eMath 3.0: Building Blocks for a Social and Semantic Web for Online Mathematics & eLearning

Catalin David*, Deyan Ginev*, Michael Kohlhase*, Joseph Corneli†

October 27, 2010

Abstract

In this paper we present recent developments in content markup for mathematics, and a corresponding software stack that functions as an enabling technology for a social and semantic web for the STEM disciplines. We show the potential of this technology in two eMath 3.0 applications: PlanetMathRedux, a re-implementation of the mathematical encyclopedia PlanetMath.org, and PantaRheiRedux, a community reader for course materials. These applications indicate both present and potential uses for this software as a basis for eLearning applications in Science, Technology, Engineering and Mathematics through the addition of suitable pedagogies.

1 Introduction

The Internet has revolutionized our access to information: much of what we need to know is available online, and can be found via search engines. In the last decade, this trend has been accelerated by the advent of the social and semantic web.

The *social web* (also called Web 2.0) fills various online commons with user-generated content and peer-based interactions. The social web has greatly extended the material available on the Internet; for instance, Wikipedia has accumulated more than 16 million articles in almost all the world's languages over the last 10 years.

The *Semantic Web* (by convention, referred to with capital letters) adds formal descriptions to web resources, so that the information they contain becomes machine-understandable. Semantic information retrieval can be used to combine different information sources to obtain facts and functionalities that are entailed, but nowhere explicitly represented, in the original sources themselves.

*Computer Science, Jacobs University, Germany; `http://kwarc.info/people/`

†Knowledge Media Institute, The Open University, UK; `http://metameso.org/~joe`

intro.tex 300 2010-10-26 11:44:03Z jcorneli

The emerging combination of social and semantic web techniques has sometimes been called Web 3.0. Applications like semantic wikis (see e.g. [Lan+10] and earlier workshops) and DBPedia [Dbp] (a semantic query engine based on the content of Wikipedia) are showing the considerable promise of this mixed approach.

Perhaps surprisingly, the Semantic Web has not taken off for the Science, Technology, Engineering & Mathematics (STEM) disciplines, even though STEM documents are ostensibly more rigorous, have more overt formal structure, and should therefore be amenable to formal, programmatic, treatment via Semantic Web technologies. Three possible impediments seem to be likely explanations for the slow uptake:

1. Both the STEM and Semantic Web user communities are small compared to the social web as a whole.
2. There is significant inherent difficulty and complexity in STEM subject matter.
3. Special representation languages are used by both practitioners (e.g. mathematical and chemical formulae) and technologists (e.g. XML and ontologies).

We believe that the first two factors are in fact offset by the dedication of STEM practitioners, and we claim that the third one – which is the main problem – has been solved, for the mathematical sciences, by recent technological developments. We will survey these broadly in the next section, and describe in detail a specific set of related technologies developed at Jacobs University Bremen over the last decade in section 3. In section 4 we will then show how these building blocks can be assembled into an *eMath 3.0 system* system for developing and exploiting social and semantic mathematics on the Web. Section 5 concludes the paper, with an emphasis on planned future work.

2 Mathematics on the Web: MathML, OpenMath and RDFa

Technologies that provide support for Mathematics on the Web have been increasing in prevalence over the last decade, and have reached a state, where they will become universally usable out of the box. This has been especially driven forward by the continued development of MathML and OpenMath as representation formats for mathematical formulae, and by the integration of Semantic Web technologies into widely-accepted web standards.

2.1 Mathematical Formulae on the Web: State of the Art

For mathematical formulae, the relevant developments are:

M1: MathML3, the upcoming W3C recommendation [Aus+10] extends MathML with line-breaking support (important e.g. for mobile applications), markup for elementary mathematics (for high school eLearning), and completely reworks the integration of content markup into MathML. Furthermore, the improved integration of MathML into host languages and environments gives a more solid basis for Math-enabled web applications.

M2: MathML is in the main browsers: natively in Firefox, via the MathPlayer plugin in Internet Explorer, via CSS in Opera, and native MathML support in Webkit has recently been announced, so we can expect it in the Webkit-based browsers (Chrome, Safari and Konqueror) in the near future. Additionally, Math support for all browsers can be enabled via MathJax [Mat], a JavaScript-based solution for displaying MathML or LaTeX in the browser.

M3: MathML is part of the upcoming HTML5 standard [Hic10], which is expected to drive the application-centered Web in the future: MathML elements (both content and presentation) are placed in the DOM, in the MathML namespace.

