

# DOMAIN MODEL AND META-LANGUAGE FOR PEER REVIEW AND ASSESSMENT

*Dmytro Babik, Department of Computer Information Systems and Business Analytics,  
James Madison University, 421 Bluestone Drive, Harrisonburg, VA 22807, 540-568-3064,  
[babikdx@jmu.edu](mailto:babikdx@jmu.edu)*

*Edward F. Gehringer, Department of Computer Science, North Carolina State University,  
890 Oval Drive, Raleigh, NC 27695, 919-515-2858, [efg@ncsu.edu](mailto:efg@ncsu.edu)*

*David Tinapple, Herberger Institute for Design and the Arts, Arizona State University,  
Dixie Gammage Hall, Tempe, AZ 85287, 480-965-3122, [david.tinapple@asu.edu](mailto:david.tinapple@asu.edu)*

*Ferry Pramudianto, Department of Computer Science, North Carolina State University, 890 Oval Drive,  
Raleigh, NC 27695, 919-515-2858, [fferry@ncsu.edu](mailto:fferry@ncsu.edu)*

*Yang Song, Department of Computer Science, University of North Carolina Wilmington,  
601 South College Road, Wilmington, NC 28403, 910-962-7701, [songy@uncw.edu](mailto:songy@uncw.edu)*

## ABSTRACT

We propose a framework for research and development of Online Peer Review and Assessment (OPRA) systems that combines a common research-oriented data model labeled Peer Review Domain Model (PRDM) and an open-source data exchange specification labeled Peer Review Meta-Language (PRML). This framework received the approbation of several OPRA systems and have been used to amass research data via the Educational Peer Review Data Warehouse. We invite feedback on the structure of the framework and the use of the data warehouse from researchers interested in educational and academic peer review.

**Keywords:** peer assessment, learning analytics, domain model, meta language, data exchange protocol.

## INTRODUCTION

Over the last several decades, practices of educational peer review and assessment have evolved from simple face-to-face paper-and-pencil exercises within small groups of learners to sophisticated online learning environments with a wide spectrum of functionality, capable of serving large traditional and online classes [27] [13]. Multiple studies have demonstrated learning benefits of online peer review in assessment in a variety of disciplines, e.g., accounting [19], engineering [14] [20], and social science [1] [8]. These benefits include better understanding of assessment criteria, timely and varied feedback, reducing instructors' overload, promoting active learning, developing transferable skills, and cultivating life-long learners. Since the late 1990s, plethora of web applications have been designed and implemented around the world to enable more efficient and effective student peer review. The better known, professionally built and widely used systems include Aropä [10], Calibrated Peer Review [21], CritViz [26], CrowdGrader [7], Expertiza [9], Mobius SLIP [4], Peerceptiv/SWoRD [6], peerScholar [11], and SPARKPlus [24]. These systems are used by educators in a wide variety of courses and disciplines, such as English and writing [6], computer science [22], business [5], visual art and design [26]. Similarly, in other domains, such as academic journals and conference proceedings, where peer review is the primary approach to assessing manuscript quality and making decisions regarding publishing, web-based systems (e.g., ManuscriptCentral.com, ScholarOne.com) replace paper- and email-based processes. Moreover, the open-access movement [17], [25], [18] and academic social

networks (such as ResearchGate.net, Academia.edu) further transform peer-based social creation and refinement of academic knowledge by allowing more agile idea exchange and feedback communication in the research community. Given the increasing use and impact of IT-enabled peer review on education and academia, it is relevant and timely to develop a common framework for researching phenomena occurring in these socio-technical systems of learning and knowledge creation.

We use the term *Online Peer Review and Assessment (OPRA) system* to describe a web-based software application purposefully designed and developed to facilitate and automate peer review and assessment process in any domain. OPRA systems are used to enable and support elaborate logistics involved in the peer review process, such as collecting submission artifacts, assigning reviewers who critique and/or evaluate designated artifacts submitted by peers, setting deadlines, guiding reviewers on the format of critiques and evaluations, and communicating feedback to authors. OPRA systems are a subclass of social computing systems used for peer-based social knowledge creation, including social networking and social media applications, such as wikis, blogs, and discussion forums. OPRA systems, however, are distinguished by having specific workflow constraints and being directed at specific educational or publishing goals.

