



# The Collage Authoring Environment: from proof-of-concept prototype to pilot service

Eryk Ciepiela<sup>1</sup>, Daniel Harezlak<sup>1</sup>, Marek Kasztelnik<sup>1</sup>, Jan Meizner<sup>1</sup>,  
Grzegorz Dyk<sup>1</sup>, Piotr Nowakowski<sup>1</sup>, and Marian Bubak<sup>1,2,3</sup>

<sup>1</sup>Academic Computer Center Cyfronet AGH, Distributed Computing Environments Team

<sup>2</sup>Department of Computer Science AGH, University of Science and Technology

<sup>3</sup>Informatics Institute, University of Amsterdam

# Outline

- Motivation is to bring reproducibility to scientific papers
- Problem is how executable papers should be supported by technology
- Proposed solution is Collage Authoring Environment
- Result is a Special Issue of “Computers & Graphics” journal
- Conclusions is what and how one can benefit from Collage now and in the future

# Motivation



## executable paper grand challenge

knowledge enhancement in the **computational sciences**

- Increasing interest of the community in bringing reproducibility and reusability to scientific publishing
  - “R dimension” of research (vide: David De Roure)
  - “Science” special issue on data replication and reproducibility (December 2011)
  - “Science as an open enterprise” report by the Royal Society (June 2012)
- Intelligent openness of science induced by
  - Recent IT breakthroughs
  - Growth of the Internet
  - Changing social and political environment (preservation, data management plans, audits, sustainability policies etc.)
- Publishers seek new ways of overcoming limitations in the scholarly communication process
  - Elsevier “Article of the future” umbrella project
  - Executable Paper Grand Challenge
  - <http://www.youtube.com/watch?v=FQ5FaeHcdwo> (video)



# Problem

Propose a solution by which the **primary data** upon which scientific research is based, as well as the **computational methods** used in the course of scientific research, become **first class citizens** of the resulting publication, so that publications regain the **ability to stand on its own merits**, just as they did when the notion of scientific publications was first conceived.

Manifested in a set of specific objectives of the Elsevier Executable Paper Grand Challenge:

- Executability
- Compatibility
- Validation
- Licensing
- Computational Access
- Data Access
- Collaborative development support
- Multi-actor environment
- Evolutionary Approach
- Security
- Interoperability
- Customizability
- Extendability
- Adaptability

# Towards Solution

- GridSpace2 distributed computing platform
  - Developed since 2008
  - Web-oriented
  - Interactive workflow management system
  - Generic platform, specific applications accommodated as light “experiments” over it
- Collage prototype
  - Winner of the Executable Paper Grand Challenge in June 2011
  - Based on GridSpace2 platform
  - Publications associated with GridSpace2 experiments
  - Demo executable paper
- Collage pilot
  - Used by 3D object retrieval community for authoring first executable papers since May 2012
  - Full journal article life-cycle: submission, peer review, corrections and press in May 2013
  - Go-live on Elsevier ScienceDirect portal and articles made available to readers in May 2013
- GridSpace2 still used in the scope of other projects
  - MAPPER – Multiscale Applications on European e-Infrastructures since 2010  
<http://www.mapper-project.eu/>
  - PL-Grid/PL-Grid+ - Polish National Grid Initiative since 2008  
<http://www.plgrid.pl>