M4: TEX/LATEX documents can be transformed to XHTML+MathML automatically by systems like Tralics [Tra] or LaTeXML [Mil]; see [Sta+09] for an overview.

M5: More and more word processors and office applications include support for mathematical formulae (most prominently MS Word since Office 2007), and allow an export to MathML [Car07].

M6: As a consequence, very large collections of mathematical documents have become available online marked up with standard web formatting, most prominently, Zentralblatt Math [Zbl] and a web-enabled version of the Cornell ePrint arXiv [Sta+10; Arx].

2.2 Semantic Web Technologies for Mathematical Documents

Semantic web technologies are able to describe relationships within and between mathematical documents above the formula level.

S1: Text fragments can be classified by their mathematical role and related to each other via RDF triples [MM04], which can be efficiently queried with RDF query systems like SPARQL [PS08].

S2: RDF triples can be embedded into XHTML documents via RDFa; the upcoming version RDFa 1.1 [Adi+10] extends this to arbitrary XML languages.

S3: HTML5 embeds similar functionality in web pages via "microformats"; see [Ten09] for an analysis.

3 The KWARC Technology Stack

In the last five years, the KWARC research group at Jacobs University has developed a stack of technologies for content-oriented representation of mathematics in Web contexts together with tools for processing and interacting with these representations in sophisticated ways (see Figure 1). They have arisen as generalizations of system components in the course of developing systems like the SWiM semantic Wiki for mathematics [Lan08], Panta Rhei, a semantic community reader [MK08], and the ActiveMath eLearning System [Mel+03a]. We will review them before we show in the next section how they can be re-combined in novel applications.

3.1 Representation: OMDoc & sTeX

OMDoc is an XML-based content-oriented representation format for scientific documents, which is now used in a large set of projects in Automated Theorem Proving [Mül06], eLearning [Mel+03b; KK08b; Koh07], eScience [HKS06], Document Retrieval [KŞ06], User Assistance [KK08a;

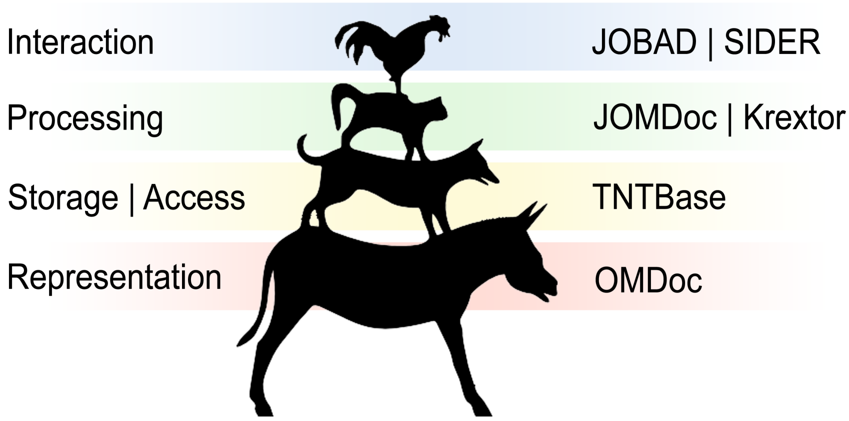

Figure 1: The KWARC Software Stack

KK09], and in Formal Digital Libraries [Url]. The OMDoc format builds on existing semantic representation formats for mathematical formulae (OpenMath [Bus+04] objects and Content MathML [Aus+03] representations), and extends them by an infrastructure for context and domain models from Formal Methods for important structural properties as well as for semi-formal content. Work on the OMDoc format shows that many added-value services in Knowledge Management do not need tedious formalization, but can be based at the structural/semantic level.

We have developed two approaches to ease the non-trivial task of authoring OMDoc documents. The first is an OMDoc-based semantic Wiki, which integrates server-based editing with user-adaptive and context-based presentation [LK08; Lan07]. The second approach we call an "invasive technology" [Koh05], since we build OMDoc-aware editing facilities into existing editing frame-works to make the most of existing functionalities and get around the learning curve involved with a new editor. We have evaluated this approach for MS PowerPoint [KK04; Koh05] and sTeX [Koh08; KKL10], a semantic variant of LaTeX; in both formats we can embed OMDoc markup as well as generate pure OMDoc as an export format.