The availability of massive data on peer review and assessment in various domains grants opportunities for “big data” research to answer important general research questions related to validity, reliability, motivation, use of different scales, rubrics, workflows and other aspects of peer review. The major challenge to such endeavor, however, is a great variety of peer review practices and processes, as well as the lack of the common data model and meta-language to describe them. To address this challenge, we propose a framework for research and development of OPRA systems that combines a common research-oriented data model and a meta-language. The purpose of this paper is to present a community-based data model labeled *Peer Review Domain Model (PRDM)*, pronounced “pridem” /'pri:.dem/) and an open-source data exchange specification labeled *Peer Review Meta-Language (PRML)*, pronounced “primal” /'prīməl/), which have received the approbation of several OPRA systems and have been used to amass research data via the Educational Peer Review Data Warehouse.

## **BACKGROUND OF THIS STUDY AND RELATED RESEARCH**

Despite the fact that dozens of online peer assessment systems are in use, virtually all studies on online peer assessment derive their conclusions from data taken from a single system. A few surveys have tried to compare different systems, but these comparisons were very superficial because they were based on the systems' feature specifications without any analyses of data generated by different systems [3], [13]. This is a consequence of lacking protocols for sharing data generated by different peer assessment systems.

Moreover, different systems implement the process of peer review in different ways. For example, some systems (CritViz, Mobius SLIP) give each reviewer several submissions and ask to rank them, whereas other systems (Calibrated Peer Review PeerCeptiv, Expertiza) give each reviewer one submission at one time and ask the reviewer to rate it based on different aspects of the work. There is no easy way to compare ranking-based evaluations from one system with ratings from another since these data structures vary widely. OPRA system designs also vary in many other respects. To produce data sets combining data from different systems and to conduct meaningful analyses, researchers need to thoroughly understand the design and intricacies of each system. Many research questions in peer review, such as comparing the effects of using different types of scales, rubrics, reviewer allocation

methods, anonymity, cannot be addressed without analyzing large data sets containing sufficiently general data obtained from different systems.

Only a few attempts have been made to create *generalized models* of peer review and assessment. Millard and colleagues [15] [16] analyzed various peer review processes (and, more specifically, reviewer *allocation patterns*) and proposed a “*canonical model*” integrating a set of “*peer review cycles*”, each of which defined by a set of “*peer review transforms*”. Based on this model they created a prototype of a generalist web-based OPRA system called PeerPigeon and a Domain-Specific Language (DSL). The project, however, has not been completed and apparently has been discontinued. Babik [2] proposed a general model of social, peer-review-based, knowledge creation, evaluation and refinement model, with the focus on analyzing intersubjectivity of peer evaluations. This projects, however, did not address the creation of a unified data model and data exchange specification. Our research aims at filling this gap.

### **PROPOSED PEER REVIEW DOMAIN MODEL (PRDM)**

To facilitate amassing and sharing research data, originators of five educational OPRA systems (Expertiza, CritViz, Mobius SLIP, CrowdGrader, Peerceptiv/SWoRD, Calibrated Peer Review), collaborating under the umbrella of the NSF-sponsored *PeerLogic* project ([www.peerlogic.org](http://www.peerlogic.org)), developed *Peer Review Domain Model* (PRDM). The class diagram in figure 1 is a graphic representation of PRDM. This domain model evolved from, but is not confined to, educational and academic peer review and assessment. It is envisioned as a generic formal conceptual model that represents concepts, actors, relationships and data types pertaining to any peer review process. This common vocabulary allows researchers, system designers, practitioners and other stakeholders from any educational or academic domain to describe peer review processes and outcomes using the same vocabulary. Furthermore, by using its common set of concepts, this PRDM provides an overview of how online peer assessment works in practice across different systems. Finally, PRDM is a foundation of the Peer Review Meta-Language that serves as a data-sharing specification and can be used to build research data sets and shareable “data warehouses” to hold data produced by different OPRA systems.

### **PROPOSED PEER REVIEW META-LANGUAGE (PRML)**

To facilitate data exchange, creation of research data sets and building the Educational Peer Review Data Warehouse, based on the PRDM described above, we are developing the Peer Review Meta-Language (PRML). Using this meta-language, researchers should be able to easily convert data from any specific OPRA system to the common meta-language specification data set and merge data from different systems. The major advantages of using data obtained from a variety of systems are that (a) such data sets containing numeric and textual entries are larger than any data set from an individual system; (b) the variety of different peer-review set-ups allows researchers to test general hypotheses and make more generalizable conclusions.

