1. PREFACE
This volume contains the proceedings of DChanges 2013, the first International Workshop on Document Changes: Modeling, Detection, Storage and Visualization. The workshop will be held at the 13th ACM Symposium on Document Engineering (DocEng 2013) in Florence, Italy, in September 2013.

The goal of this workshop is to share ideas, common issues and principles about diff models and algorithms, change tracking, collaborative editing and document versioning. We want to look at these topics from different perspectives and want to understand which are the most common issues and which are the peculiarities of each domain and each approach.

1.1 Keynote
There is a great overlap between document engineering and software engineering with regard to these issues. The keynote, given by Ethan Munson and titled Collaborative Authoring Requires Advanced Change Management, is about such overlap and possible synergies.

The core idea is that authors collaborating on textual documents should have at least the same tools that software engineers use when collaborating on source code. With some important differences mainly due to some uncertainty and variability (for instance, the authors might not trust each other and might require a third-party validation) and lack of rigorous validating techniques (like compilers for software engineers). More important, non-technical users are not expected to have the same expertise of software engineers in using versioning control systems. The talk envisions a new generation of versioning and merging/branching tools tailored for non-expert users with simplified interfaces, more precise algorithms and more precise models for representing changes. It also describes some results on diff and merging algorithms (and in particular the idea of version-aware documents) that are going in this direction.

Bio. Ethan V. Munson is Professor and Co-Chair for Computer Science in the Department of EECS at the University of Wisconsin-Milwaukee. He received a Ph.D. in Computer Science from the University of California, Berkeley in 1994. He also holds a MS in Computer Science (UC Berkeley, 1989) and Bachelor degrees in Computer Science (UCSD, 1986) and Psychology (UCSD, 1978).

Dr. Munson’s research has focused on tools for managing software documents (style sheets, program editors, version control, and build systems) and on human-computer interaction (programming interfaces and system latency). Dr. Munson is a recipient of a National Science Foundation CAREER award, several research grants from industry, and four NSF educational grants. He is active in conference organization, especially with the ACM Symposium on Document Engineering, and was Chair of ACM SIGWEB from 2006 to 2011.

1.2 Research Papers
The core sessions of the workshop are on research papers. We received 12 submissions from all around the world, among which 7 papers were selected after a single-blind review process. The papers cover both practical and theoretical issues and can be clustered around three main topics: XML change management, understanding the evolution of non-textual documents and data structures and distributed collaborative authoring.

The first session deals with the management of changes in XML documents.

In Merging Uncertain Multi-Version XML Documents, the authors focus on versioning uncertain XML documents. The
In the second session, the focus is on documents and data structures that are more complex than simple text.

The paper *An Algorithm for Transforming XPath Expressions According to Schema Evolution* deals with evolving XML schemas and documents. It studies how to automatically update queries on XML documents, when these documents change to meet changes of their validating schemas. The authors present a novel algorithm based on tree automata, together with some experimental results. Though limited to only few XPath axes, the algorithm is very efficient and extensible.

The third session is on distributed collaborative editing. The paper *Concurrency Effects Over Variable-size Identifiers in Distributed Collaborative Editing* tackles the problem of building distributed editors with CRDT (Conflict-free Replicated Data Type). This approach consists of modeling a document as a sequence of items, each with a global unique identifier, and providing insert/delete operations on those items. The dynamic generation of these identifiers is a challenging problem, since limited consumption of resources and reliable propagation of identifiers must be guaranteed. The paper presents a novel strategy for such a generation that works well with a large number of users and has a very limited impact on latency.

In *Tracking Changes Through EARMARK: a Theoretical Perspective and an Implementation* the authors deal with changes on markup structures. Their goal is to define in a precise and unambiguous way when (and how) the same markup element has to be considered as changed, if its content changes. The authors propose a theoretical representation of change tracking information based on FRBR (Functional Requirements for Bibliographic Records), that also provides support for expressing provenance information. Their implementation of the framework is based on EARMARK, a Semantic Web-based meta-model that allows a fine-grained definition of overlapping structures on plain content.

### 1.3 Round-table Session

The workshop also includes a round-table session. The outcome of the discussion is not reported in these proceedings, since the workshop is not yet held. Some topics that will be discussed are: distributed editing issues (following-up the keynote), human-interpretation of changes and quality of deltas, and identification of editing patterns in other domains like law-making and humanities. More updated details will be published on the workshop web page [http://diff.cs.unibo.it/dchanges2013/roundtable/](http://diff.cs.unibo.it/dchanges2013/roundtable/).

Suggestions from the audience will be encouraged throughout the workshop. Our goal is to foster research collaboration and to also identify topics for a second edition. We hope to have more and more interesting editions of DChanges in the future and to gather a lively community of researchers around these themes.

### 1.4 Acknowledgements

In conclusion, we would like to thank all people who had expressed interest in DChanges and the organizers of DocEng – in the first place Simone Marinai and Kim Marriott – for giving us the possibility of organizing it and for supporting us continuously.

Our thanks go to the committee members, for their hard work in circulating the call for papers and reviewing papers (perfectly on time!).

A special thank goes to Ethan Munson, for his illuminating keynote.

We wish you a very good read,

The DChanges chairs
2. COMMITTEE
The workshop has been organized by four people, from two research groups. They have been helped by a committee of experts from all around the world.

Organizers
- Gioele Barabucci is a research fellow at Università di Bologna. He recently received his PhD with a thesis on diff algorithms and delta models.
- Uwe M. Borghoff is a full professor of Computer Science at Universität der Bundeswehr München. With his research group, he published various papers on algorithms for comparing textual documents and on related topics.
- Angelo Di Iorio is an assistant professor at Università di Bologna. He worked on various systems for document versioning and publishing, and collaborative editing.
- Sonja Maier is a Postdoc at Universität der Bundeswehr München. In her research, she focuses on tool creation and tool integration for (visual) domain-specific languages, and is interested in tracking the evolution of text and diagrams.

Committee members
- Serge Autexier, DFKI Bremen
- Stéphane Ducasse, INRIA Lille Nord Europe research center
- Boris Konev, University of Liverpool
- John Lumley
- Pascal Molli, Université de Nantes - LINA
- Sebastian Rönnau
- Wolfgang Stürzlinger, York University
- Yannis Tzitzikas, University of Crete and FORTH-ICS
- Fabio Vitali, Università di Bologna
- Jean-Yves Vion-Dury, Xerox Research Centre Europe

Additional reviewers:
- Emmanuel Desmontils
- Christina Lantzaki
- Brice Nédelec
Table of Contents

Merging Uncertain Multi-Version XML Documents
  Mouhamadou Lamine Ba, Talel Abdessalem and Pierre Senellart

Identifying Change Patterns in Software History
  Jason Dagit and Mathew Sottile

The Concept Difference for EL-Terminologies using Hypergraphs
  Andreas Ecke, Michel Ludwig and Dirk Walther

An Algorithm for Transforming XPath Expressions According to Schema Evolution
  Kazuma Hasegawa, Kosetsu Ikeda and Nobutaka Suzuki

Concurrency Effects Over Variable-size Identifiers in Distributed Collaborative Editing
  Brice Nédélec, Pascal Molli, Achour Mostefaoui and Emmanuel Desmontils

Tracking changes Through EARMARK: a Theoretical Perspective and an Implementation
  Silvio Peroni, Francesco Poggi and Fabio Vitali

Staged Evolution With Quality Gates for Model Libraries
  Alexander Roth, Andreas Ganser, Horst Lichter and Bernhard Rumpe