3.2 Storage: TNTBase

Large scale collaborative authoring of mathematical documents requires versioned storage. On the language end, OMDoc supports this by making all identifiers URIs so that OMDoc documents can be distributed among authors and networks and reference each other. On the storage end, we use the TNTBase system [ZK09], a versioned XML-database with a client-server architecture. It integrates Berkeley DB XML with a Subversion server [Apa]. DB XML stores HEAD revisions of XML files; non-XML content like PDF, images or LaTeX source files, as well as differences between revisions, directory entry lists and other repository information are retained in a usual SVN back-end storage (Berkeley DB, in our case). Keeping XML documents in DB XML allows access to files via not only an SVN client, but also through the DB XML API, which supports efficient querying of XML content via XQuery, as well as (versioned) modification of the XML content via XQuery Update.

Versioning and distribution can also be realized with a plain SVN server, but for mathematics, it is additionally important that the storage backend is aware of at least some aspects of the mathematical semantics. For example, in large-scale authoring processes, an important requirement is to guarantee consistency, i.e., it should be possible to reject commits of invalid documents. TNTBase can support format-specific validation of language-specific constraints and invariants that cannot be expressed in the XML schema languages.

For document management, TNTBase provides Virtual Documents. The author writes a Virtual Document skeleton document that embeds XQueries which are materialized by TNTBase when the content is served. This is particularly useful in eLearning applications, since Virtual Documents make it easy to generate aggregated- and user-adapted documents, as well as other kinds of document variants. TNTBase permits committing back changed Virtual Documents; the changes are distributed to the original files the Virtual Document was assembled from; see [ZK10] for details.

3.3 Processing: JOMDoc

JOMDoc [Jom] is a Java API for OMDoc documents, which facilitates parsing OMDoc XML documents into an internal Java data structure, allowing for convenient manipulation, and ultimately, serialization of this internal representation back to XML. JOMDoc has been integrated into TNTBase via the latter's plugin architecture for document format-specific customizations [ZKR10]. This makes made TNTBase OMDoc-aware so that data-intensive JOMDoc algorithms can be executed within the database, alleviating the need of sending the contents over the network for processing. Additionally, JOMDoc is used as a presentation framework for OMDoc. With its notation service [KMR08], it allows for context-sensitive rendering of XML documents containing mathematical formulae in content markup (Content MathML or OpenMath) into Presentation MathML. Optionally this can be presented as parallel markup, i.e., interlinked with the original content markup. Transformation of OMDoc documents to XHTML is supported by bundled XSLT stylesheets. In particular, this presentation service can be used to serve OMDoc documents in a human-readable presentation.

3.4 Interaction: JOBAD

Our JOBAD architecture embeds interactive mathematical services into XHTML+MathML documents. JOBAD is a modular JavaScript framework for interactive services such as term folding or definition lookup.

Our vision of a document is that it should be something the user can adapt according to his or her preferences and interests while reading it. This goes beyond customizing the display of the rendered document in the browser, to include, for example, changing notations (which requires re-rendering at least portions of the document), or retrieving additional information from services on the web to enhance the document with annotations. Consider a student reading lecture notes: whenever an unfamiliar mathematical symbol occurs in some formula, JOBAD enables the look up its definition without opening another document, but adds an explanation right into the current reading context. Converting between physical units (e.g. imperial and SI) can also be effected automatically, in-place and on the fly.

4 The Planetary System: Assembling Applications

In August 2010, the authors started using the building blocks described in the last section to build a new front-end system for eMath 3.0 applications: the Planetary System. The starting point of this development project was our aim to make PlanetMath.org [Plab], one of the original eMath2.0 systems, more semantic, by integrating it with the KWARC technologies described in the last section.

PlanetMath.org is a relatively well-known online community devoted to mathematics. At present, its central feature is a mathematics encyclopedia with around 9K entries, which has been built and peer reviewed through effort of several hundred of volunteers since the site went online in 2001. PlanetMath also includes several general-purpose discussion forums which have received around 15K posts to date. Its most popular forums, containing about half of these posts, are devoted to Q&A about mathematics at the university, post-graduate, and research levels. Notably, each encyclopedia article also has its own attached discussion forum.

While the PlanetMath concept, community and vision remain alive and active, the Noösphere web application that underlies the site is showing its age. In particular, PlanetMath does not currently make significant use of any of the state of the art technologies described in section 2, other than using LaTeXas an input syntax.

To bring PlanetMath up to date, and simultaneously make its software easier to extend in the future, we decided to recreate the relevant functionality of Noösphere by integrating contemporary mathematical communication features into the existing open source web platform, Vanilla Forums. Vanilla offers a general-purpose online infrastructure, including user management and discussion forums, together with a plugin system that makes it relatively easy to adapt different components to a given special-purpose use.