The PRML specification is grounded in the following *guiding principles*. It should be:

- *generic* (i.e., able to describe outcomes of any formal or informal peer review process);
- *free and non-proprietary* (the format should reside in the public domain and be maintained and updated by its users);
- *human intelligible* and *self-descriptive*, without extensive technical documentation (tags are explicit English words or phrases; users should be able to understand how the data are organized by reading the tags);

- *easily transformable/convertible* and *machine readable* (the specification can be easily converted into JSON, XML or plain-text formats);
- *compact* (using as few tags as possible with as short keywords as possible);
- *complete* (using enough tags to describe all possible variations of the process and product);
- *extensible* (keywords for additional entities and attributes can be introduced without jeopardizing the overall integrity of the PRDM and PRML; the *PeerLogic* project assumes the role of the specification governance agency that can endorse changes and document history of modifications).

In addition, dissemination of the PRML format should be based on the community buy-in rather than the formal requirement to use it.

The current PRML tag dictionary is available online at [goo.gl/CVuSTu](http://goo.gl/CVuSTu).

## **EDUCATIONAL PEER REVIEW DATA WAREHOUSE**

To validate proposed PRDM and PRML, we developed the Educational Peer Review Data Warehouse that can be used to amass and share data from different OPRA systems. Currently, the data warehouse contains tens of thousands of anonymized peer-review records from two systems (with other four systems getting ready to contribute their data), that can be mined for research purposes. The data warehouse was designed based on the dimensional modeling approach [12]. It stores facts (artifacts, measurements and metrics) from the peer-review process in fact tables that hold references (foreign keys) to the dimension tables. The dimension tables contain groups of hierarchies and descriptors that define the facts. The dimensions can be used to group the facts into multidimensional arrays of data, known as *OLAP cube*, or *hypercube*. Because the dimensional modeling approach was chosen, the data warehouse schema permits a star topology, in which fact tables are placed in the center [23].

We implemented the initial version of the data warehouse using MySQL for several following reasons. First, it allows to use a mainstream query language SQL and more mature tools compared to NoSQL databases. Secondly, most peer review systems that we know still use relational databases, therefore mapping them to the relational data warehouse would be simpler and less risky than using NoSQL approach. Third, we anticipated, based on the past growth of several systems, that the amount of aggregated data will not exceed 100GB within the next three years, and therefore MySQL would be adequate to serve our needs.

The data warehouse is currently accessible through the MySQL server at [www.peerlogic.org/data-warehouse](http://www.peerlogic.org/data-warehouse). Read-only credentials are provided upon request. In the future, these data will be accessible through a RESTful web service and visualized through the [peerlogic.org](http://peerlogic.org) website. Since the data warehouse receives data from widely used systems, the amount of data is extensive in many dimensions: the number of participants, the number of peer reviews, the diversity of peer-review processes, and the range of students' background and level. Researchers can mine this dataset to derive general conclusions rather than those that are just class- or task-specific.

## **EXPECTED RESULTS AND FUTURE DIRECTIONS**

The described work is currently in progress. We actively solicit feedback from researchers and practitioners interested in educational and academic peer review regarding structure of PRDM and PRML and the use of the data warehouse. The PeerLogic project currently involves six OPRA systems, of which two have already contributed data and four others are in the process of preparing their data. In the future, we expect more OPRA systems and researchers to participate in this project. The data

warehouse will offer extensive data sets with more dimensions, such as number of users, range of educational levels, number of majors, etc. Researchers will have access to these data sets for mining and analyses. Currently, data can only be accessed via database queries. We are also creating a web visual query builder to facilitate data access to the data warehouse for researchers seeking to test their hypotheses.

The paradigm of education is gradually shifting towards greater focus on higher-level competencies, such as creative problem solving, critical thinking, communication and collaboration. Technologies that enable students' social learning interactions to build these competencies play the key role. The scope of developing these technologies and researching learning outcomes they produce reaches beyond any individual system. In addition, synergies of researching educational and academic peer review practices offer new opportunities for building generalized knowledge. Therefore, we believe that the proposed open-source, community-supported framework – the domain model and the meta-language – will transform the ways in which research of academic and educational peer review and assessment is conducted and invite contributions based on this framework.

### ACKNOWLEDGEMENT

This research is a part of the PeerLogic project sponsored by the NSF IUSE grants DUE-1431856, DUE-1431975, DUE- 1432347, DUE- 1432580, DUE- 1432690.

