

# Benchmarking Forum

Kresimir Duretec  
Vienna University of Technology  
Karlsplatz 13,  
1040 Wien,  
Austria  
duretec@ifs.tuwien.ac.at

Artur Kulmukhametov  
Vienna University of Technology  
Karlsplatz 13,  
1040 Wien,  
Austria  
artur.kulmukhametov@tuwien.ac.at

Andreas Rauber  
Vienna University of Technology  
Karlsplatz 13,  
1040 Wien,  
Austria  
rauber@ifs.tuwien.ac.at

Christoph Becker  
Vienna University of Technology & University of Toronto  
27 King's College Cir,  
Toronto, ON M5S,  
Canada  
christoph.becker@utoronto.ca

## ABSTRACT

This proposal is for a full day workshop to be held at the IPRES 2015 conference. The focus of the workshop will be on areas in the digital preservation field which could benefit from benchmarking. Benchmarking is a method of comparing entities against a well-defined standard (i.e. benchmark). The workshop is focused on discussing software benchmarking practices in digital preservation and how these can contribute to improving digital preservation tools

## General Terms

Frameworks for digital preservation; Preservation strategies and workflows; Innovative practice.

## Keywords

Digital preservation, benchmarks, benchmarkDP, evaluation, software tools, empirical evidence, workshop.

## 1. INTRODUCTION

The Digital Preservation field is characterized by a variety of tools used to collect and process digital information. The quality of those tools is of great importance to the preservation community. However, quality assessment is often done in an isolated way independent of systematic and community driven initiatives. This results in a lack of solid evidence to support quality claims.

iPres 2015 conference proceedings will be made available under a Creative Commons license.

With the exception of any logos, emblems, trademarks or other nominated third-party images/text, this work is available for reuse under a Creative Commons Attribution 3.0 unported license. Authorship of this work must be attributed. View a [copy of this license](#).

Benchmarking is a systematic method of facilitating comparisons of software artifacts according to a well-defined standard (benchmark). It has shown itself as a valuable empirical method for evaluating software tools. Various fields such as software engineering and information retrieval have reported major benefits from properly established benchmarking initiatives.

Several of the most successful, such as TREC<sup>1</sup> and CLEF<sup>2</sup>, have annual meetings where they report on benchmark results, refine already existing benchmarks and define new ones. However the overall successfulness of benchmarking is dependent on the readiness of the community to accept and drive the whole process.

There are several indicators showing that the digital preservation field and the community are mature enough and ready for establishing benchmarking as a community driven method. This would generate benefits, such as provision of solid evidence around the quality of digital preservation tools to enable improvements of those tools [2].

The goal of this workshop is to bring together key stakeholders in digital preservation to discuss the needs of benchmarking in the field, and to define and prioritize initial benchmarks. The workshop is imagined as the first workshop in a series of workshops where community members would report on benchmark results, define new and refine already existing benchmarks.

It is distinct from the Capability Assessment and Improvement workshop at IPRES 2013<sup>3</sup> which focused on the assessment and improvement of organizational capabilities [1]. In contrast, the focus of this workshop is on benchmarking software tools.

## 2. WORKSHOP FOCUS & OUTCOMES

The workshop is planned as a full day event. We would like to see participants from various domains with a common interest in software quality challenges in digital preservation. We will encourage submission of short position papers (two pages maximum) defining a specific benchmarking need, articulating the

---

<sup>1</sup> <http://trec.nist.gov/>

<sup>2</sup> <http://trec.nist.gov/>

<sup>3</sup> <http://benchmark-dp.org/caiwipres/>

challenges behind that benchmark and the benefits expected from implementing that benchmark in the community.

The workshop will be divided into two parts. The first part will consist of:

- An introduction session, where we give an overview of the workshop together with a brief introduction into benchmarking as a method, with illustrations and experiences from other fields;
- A session where participants present their benchmark proposals and perspectives. Each presentation will be accompanied by time for discussion.