Soon after we began to carry out this plan, we realized that Vanilla's plugin architecture would allow us to build a system that could be configured into multiple different eMath3.0 applications. This gave rise of the concept of the Planetary System. We are currently exploring its possibilities in two main applications: PlanetMathRedux [Plaa] (a new PlanetMath.org based on the Planetary System), and PantaRheiRedux [Pan].

4.1 The Software Base: Vanilla Forums

Vanilla Forums is an open-source, standards-compliant discussion forum platform with a very large user base (around 390K communities). Using this "off the shelf" forum software gave us a lot "for free", including user and permission management, and an extensive set of existing plugins to enhance the content and display. Most importantly, Vanilla's plugin architecture nicely complements the software stack that we already have which can be integrated into the system with plugins and applications, as described in the following subsections. The only significant problem we encountered with Vanilla was that while it is advertised to be XHTML-compatible, it seems to be only served with the `text/html` media type (as tag soup) in practice. Thus we had to correct numerous XHML validity errors when changing to the `application/xhtml+xml` media type, a prerequisite for embedding MathML into Vanilla.

4.2 Adding Math to Vanilla

Vanilla Forums does not provide authoring tools for mathematical formulae, hence we had to extend the forum functionality by adding a LaTeX-editor plug-in via the LaTeXML LaTeX-to-XHTML+MathML converter [Mil]. There were many alternatives to this choice (see [Sta+09]), but none of them scale to the full expressivity of LATEX, which is the input format used by PlanetMath.org. We were able to make use of our extensions to the batch-mode LaTeXML converter that turned it into a daemon web-service, to decreasing startup latency. This allows high-throughput conversion of arbitrary LaTeX fragments – ranging from simple expressions, to entire chapters or books.

We developed a Vanilla plugin that integrates the LaTeXML daemon with forum posts in the following manner: if a post is categorized as LaTeX then instead of letting Vanilla embed it into HTML directly, we first have it transformed by the resident LaTeXML daemon and integrate the XHTML+MathML result instead. Due to the swiftness of the conversion, it is possible to create "on-the-fly" editors like [Sta], in which the author can see the produced content as they type.

Using the LaTeX syntax significantly increases the expressivity of the authoring process when compared with Rich Text and Wiki syntax. LaTeX supports mathematical formulas, creating graphics and charts, easy fine-tuning of tables, complex page partitioning, custom commands, preambles, abbreviations, invasive preloading of semantics and more. Probably even more importantly, LaTeX continues to be deeply ingrained in contemporary mathematical publication and communication processes.

4.3 Encyclopedia Articles

The probably largest difference between standard Vanilla forums and PlanetMath.org is that the main content of PlanetMath is organized in and around "encyclopedia articles". Each of these is a versioned, LaTeX-encoded description of a particular mathematical object or topic, with an attached discussion, as we mentioned above. We added functionality for "encyclopedia articles" to Vanilla forums via a new "application" (Vanilla's term for a complex style of plugin that hooks into the same core, but adds new features, instead of merely changing old the way old features work). In contrast to unversioned forum posts, articles are stored in an associated Subversion repository, whose versioning functionalities are then exposed in the new interface. Vanilla's own database is used as a cache for efficient web publishing.

The default mode of navigating a content collection in Vanilla (based on listing articles by name) is unsuitable for the 9K articles in PlanetMath.org. Access via the Math Subject Classification (MSC [Msc]), as in the original version of PlanetMath.org, can be realized via the Vanilla Metadata scheme; we can even make these metadata properties versioned by encoding them into Subversion properties.

We have also added a new navigation method, called a "virtual bookshelf" which contains "books" created by users (here, authors, aggregators, or readers) according to their interest. The hierarchical document structure of a "book'' is encoded by allowing inclusion primitives (a variant of "`\input`" from LaTeX) in articles, which then become sectioning nodes (e.g. chapters, sections), which can be mixed with transitional text. A narrative structure can be represented

planetary.tex 301 2010-10-27 06:04:43Z kohlhase

simply as a "next" relation that links such articles. In a course setting, our Articles application can be used to highlight important books together with the forum interface (e.g. these may be portions of the course notes or other required readings).

With the JOBAD system we can already add a very nice feature to PlanetMathRedux: fine-grained forum posts. As JOBAD has access to the document object model (including that of the mathematical formulae in MathML), we can use it to attach forum posts to arbitrary document sub-structures, so that a user can e.g. ask a question about a particular definition, make a line-level proposal for an alternative proof, or offer a correction to a subformula. This fine-grained embedding of the forum into the documents allows Planetary System to be used as a "community reader", supporting discussion, document inspection, and refereeing.

