work represents progress, an agreement on a STEM OER metadata standard has not been reached within the education community. But interoperability does not stop at metadata. NSDL has begun to promote a standard set of community input called "paradata" (*18*) as another way to share contextualized usage data about a resource and other relevant information across multiple online platforms. The hope is that STEM OERs listed on multiple platforms will retain usage information to establish a common set of digital footprints to be viewable at all points of access. Interoperability must stretch to the technical back-end systems of online platforms to include standard metadata harvesting, resource vitality, and analytics protocols. Following good Linked Open Data (*19*) practices will better interconnect STEM OERs and enable greater flexibility for application development based on data from online platforms.

STEM OERs are becoming important parts of teaching experiences inside and outside the classroom. As search and discovery for STEM OERs become more seamless and natural to the online workflow of the educator and student, the four critical components described above will drive advanced architecture to streamline delivery within browsers, mobile devices, smart boards, and future educational technologies. A distributed, yet fully interconnected, online landscape of STEM OERs will improve quality and the user experience and will help avoid inadvertently recreating resources that already exist during times of dwindling budgets for new materials. When creating or disseminating new STEM OERs, a sensible strategy would be to leverage existing projects to deliver OERs in a responsible and efficient manner online.

References and Notes

1. Project Tomorrow, 2011 Speak-Up Report (Project Tomorrow, Irvine, CA, 2012); www.tomorrow.org/speakup/zSU11_PersonalizedClassroom_EducatorsReport.html.
2. 2010 Speak-Up Report (2011); www.tomorrow.org/speakup/pdfs/SU10_3EofEducation_Educators.pdf.
3. L. L. Zia, *D-Lib* **6**, (2000). www.dlib.org/dlib/october00/zia/10zia.html10.1045/october2000-zia.
4. Including comPADRE (Physics), ChemEd DL (Chemistry), CLEAN (Climate and Energy), AMSER (Technical and Community Colleges), Engineering Pathway, Teachers' Domain (Media), MSP2 (Middle School Teachers), CPALMS (Florida K-12 Teachers), DLESE (Geosciences), and Howtosmile.org.
5. NSDL LAR, https://wiki.ucar.edu/display/nsdldocs/LAR+Concepts.
6. LRMI, www.lrmi.net/about.
7. Wikimedia Foundation, 2011–2012 Annual Report; http://wikimediafoundation.org/wiki/FAQ/en.
8. NASA Wavelength.org Review Criteria; http://nasareviews.strategies.org/reviewcriteria.html.
9. Project 2016, www.project2061.org/publications/bsl/.
10. CPALMS review process: www.cpalms.org/cpalms/review.aspx.
11. Howtosmile.org lists: http://howtosmile.org/userlists/recent.
12. Gooru, http://about.goorulearning.org/.
13. InBloom, https://www.inbloom.org/.
14. Sparticl from Twin Cities Public Television, www.sparticl.org/about/.
15. CAISE's Informalscience.org project, http://informalscience.org/perspectives/blog/welcome-to-informalscienceorg.
16. The Learning Registry, http://learningregistry.org.
17. Open Science Resources (OSR) Project, www.openscienceresources.eu.
18. NSDL, Paradata; https://wiki.ucar.edu/display/nsdldocs/Paradata.
19. Linked Open Data—What is it? A video from Europeana; http://vimeo.com/36752317.

Acknowledgments: Authors are affiliated with Howtosmile.org, NASAWavelength.org, and Informalscience.org.

10.1126/science.1234722