It is expected that these two sessions will fill the time until lunch. After lunch we will focus on providing a discussion forum to brainstorm and discuss possible benchmarks. This will build on the morning session and will include defining each benchmark according to a common structure with defined components. In the case of a bigger number of submissions, proposals will be grouped and worked on in breakout sessions and each group will present their outcomes.

In the final wrap-up session, we plan to discuss the workshop outcomes, reflect on the possibility and challenges of establishing benchmarking in the digital preservation field, and distill specific steps forward and a future roadmap.

All the proposals submitted by the participants will be published online before the workshop. Following the workshop, we will write a workshop report for the D-Lib Magazine.<sup>4</sup> Finally, each defined benchmark will be released in a full specification published online to enable wider community involvement.

### 3. INTENDED AND EXPECTED AUDIENCE

This workshop will be open for all participants. Certain categories of expected participants can be listed, however:

- Stakeholders whose digital preservation operations are dependent on the high quality of software tools, who know specific aspects that could benefit from benchmarking. We encourage the proposal of real world scenarios considered important to the community.
- Tool developers familiar with a specific domain (e.g., validating PDF files or component integration in preservation systems).
- Researchers with an interest in software benchmarking, testing and test data generation are expected to bring fresh ideas for building certain benchmarks and addressing the challenges defined by other participants.

### 4. ORGANIZERS`

Kresimir Duretec is a Project Assistant at the Department of Software Technology and Interactive Systems (IFS) at the Vienna University of Technology. He is currently pursuing his PhD at the same department. Previously he graduated with an MSc and BSc in Computer Science from the University of Zagreb in 2011 and

2009 respectively. He previously worked as a Sub-project lead of the SCAPE Planning and Watch sub-project. Currently he is working on the project BenchmarkDP, where his main focus is on benchmarking digital preservation tools and automatic test dataset generation using model driven engineering principles.

Artur Kulmukhametov is Project Assistant and PhD student at the Software and Information Engineering Group, Vienna University of Technology. He was involved in the European research project SCAPE. His current focus in the project BenchmarkDP is the systematic evaluation of software tools for digital preservation, content profiling and quality assurance of migration processes.

Andreas Rauber is Associate Professor at the Department of Software Technology and Interactive Systems (IFS) at the Vienna University of Technology (TU-Wien). He furthermore is president of AARIT, the Austrian Association for Research in IT. His research interests cover the broad scope of digital libraries and information spaces, including specifically text and music information retrieval and organization, information visualization, data analysis, neural computation and digital preservation.

Christoph Becker is an Assistant Professor at the University of Toronto where he leads the Digital Curation Institute, and a Senior Scientist at the Software and Information Engineering Group at Vienna University of Technology in Austria. He was involved in the European research projects DELOS, PLANETS, DPE, and SHAMAN. He led the sub-project Scalable Planning and Watch in the SCAPE project and he is Principal Investigator of BenchmarkDP. His research focuses on digital libraries, digital curation and digital preservation, and sustainability in software engineering and information systems design.

### 5. PROCESS FOR CONTRIBUTION

The workshop is an interactive event focusing on discussions and inspirational exchanges between participants on software benchmarking, related challenges and opportunities in digital preservation. To achieve this goal, we will publish a call for participation and ask participants to submit a short abstract with their experiences and thoughts on benchmarks they wish to see discussed at the workshop. The review criteria for these submissions will be based upon for the overall suitability of the topics proposed.

### 6. ACKNOWLEDGMENTS

The workshop is supported by the Vienna Science and Technology Fund (WWTF) through the project BenchmarkDP (ICT12-046).

### 7. REFERENCES

- [1] C. Becker and E. Cardoso. Report on the Capability Assessment and Improvement Workshop (CAIW) at iPres 2013. D-Lib Magazine, 20(3/4), Apr. 2014.
- [2] K. Duretec, A. Kulmukhametov, A. Rauber, and C. Becker. Benchmarks for digital preservation tools. In Proceedings of IPRES 2015, 2015.

---

<sup>4</sup> <http://www.dlib.org/>