The Articles application, together with the math editing plugin described above (and other some off-the-shelf Vanilla plugins) are enough to replicate most of the relevant parts of the Noösphere functionality, and they form the core of the PlanetMathRedux system. Note that PlanetMathRedux does not yet make any essential use of the semantics implicit in the articles.

4.4 Semantic Interaction

We consider mere "reading" of an article to be a deeply eMath2.0 activity and note that mathematics consumers will want to engage with the content of the article in more interactive ways, for more efficient and enjoyable learning experience. In mathematics, a dialogue with the expert is often a better way to settle an uncertain matter than reading a book, or searching through the library or the internet. Reading without thinking, computing and proving intermediate results is generally not the way to learn mathematics.

A prerequisite for the Planetary System to offer solid support for mathematical interaction is for it to have the content markup for articles available (and maybe eventually the corresponding content markup for forum posts as well). To support this, we have extended the articles application to handle sTeX. In this extension, we use TNTBase instead of Subversion and transform sTeX articles to OMDoc, which is then managed in TNTBase.

Our goal is to create "active documents" which adapt to the environment and can richly support user interaction. Some interactions only depend on information that is only related to the document at hand, e.g. the elision of formulae parts like brackets, types or inferable arguments. These can be implemented in the browser: JOMDoc exports the respective semantic information from the OMDoc representation into the XHTML+MathML documents as RDFa annotations [Adi+08], where they can be picked up via the document-embedded JOBAD services. For interactions that depend on larger amounts of data from outside the document itself, JOBAD implements call-backs to TNTBase. For instance, this is used for definition lookup, and for the generation of a concept graph of an exercise problem; see [Dav+10] for details on other services. In some cases, we need mixed computation models for semantic services, e.g. where some information is only present on the client (personal information which cannot be transferred to the server due to privacy concerns).

concl.tex 300 2010-10-26 11:44:03Z jcorneli

5 Conclusion & Future work

We have presented a set of content-based technologies (the KWARC stack) for building semantically and socially enabled math-aware online applications (eMath3.0 applications). We have shown how this can be done using our new Planetary System as an example. Originally planned as a re-implementation of the software underlying PlanetMath.org, it quickly grew more general and is now also used as the basis of PantaRheiRedux, a semantic eLearning platform in active use at Jacobs University.

We do not view the Planetary System (or the KWARC stack for that matter) as an eLearning system in its own right, but rather as an enabling technology for eLearning in the STEM disciplines. An eLearning platform additionally requires pedagogy. In PantaRheiRedux, the pedagogical aspects are minimal: the system gives access to the course materials, allows students to discuss them, and gives access to semantic services. It would be possible to add more pedagogy, e.g. formative assessments, learner modeling, adaptiveness and instant feedback based on the learner models.

We are planning additional applications of the KWARC stack and the Planetary System: in the arXMLiv project [Sta+10; Arx], we have transformed a large corpus of scientific papers to XHTML+MathML. We are currently working towards extracting a subset of the OMDoc format automatically from these, which would allow us to use the Planetary System as a lightweight community reading platform for the arXiv, where readers can discuss scientific questions, annotate semantic relations in the papers, and interact with the content of these papers more fully than in the current PDF-based system.

Finally, we are working on a version of the Planetary System for Formal Methods: we want to use the Planetary System as a front-end for a knowledge base of modular logic representations and logic transformations represented in an upcoming version of OMDoc; see [KMR] for details. The formal documents can support very powerful semantic services like borrowing automated theorem provers or the automated translation between formalizations in different logics. The formal setting provides a compelling angle on the useful range of this system, which we have seen includes support for both lightweight social to very intense semantic interaction.

References

[Adi+08] Ben Adida et al. *RDFa in XHTML: Syntax and Processing*. W3C Recommendation. World Wide Web Consortium (W3C), Oct. 2008. URL: http://www.w3.org/TR/2008/REC-rdfa-syntax-20081014/.

[Adi+10] Ben Adida et al. *RDFa Core 1.1. Syntax and processing rules for embedding RDF through attributes*. W3C Working Draft. World Wide Web Consortium (W3C), Aug. 3, 2010. URL: http://www.w3.org/TR/2010/WD-rdfa-core-20100803/.

[Apa] Apache Software Foundation. *Apache Subversion*. URL: http://subversion.apache.org/ (visited on 10/22/2009).

