The Collage Authoring Environment: A Platform for Executable Publications

Eryk Ciepiela[1], Piotr Nowakowski[1], Daniel Harężlak[1], Marek Kasztelnik[1], Grzegorz Dyk[1], Jan Meizner[1], Marian Bubak[1,2]

[1] AGH University of Science and Technology, ACC Cyfronet AGH, Nawojki 11, 30-950, Kraków, Poland
[2] AGH University of Science and Technology, Institute of Computer Science AGH, Department of Computer Science, al. Mickiewicza 30, 30-059 Kraków, Poland
emails: bubak@agh.edu.pl
{e.ciepiela,p.nowakowski,d.harezlak,m.kasztelnik,g.dyk, j.meizner}@cyfronet.pl

Keywords: executable publications, scientific publishing, reproducible research, e-science, virtual laboratory

1. Introduction

The Collage Authoring Environment [1] is a software platform which enables e-scientists to carry out research and publish their research in the form of scientific executable papers. Recent developments in both e-science and computational technologies call for a novel publishing paradigm where static content of a traditional scientific publication is to be supplemented with elements that enable reviewers and readers to reexamine the reported computational method. This includes re-execution of a method against original or user-provided data, modification and customization of a method to suit specific use case, and, ultimately, reproducibility of a method.

2. Description

We propose a platform which enables authors to seamlessly embed chunks of executable code and data into scientific publications and allow repeated execution of such assets on underlying computing and data storage resources, as required by scientists who wish to build upon the presented results. The Collage Authoring Environment is based on GridSpace2 Virtual Laboratory [2] that is used as an execution engine and gateway to arbitrary computational resources, including those belonging to high performance computing centers, scientific e-infrastructures and resources contributed by the scientists themselves. The platform provides access to static content, primary datasets and executable assets.

3. Results

The platform allowed for implementation of a number of executable papers as one depicted in Fig. 1. Every GridSpace2 experiment can be released as an executable paper by using special “embed codes” that place Collage Widgets on an arbitrary web site. The example executable paper [3] is published on Wordpress web blog platform and is backed by a dedicated installation of the GridSpace2 Virtual Laboratory [4], that is available for demo users.

4. Conclusions and future work

The platform on its prototype stage proved promising winning Executable Paper Grand Challenge in 2011. Since that time, additional effort has been invested in completeness and robustness of the solution. Nowadays it's delivered as a service in ACC CYFRONET AGH. Moreover, the pilot project has been launched in cooperation with Elsevier – worldwide scientific publisher, whose aim are special issues of journals featured by executable papers powered by the Collage Authoring Environment.

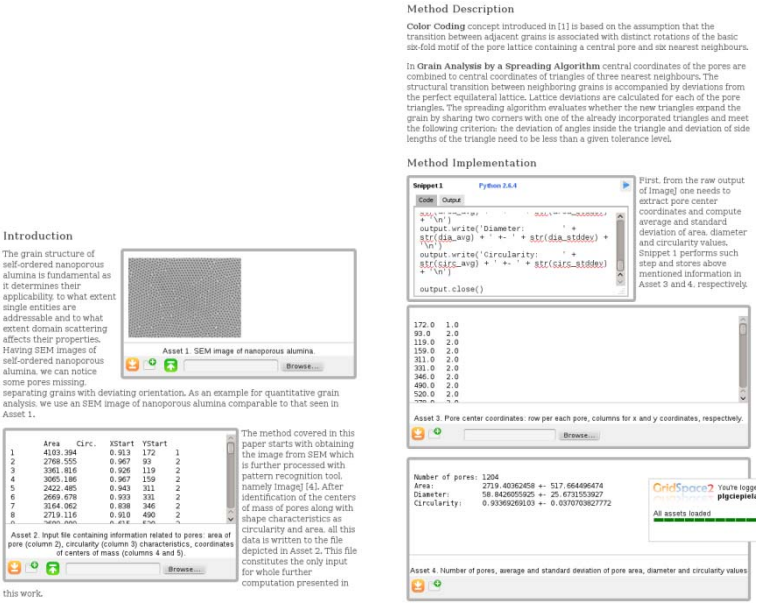

Fig. 1. Example on-line executable paper (fragment): "Implementation of Algorithms of Quantitative Analysis of the Grain Morphology in Self-Assembled Hexagonal Lattices according to Hillebrand method" [3]. Web page content is enriched with Collage Widgets that enables interactive execution of computations and insight into experiment input out output files.

References

1. P. Nowakowski, E. Ciepiela, D. Harężlak, J. Kocot, M. Kasztelnik, T. Bartyński, J. Meizner, G. Dyk, M. Malawski: The Collage Authoring Environment. In: Proceedings of the International Conference on Computational Science, ICCS 2011 (2011).
2. E. Ciepiela, P. Nowakowski, J. Kocot, D. Harężlak, T. Gubala, J. Meizner, M. Kasztelnik, T. Bartynski, M. Malawski, M. Bubak: Managing Entire Lifecycles of e-Science Applications in the GridSpace2 Virtual Laboratory - From Motivation through Idea to Operable Web-Accessible Environment Built ona Top of PL-Grid e-Infrastructure. In: M. Bubak, T. Szepieniec, K. Wiatr (Eds) Building a National Distributed e-Infrastructure - PL-Grid - Scientific and Technical Achievements, Springer 2012, ISBN 978-3-642-28266-9, pp. 228-239 (2012).
3. E. Ciepiela, L. Zaraska, G. D. Sulka: Implementation of Algorithms of Quantitative Analysis of the Grain Morphology in Self-Assembled Hexagonal Lattices according to Hillebrand method, example executable paper powered by Collage, `http://gs2.cyfronet.pl/epapers/hillebrand-grains/`
4. Dedicated instance of GridSpace2 Virtual Laboratory provided by ACC CYFRONET AGH, `https://gs2.cyfronet.pl`