### REFERENCES

- [1] Althausen, R. & Darnall, K. Enhancing critical reading and writing through peer reviews: An exploration of assisted performance. *Teaching Sociology*, 2001, 29(1), 23-35.
- [2] Babik, D. Investigating intersubjectivity in peer-review-based, technology-enabled knowledge creation and refinement social systems. *Dissertation, University of North Carolina at Greensboro*, 2015.
- [3] Babik, D., Gehringer, E., Kidd, J., Pramudianto, F. & Tinapple D. Probing the landscape: Toward a systematic taxonomy of online peer assessment systems in education. *In Educational Data Mining Conference Workshops*, 2016.
- [4] Babik, D., Iyer, L. & Ford E. Towards a comprehensive online peer assessment system. *In International Conference on Design Science Research in Information Systems 2012*, Springer Berlin Heidelberg, 1-8,
- [5] Campbell, K., Mothersbaugh, D., Brammer, C. & Taylor T. Peer versus self assessment of oral business presentation performance. *Business Communication Quarterly*, 2001, 64(3), 23-40.
- [6] Cho, K. & Schunn, C. Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 2007, 48(3), 409-26.
- [7] de Alfaro, L. & Shavlovsky, M. CrowdGrader: A tool for crowdsourcing the evaluation of homework assignments. *In Proceedings of the 45th ACM technical symposium on Computer science education, ACM*, 2014, 415-420.
- [8] Falchikov, N. Peer feedback marking: Developing peer assessment. *Innovations in Education Training International*, 1995, 32(2), 175-187.
- [9] Gehringer, E., Ehresman, L., Conger, S. & Wagle, P. Reusable learning objects through peer review: The Expertiza approach. *Journal of Online Education*, 2007, 3(5), 4.
- [10] Hamer, J., Kell, C. & Spence, F. Peer assessment using Aropä. *In Proceedings of the 9th Australasian Conference on Computing Education*, Australian Computer Society, Inc., 2007, 66, 43-54.

- [11] Joordens, S., Desa, S. & Paré, D. The pedagogical anatomy of peer-assessment: Dissecting a peerScholar assignment. *Journal of Systemics, Cybernetics & Informatics*, 2009, 7(5).
- [12] Kimball, R. & Ross, M. The data warehouse toolkit: the complete guide to dimensional modeling. *John Wiley & Sons*, 2011.
- [13] Luxton-Reilly, A. A systematic review of tools that support peer assessment. *Computer Science Education*, 2009, 19(4), 209-232.
- [14] MacAlpine, J. Improving and encouraging peer assessment of student presentations. *Assessment & Evaluation in Higher Education*, 1999, 24(1), 15-25.
- [15] Millard, D., Fill, K., Gilbert, L., Howard, Y., Sinclair, P., Senbanjo, D. & Wills, G. Towards a canonical view of peer assessment. In *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies (ICALT)*, 2007, 793-797.
- [16] Millard, D., Sinclair, P. & Newman, D. PeerPigeon: A web application to support generalised peer review. 2008.
- [17] Mizzaro, S. Quality control in scholarly publishing: A new proposal. *Journal of the Association for Information Science and Technology*, 2003, 54(11), 989-1005.
- [18] Pandita, R. Open Access: A Global Public Movement. *Journal of Indian Library Association*, 34.
- [19] Persons, O. Factors influencing students' peer evaluation in cooperative learning. *Journal of Education for Business*, 1998, 73(4), 225-229.
- [20] Rafiq, Y. & Fullerton, H. Peer assessment of group projects in civil engineering. *Assessment & Evaluation in Higher Education*, 1996, 21(1), 69-81.
- [21] Russell, A. Calibrated peer review-a writing and critical-thinking instructional tool. *Teaching Tips: Innovations in Undergraduate Science Instruction*, 2004, 54.
- [22] Søndergaard, H. & Mulder, R. Collaborative learning through formative peer review: pedagogy, programs and potential. *Computer Science Education*, 2012, 22(4), 343-67.
- [23] Song, Y., Pramudianto, F. & Gehringer, E. A markup language for building a data warehouse for educational peer-assessment research. In *IEEE Frontiers in Education Conference (FIE)*, 2016, 1-5.
- [24] SparkPlus, University of Technology, Sydney, 2006, Accessed November 27, 2016, [sparkplus.com.au](http://sparkplus.com.au)
- [25] Suber, P. Open access overview, 2007.
- [26] Tinapple, D., Olson, L., Sadauskas, J. CritViz: Web-based software supporting peer critique in large creative classrooms. *Bulletin of the IEEE Technical Committee on Learning Technology*, 2013, 15(1), 29
- [27] Topping, K. Peer assessment between students in colleges and universities. *Review of Educational Research*, 1998, 68(3), 249-276.

# Figure 1 – Peer Review and Assessment Domain Model (PRDM) Class Diagram