[Arx] *arXMLiv Build System*. URL: http://arxivdemo.mathweb.org (visited on 09/27/2010).

[Aus+03] Ron Ausbrooks et al. *Mathematical Markup Language (MathML) Version 2.0 (second edition)*. W3C Recommendation. World Wide Web Consortium (W3C), 2003. URL: http://www.w3.org/TR/MathML2.

[Aus+10] Ron Ausbrooks et al. *Mathematical Markup Language (MathML) Version 3.0*. W3C Proposed Recommendation of 10. August 2010. World Wide Web Consortium (W3C), 2010. URL: http://www.w3.org/TR/MathML3.

[Bus+04] Stephen Buswell et al. *The Open Math Standard, Version 2.0*. Tech. rep. The OpenMath Society, 2004. URL: http://www.openmath.org/standard/om20.

[Car07] David Carlisle. *XHTML and MathML from Office 2007*. Apr. 10, 2007. URL: http://dpcarlisle.blogspot.com/2007/04/xhtml-and-mathml-from-office-20007.html (visited on 09/29/2010).

[Dav+10] Catalin David et al. "Publishing Math Lecture Notes as Linked Data". In: *The Semantic Web: Research and Applications (Part II)*. 7th Extended Semantic Web Conference (ESWC) (Heraklion, Crete, Greece, May 30–June 3, 2010). Ed. by Lora Aroyo et al. Lecture Notes in Computer Science 6089. Springer Verlag, 2010, pp. 370–375. arXiv:1004.3390v1 [cs.DL].

[Dbp] *DBpedia*. URL: http://dbpedia.org (visited on 01/23/2010).

[Hic10] Ian Hickson. *HTML5 (including next generation additions still in development)*. Draft Standard. Web Hypertext Application Technology Working Group (WHATWG), 2010. URL: http://whatwg.org/html5 (visited on 02/02/2010).

[HKS06] Eberhard Hilf, Michael Kohlhase, and Heinrich Stamerjohanns. "Capturing the Content of Physics: Systems, Observables, and Experiments". In: *Mathematical Knowledge Management (MKM)*. Ed. by Jon Borwein and William M. Farmer. LNAI 4108. Springer Verlag, 2006, pp. 165–178. URL: http://kwarc.info/kohlhase/papers/mkm06physml.pdf.

[Jom] *JOMDoc Project — Java Library for OMDoc documents*. URL: http://jomdoc.omdoc.org (visited on 10/22/2009).

[KK04] Andrea Kohlhase and Michael Kohlhase. "CPoint: Dissolving the Author's Dilemma". In: *Mathematical Knowledge Management, MKM'04*. Ed. by Andrea Asperti, Grzegorz Bancerek, and Andrej Trybulec. LNAI 3119. Springer Verlag, 2004, pp. 175–189. URL: http://kwarc.info/kohlhase/papers/mkm04.pdf.

[KK08a] Andrea Kohlhase and Michael Kohlhase. "Compensating the Semantic Bias of Spreadsheets". In: *Wissens- und Erfahrungsmanagement LWA (Lernen, Wissensentdeckung und Adaptivität) Conference Proceedings*. Ed. by Joachim Baumeister and Martin Atzmüller. Vol. 448. Oct. 2008. URL: http://omdoc.org/pubs/kohkoh-lwa08.pdf.

[KK08b] Andrea Kohlhase and Michael Kohlhase. “Semantic Knowledge Management for Education”. In: *Proceedings of the IEEE; Special Issue on Educational Technology* 96.6 (June 2008), pp. 970–989. URL: `http://kwarc.info/kohlhase/papers/semkm4ed.pdf`.

[KK09] Andrea Kohlhase and Michael Kohlhase. “Spreadsheet Interaction with Frames: Exploring a Mathematical Practice”. In: *MKM/Calculemus Proceedings*. Ed. by Jacques Carette et al. LNAI 5625. Springer Verlag, July 2009, pp. 341–256. URL: `http://kwarc.info/kohlhase/papers/mkm09-framing.pdf`.

[KKL10] Andrea Kohlhase, Michael Kohlhase, and Christoph Lange. “sTeX – A System for Flexible Formalization of Linked Data”. In: *Proceedings of the 6th International Conference on Semantic Systems (I-Semantics) and the 5th International Conference on Pragmatic Web*. Ed. by Adrian Paschke et al. ACM, 2010. ISBN: 978-1-4503-0014-8. arXiv:`1006.4474v1 [cs.SE]`.

[KMR] Michael Kohlhase, Till Mossakowski, and Florian Rabe. *LATIN: Logic Atlas and Integrator*. URL: `http://latin.omdoc.org` (visited on 09/15/2010).

[KMR08] Michael Kohlhase, Christine Müller, and Florian Rabe. “Notations for Living Mathematical Documents”. In: *Intelligent Computer Mathematics*. 9th International Conference, AISC, 15th Symposium, Calculemus, 7th International Conference MKM (Birmingham, UK, July 28–Aug. 1, 2008). Ed. by Serge Autexier et al. LNAI 5144. Springer Verlag, 2008, pp. 504–519. URL: `http://omdoc.org/pubs/mkm08-notations.pdf`.

[Koh05] Andrea Kohlhase. “Overcoming Proprietary Hurdles: CPoint as Invasive Editor”. In: *Open Source for Education in Europe: Research and Practise*. Ed. by Fred de Vries et al. Proceedings at `http://hdl.handle.net/1820/483`. Open Universiteit Nederland. Heerlen, The Netherlands: Open Universiteit Nederland, Nov. 2005, pp. 51–56. URL: `http://hdl.handle.net/1820/483`.

[Koh07] Andrea Kohlhase. “Semantic PowerPoint: Content and Semantic Technology for Educational Added-Value Services in MS PowerPoint”. In: *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications 2007 (ED-MEDIA’07)*. Ed. by Craig Montgomerie and Jane Seale. AACE, June 2007, pp. 3576–3583. URL: `http://go.editlib.org/p/25890`.

[Koh08] Michael Kohlhase. “Using LaTeX as a Semantic Markup Format”. In: *Mathematics in Computer Science* 2.2 (2008), pp. 279–304. URL: `https://svn.kwarc.info/repos/stex/doc/mcs08/stex.pdf`.

[KŞ06] Michael Kohlhase and Ioan Şucan. “A Search Engine for Mathematical Formulae”. In: *Proceedings of Artificial Intelligence and Symbolic Computation, AISC’2006*. Ed. by Tetsuo Ida, Jacques Calmet, and Dongming Wang. LNAI 4120. Springer Verlag, 2006, pp. 241–253. URL: `http://kwarc.info/kohlhase/papers/aisc06.pdf`.

[Lan07] Christoph Lange. “Towards Scientific Collaboration in a Semantic Wiki”. In: *Bridging the Gap between Semantic Web and Web 2.0 (SemNet)*. Ed. by Andreas Hotho and Bettina Hoser. June 2007.

[Lan08] Christoph Lange. “SWiM – A semantic wiki for mathematical knowledge management”. In: *The Semantic Web: Research and Applications*. 5$^{\text{th}}$ European Semantic Web Conference (ESWC) (Tenerife, Spain, June 1–5, 2008). Ed. by Sean Bechhofer et al. Lecture Notes in Computer Science 5021. Springer Verlag, 2008, pp. 832–837. arXiv:`1003.5196v1 [cs.DL]`.

[Lan+10] Christoph Lange et al., eds. *Proceedings of the 5th Workshop on Semantic Wikis, Extended Semantic Web Conference*. (Hersonissos, Greece, May 31, 2010). CEUR Workshop Proceedings 632. 2010.

[LK08] Christoph Lange and Michael Kohlhase. “A Semantic Wiki for Mathematical Knowledge Management”. In: *Emerging Technologies for Semantic Work Environments: Techniques, Methods, and Applications*. Ed. by Jörg Rech, Björn Decker, and Eric Ras. IGI Global, Apr. 2008, pp. 47–68. URL: `http://www.igi-global.com/Bookstore/Chapter.aspx?TitleId=10143`.

[Mat] *MathJax: Beautiful Math in all Browsers*. URL: `http://mathjax.com` (visited on 09/27/2010).

[Mel+03a] E. Melis et al. “Knowledge Representation and Management in ACTIVEMATH”. In: *International Journal on Artificial Intelligence and Mathematics, Special Issue on Management of Mathematical Knowledge* 38.1–3 (2003), pp. 47–64.

[Mel+03b] Erica Melis et al. “Knowledge Representation and Management in ActiveMath”. In: *Annals of Mathematics and Artificial Intelligence* 38 (2003). see `http://www.activemath.org`, pp. 47–64.

[Mil] Bruce Miller. *LaTeXML: A LaTeX to XML Converter*. URL: `http://dlmf.nist.gov/LaTeXML/` (visited on 05/08/2010).

[MK08] Christine Müller and Michael Kohlhase. “Towards A Community of Practice Toolkit Based On Semantically Marked Up Artifacts”. In: *Proceedings of the 1st World Summit of the Knowledge Society: Emerging Technologies and Information Systems for the Knowledge Society*. Ed. by M. D. Lytras et al. LNAI 5288. Springer-Verlag Berlin Heidelberg, 2008, pp. 41–50.

[MM04] Frank Manola and Eric Miller. *RDF Primer*. W3C Recommendation. World Wide Web Consortium (W3C), Feb. 2004. URL: `http://www.w3.org/TR/2004/REC-rdf-primer-20040210/`.

[Msc] *Mathematics Subject Classification MSC2010*. 2010. URL: `http://msc2010.org` (visited on 11/16/2009).

[Mül06] Normen Müller. "OMDoc as a Data Format for VeriFun". In: OMDOC – *An open markup format for mathematical documents [Version 1.2]*. LNAI 4180. Springer Verlag, Aug. 2006. Chap. 26.20, pp. 329–332. URL: http://omdoc.org/pubs/omdoc1.2.pdf.

[Pan] *PlanetMath Redux.org – Math for the people, by the people*. URL: http://trac.mathweb.org/planetary/wiki/PantaRheiRedux (visited on 09/30/2010).

[Plaa] *PlanetMath Redux.org – Math for the people, by the people*. URL: http://planetmath.mathweb.org (visited on 09/30/2010).

[Plab] *PlanetMath.org – Math for the people, by the people*. URL: http://planetmath.org (visited on 07/14/2010).

[PS08] Eric Prud'hommeaux and Andy Seaborne. *SPARQL Query Language for RDF*. W3C Recommendation. World Wide Web Consortium (W3C), Jan. 15, 2008. URL: http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/.

[Sta] Heinrich Stamerjohanns. *An in-browser editor for LaTeX fragments*. URL: http://tex2xml.kwarc.info/test/edit.php (visited on 10/10/2010).

[Sta+09] Heinrich Stamerjohanns et al. "MathML-aware article conversion from LaTeX, A comparison study". In: *Towards Digital Mathematics Library, DML 2009 workshop*. Ed. by Petr Sojka. Masaryk University, Brno, 2009. URL: http://kwarc.info/kohlhase/submit/dml09.pdf.

[Sta+10] Heinrich Stamerjohanns et al. "Transforming large collections of scientific publications to XML". In: *Mathematics in Computer Science* 3.3 (2010): *Special Issue on Authoring, Digitalization and Management of Mathematical Knowledge*. Ed. by Serge Autexier, Petr Sojka, and Masakazu Suzuki, pp. 299–307. URL: http://kwarc.info/kohlhase/papers/mcs09.pdf.

[Ten09] Jeni Tennison. *HTML5/RDFa Arguments*. Aug. 21, 2009. URL: http://www.jenitennison.com/blog/node/124 (visited on 02/02/2010).

[Tra] *Tralics: a LaTeX to XML translator*. URL: http://www-sop.inria.fr/miaou/tralics/ (visited on 09/27/2010).

[Url] *Logosphere: a Formal Digital Library*. web page at http://www.logosphere.org. 2006. URL: http://www.logosphere.org/.

[Zbl] *Zentralblatt MATH*. URL: http://www.zentralblatt-math.org (visited on 01/08/2010).

[ZK09] Vyacheslav Zholudev and Michael Kohlhase. "TNTBase: a Versioned Storage for XML". In: *Proceedings of Balisage: The Markup Conference*. Vol. 3. Balisage Series on Markup Technologies. Mulberry Technologies, Inc., 2009. DOI: 10.4242/BalisageVol3.Zholudev01. URL: http://www.balisage.net/Proceedings/vol3/html/Zholudev01/BalisageVol3-Zholudev01.html.

[ZK10] Vyacheslav Zholudev and Michael Kohlhase. “Scripting Documents with XQuery: Virtual Documents in TNTBase”. In: *Proceedings of Balisage: The Markup Conference*. Vol. 5. Balisage Series on Markup Technologies. Mulberry Technologies, Inc., 2010. DOI: 10.4242/BalisageVol5.Zholudev01. URL: http://www.balisage.net/Proceedings/vol5/html/Zholudev01/BalisageVol5-Zholudev01.html.

[ZKR10] Vyacheslav Zholudev, Michael Kohlhase, and Florian Rabe. *A [insert XML Format] Database for [insert cool application] (extended version)*. 2010. URL: http://kwarc.info/vzholudev/pubs/XMLPrague_long.pdf (visited on 12/21/2009).