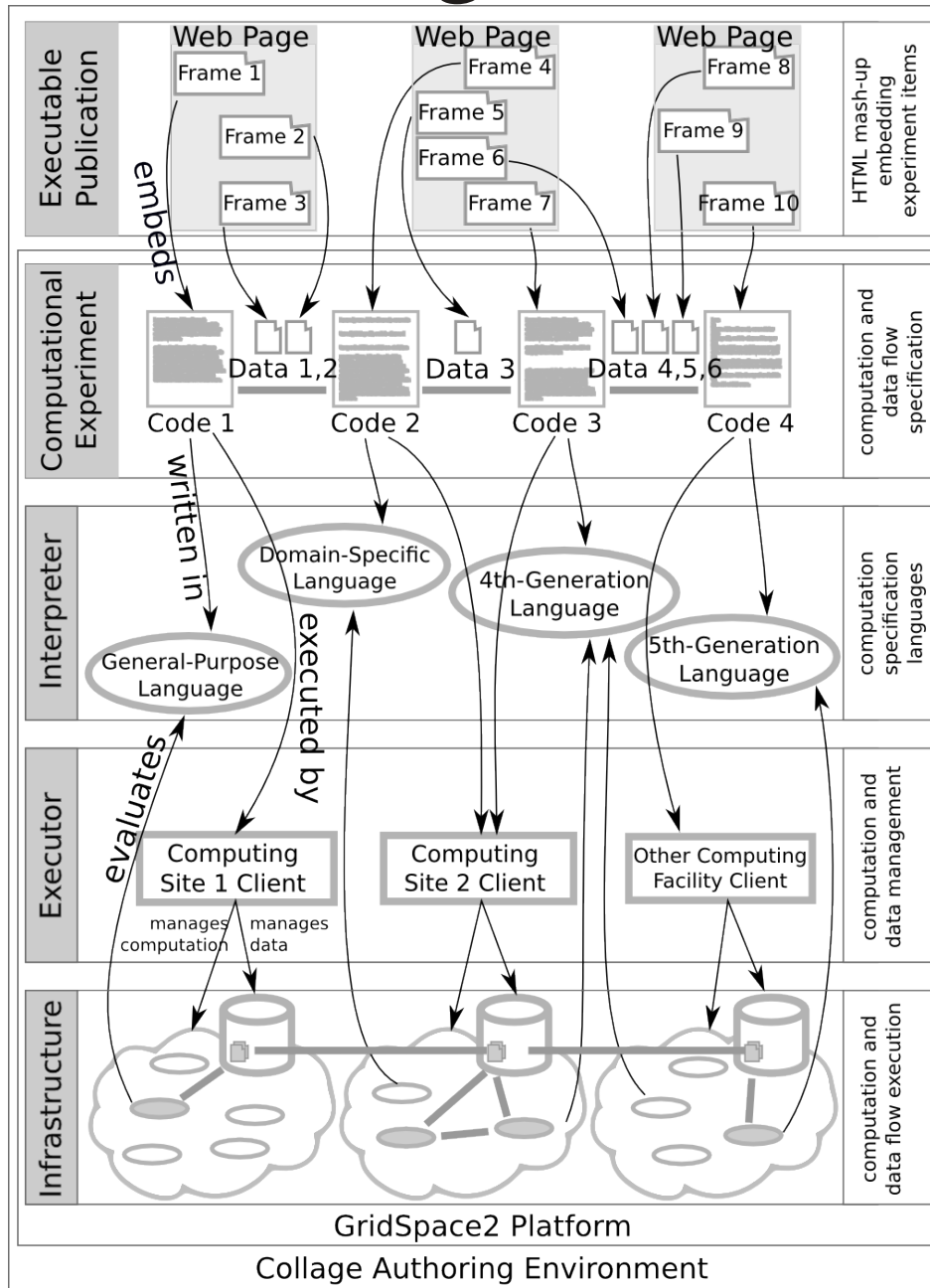
**executable paper  
grand challenge**

knowledge enhancement in the **computational sciences**

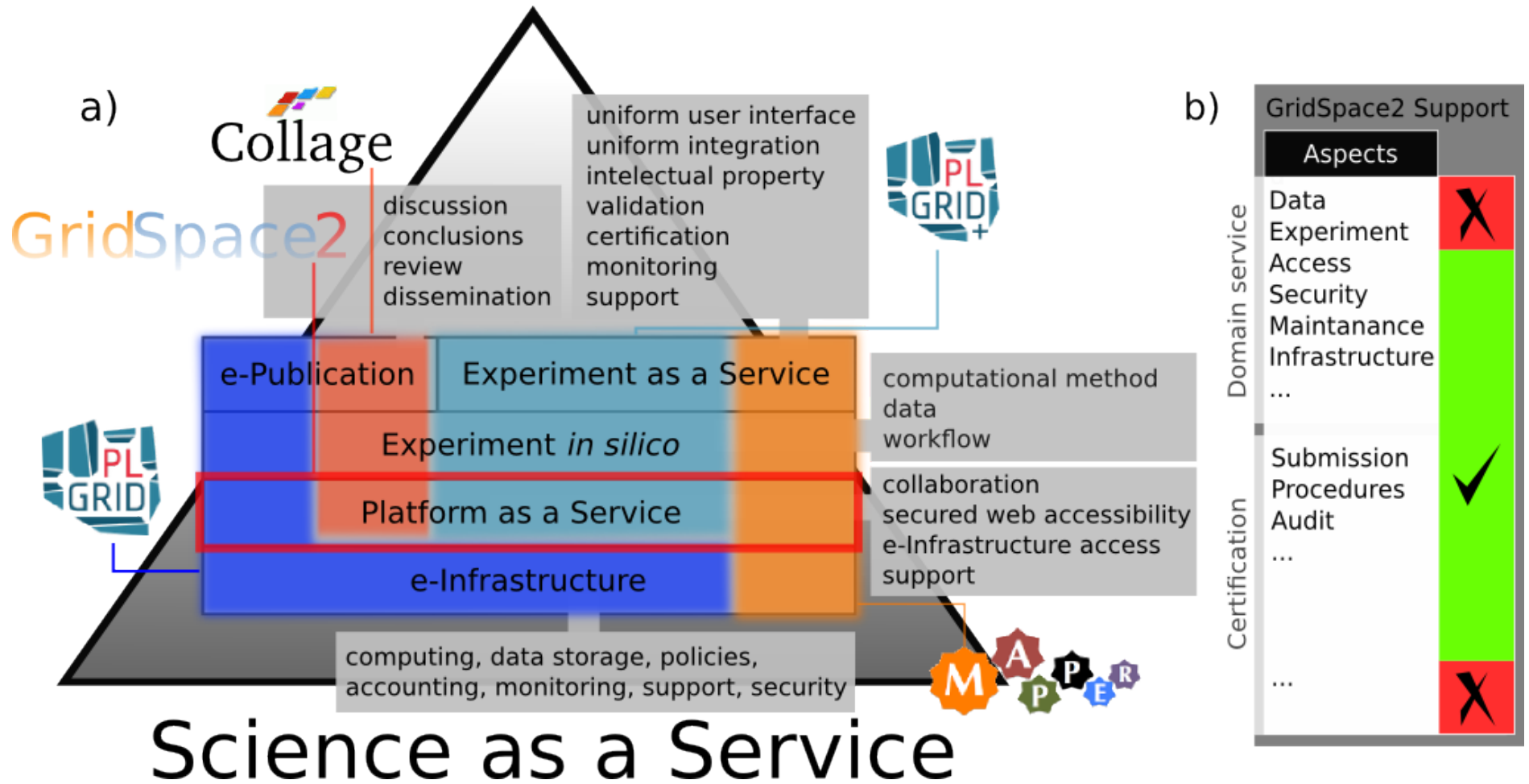
  
**Collage**  
Authoring  
Environment



# GridSpace2/Collage: Concept



# GridSpace2/Collage: “Science as a Service” Architecture



# GridSpace2/Collage: User Interfaces

Collage Authoring Workbench

You are logged in to [collage-exphost.elsevier.com](http://collage-exphost.elsevier.com) as [eciepiela](#)

© 2012 [Spectrum - IT Division](#) Please report any problems related to this portal to our [Issue Tracker](#) Acknowledgements

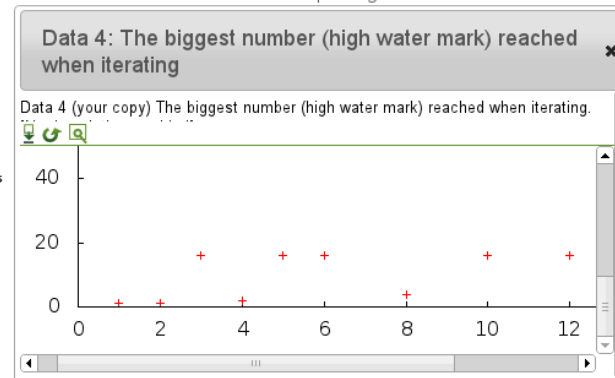
- Authoring (up) and previewing/reviewing (right) through a dedicated web platform: [Authoring Workbench](#)
- Access to article through ScienceDirect portal via [Collage SciVerse App](#) (video)

## The Collatz Conjecture

eciepiela, DOI: 10.0000/1358511059290

The experiment was released by [eciepiela](#) on **Fri Jan 18 13:10:59 CET 2013** in the **private** scope. No, below is not an article. It's only generated text with injected labels that navigate to particular experiment items to show you how in-text links work.

Euismod in pellentesque massa placerat dui ultricies lacus. Penatibus et magnis dis parturient. Ipsum consequat nisi vel pretium lectus. Interdum velit euismod in pellentesque. Nascetur ridiculus mus mauris vitae ultricies leo integer malesuada nunc. Sapien nec sagittis aliquam malesuada. Donec ultrices tincidunt arcu non sodales neque sodales. Sed velit dignissim sodales ut eu. [Data 1](#) Ac tortor vitae purus faucibus ornare. Mattis enim ut tellus elementum sagittis vitae et. Vitae semper quis lectus nulla at. Vitae purus faucibus ornare suspendisse sed nisi lacus sed viverra. Amet tellus cras adipiscing enim eu. [Code 1](#) Tellus integer



[Data 5](#) Non consectetur a erat nam at lectus urna dui. Diam volutpat commodo sed egestas egestas fringilla phasellus faucibus scelerisque. Magna eget est lorem ipsum dolor sit.



