

Chapter 6

Data Sharing in CSCR: Towards In-Depth Long Term Collaboration

Christophe Reffay

Ecole Normale Supérieure de Cachan, France & Ecole Normale Supérieure de Lyon, France

Gregory Dyke

Carnegie Mellon University, USA

Marie-Laure Betbeder

Université de Franche-Comté, France

ABSTRACT

In this chapter, the authors show the importance of data in the research process and the potential benefit for communities to share research data. Although most of their references are taken from the fields of Computer Supported Collaborative Learning and Intelligent Tutoring Systems, they claim that their argument applies to any other field studying complex situations that need to be analyzed by different disciplines, methods, and instruments. The authors point out the evolution of scientific publication, especially its openness and the variety of its emerging forms. This leads them to propose corpora as boundary objects for various communities in the scientific sphere. Data release being itself a complex problem, the authors use the Mulce¹ experience to show how sharable data can be built and made available. Once corpora are considered available, they discuss the potential of their reuse for multiple analyses or derivation. They focus on analytic representations and their combination with initial data or complementary analytic representations by presenting a tool named Tatiana. Finally, the authors propose their vision of data sharing in a world where scientists would use social network applications.

DOI: 10.4018/978-1-4666-0125-3.ch006

INTRODUCTION

In the research process, data is crucial and often hard to collect. Researchers spend a lot of time designing studies and collecting, transforming, analyzing, or interpreting data. Once analyzed and communicated in some publication by a local research team, data is often lost and can't be re-used by anybody. This means that other researchers have no access to original data to deepen their understanding by replicating an analysis or comparing their own results on the same data with a slightly different analysis method.

In this chapter, we would like to draw the state of the art in data sharing among research communities and, in particular, to report the results of the Mulce project¹. This project's main results are the design and creation of a data structure and a corresponding platform to share learning and teaching corpora. These results give the community a way to access, share, analyze and visualize learning and teaching corpora.

This work has been motivated by the lack of impact of research results in the real world of on-line learning. In the CSCL (Computer Supported Collaborative Learning) research for example, a very wide range of indicators on collaboration have been designed and prototyped in a particular context but almost none of them are reused in other situations or contexts. We argue in this chapter that our research community should be able to widen the validity of its results by sharing data, tools and analyses performed with these tools.

In their work on the coding and counting analysis methodology, Rourke, Anderson, Garisson, and Archer (2001) have pointed out the weakness of our research domain. Replicability, reliability and objectivity need to be improved in our work. The main idea of research collaboration is well expressed by (Chan, et al., 2006) in the following terms:

“There is urgent need of putting together complementary strengths and contexts and combining our

insights as rapidly as possible to make a greater impact and further elevate our research quality at the same time. Research generally has had a small voice in national educational outcomes; we can speak louder if we speak together.” (Chan et al., 2006)

Considering e-Research as an efficient way to meet and collaborate, this chapter suggests that e-collaboration could provide emerging communities with tools and virtual places to actually share their data, analyses and results in order to improve their theories, knowledge and tools. Although the focus of our work is on CSCL, we argue that this proposal is not limited to this domain or even to its contributing disciplines, and that the core ideas and benefits of our proposal can be extrapolated to other fields of research.

In the remainder of the chapter, we first examine current trends in scientific publication and the central role played by data in the scientific process. We then highlight the particular problem posed by data collection and replication in learning-related research and examine the state of the art for data sharing within this context. The Mulce proposal for constructing and sharing learning and teaching corpora is presented in detail, followed by the Tatiana framework for creating and re-using analytic representations. We conclude by drawing up our vision of data sharing within the learning sciences field and describe how other fields can draw upon our experience to construct data and analysis sharing models of their own.

Evolution in Scientific Publication

Who is producing knowledge? Nowadays, this process is no longer limited to academic researchers and prestigious journals. Civil society including local, national and international organizations is bringing its truth in various areas like economy, education, environment science, etc. Forms of publication are also evolving from classical journal articles to virtual exhibitions, datasets, software

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/data-sharing-cscr/63506

Related Content

Understanding Consumer Intention Towards Blockchain-Based Mobile Payment Adoption Services in Pakistan

Zulfiqar Aliand Hasin Bano (2022). *International Journal of e-Collaboration* (pp. 1-21).

www.irma-international.org/article/understanding-consumer-intention-towards-blockchain-based-mobile-payment-adoption-services-in-pakistan/307136

An Ontological Structure for Gathering and Sharing Knowledge among Scientists through Experiment Modeling

Luis Casillasand Thanasis Daradoumis (2012). *Collaborative and Distributed E-Research: Innovations in Technologies, Strategies and Applications* (pp. 165-179).

www.irma-international.org/chapter/ontological-structure-gathering-sharing-knowledge/63508

The Roles of Peripheral Participants and Brokers: Within and Beyond Communities of Practices

François Grimaand Emmanuel Josserand (2011). *Handbook of Research on Communities of Practice for Organizational Management and Networking: Methodologies for Competitive Advantage* (pp. 297-307).

www.irma-international.org/chapter/roles-peripheral-participants-brokers/52906

Meta-Analysis Research on Virtual Team Performance

Ying-Chieh Liu, Janice Burnand Susan Stoney (2009). *E-Collaboration: Concepts, Methodologies, Tools, and Applications* (pp. 1349-1361).

www.irma-international.org/chapter/meta-analysis-research-virtual-team/8868

Using Role Play and Team Teaching as Strategies to Add Depth to Online Discussion

Ron Lombardand Barbara Biglan (2011). *Techniques for Fostering Collaboration in Online Learning Communities: Theoretical and Practical Perspectives* (pp. 164-182).

www.irma-international.org/chapter/using-role-play-team-teaching/46912