# Results

- Technology: web platform interoperable on presentation (http) and computational back-end (SSH, Grid, Cloud, ...) layers, based on Java EE platform
- Collage Authoring Environment pilot service on <https://collage.elsevier.com>
- “Computers & Graphics” Special Issue on 3D object retrieval (open access)

# Conclusions

## From computational experiment to e-publication

- Scientific relevance
- Originality
- Reproducibility
- Verifiability
- Transparency
- Primary data and results availability
- Support for review process
- Publication factors
- Publishing medium

## From computational experiment to service

- Integration with common e-infrastructure
- User access management
- Respected intellectual property rights
- Cataloging, indexing
- Accessibility
- Reusability
- Experiment availability
- Documentation availability
- Examples availability
- Monitoring
- Accounting
- Maintenance
- User support
- Quality assurance
- Security assurance

# Further information

- Check out Special Issue of “Computers & Graphics” on  
<http://www.journals.elsevier.com/computers-and-graphics/news/special-issue-on-3d-object-retrieval/>
- Read more and sign in for trial account on  
<https://collage.elsevier.com>
- Tell us what you think on  
<http://www.surveymonkey.com/s/DCWL2FV>
- Contact us through  
[eryk.ciepiela@cyfronet.pl](mailto:eryk.ciepiela@cyfronet.pl)
- More about us on  
<http://dice.cyfronet.pl>

# Thank You

- Elsevier team: Ann Gabriel, Hylke Koers, Gail Rodney, Rebecca Capone, Beate Specker
- Computers & Graphics editors: Joaquim Jorge, Michela Spagnuolo, Remco Veltkamp
- Distributed Computing Environments Team (DICE) in Cyfronet led by Marian Bubak
- All authors and reviewers having contributed to the Special Issue of “Computers & Graphics” on 3D object retrieval
- And You for your attention!

13<sup>th</sup> international meeting

# Cracow '13 Grid Workshop

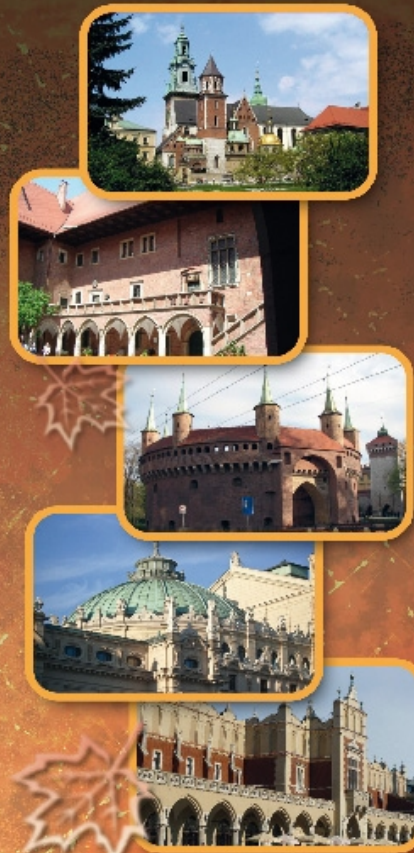
Kraków, Poland  
November 4-6, 2013

## TOPICS

- ▶ e-Science, system-level science and collaborative applications,
- ▶ models, methods and tools for collaborative applications development,
- ▶ virtual laboratories and problem solving environments,
- ▶ distributed computing infrastructures, grids and clouds,
- ▶ knowledge in e-Science and DCI systems,
- ▶ virtual organizations and security aspects,
- ▶ resource management and scheduling,
- ▶ monitoring and information management,
- ▶ software engineering aspects,
- ▶ industrial and social implications.

## KEYNOTE SPEAKERS

**I.T. Foster**, Argonne National Laboratory, US  
**F. Gagliardi**, Microsoft Research, CH  
**P. Guisset**, ERCIM Office, Brussels, BE  
**R. Hose**, University of Sheffield, UK  
**C. de Laat**, University of Amsterdam, NL  
**L. Matyska**, Masaryk University, CZ  
**A. Pasic**, Atos Research & Innovation, ES  
**B. Schuller**, Jülich Supercomputing Centre, DE  
**A.M.J. Skulimowski**, AGH Kraków, PL



## TUTORIAL

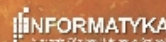
"Cloud Platform for VPN Applications"

## DEADLINES

**Sep 29, 2013** – short paper submission  
**Oct 4, 2013** – acceptance notification  
**Oct 16, 2013** – early registration  
**Nov 25, 2013** – extended paper submission

**MORE INFO** [www.cyfronet.krakow.pl/cgw13/](http://www.cyfronet.krakow.pl/cgw13/)

ORGANIZERS:



SPONSORS:

