

# System Description: MathHub.info

Mihnea Iancu, Constantin Jucovschi, Michael Kohlhase, Tom Wiesing

Computer Science, Jacobs University Bremen, <http://kwarc.info>

**Abstract.** We present the `MathHub.info` system, a development environment for active mathematical documents and an archive for flexiformal mathematics. It offers a rich interface for reading, writing, and interacting with mathematical documents and knowledge. The core of the `MathHub.info` system is an archive for flexiformal mathematical documents and libraries in the `OMDoc/MMT` format. Content can be authored or archived in the source format of the respective system, is versioned in GIT repositories, and transformed into `OMDoc/MMT` for machine-support and further into HTML5 for reading and interaction.

## 1 Introduction

As the field of Mathematical Knowledge Management (MKM) and Digital Mathematical Libraries (DML) mature, we need to shift attention from experimenting with semantic services on small and practice data sets to supporting management of large corpora of mathematical knowledge and documents. In the past, MKM and DML have latched onto existing corpora/libraries ranging from digitized mathematical articles (e.g. EuDML [EUD]) over semi-structured representations generated from  $\text{\LaTeX}$  sources [Sta+10] to fully formal theorem prover libraries, e.g. the Mizar Mathematical Library [MizLib]. But so far management support for and semantic services deployed on such libraries have been essentially insular, existing cross-library methods remain experiments, and have not been implemented into usable systems.

This insularity also applies to the work in the KWARC group. For instance, we have *i*) designed a cross-library representation language: `OMDoc` (Open Mathematical Documents), [Koh06], *ii*) developed a meta-logical framework `MMT` (A Module System for Mathematical Theories) [RK13; Cod+11]) that allows to represent the logical languages underlying the theorem prover libraries, and relate them to each other by logic morphisms, *iii*) built libraries of formalized mathematics in `OMDoc/MMT` either by manually creating content (e.g. `LATIN` [Cod+11]) or by implementing transformations from existing theorem prover libraries (e.g. Mizar Mathematical Library [Ian+13]), and *iv*) have used `OMDoc` as the basis for active mathematical documents in the Planetary system [Koh12]. But we have not integrated all of these or made them available to mathematicians in one comprehensive environment.

To change this situation – and to provide a realistic base for our own cross-library research and development efforts – we have started work to realize a universal archiving solution for formal and informal mathematical corpora/libraries.

We present the MathHub.info system and its design goals in this paper. MathHub.info must satisfy two conflicting goals: On the one hand, it must be so generic that it is open to all logics and implementations; on the other hand, it must be aware of the semantics of the formalized content so that it can offer meaningful services. These services must be independent of both the formal system and the implementation used to produce the library and offer a uniform high-level interface for both users and machines to access the combined library.

We claim that MathHub.info will resolve two major bottlenecks in the current state of the art. It will provide a permanent archiving solution that not all systems and user communities can afford to maintain separately. And it will establish a standardized and open library format that serves as a catalyst for comparison and thus evolution of systems.

Concretely, we see three ways the formal methods and mathematical knowledge management communities can benefit from MathHub.info: *i*) users can view formerly disparate developments in a common, neutral framework and compare them, *ii*) system developers can import libraries from other logical systems to extend the reach of formalizations and avoid duplicate development *iii*) the existence of a library management system (and importable content) can lower the entry hurdle for developing new logic-based systems. In the next section we present the current system architecture and realization, and Section 3 concludes the paper.

## 2 System Architecture and Realization

MathHub.info is realized as an instance of the Planetary System [Koh12], which we have substantially extended in the course of the work reported here.

The system architecture has three main components: *i*) a versioned *data store* holding the source documents *ii*) a *semantic service provider* that imports the source documents and provides services for them *iii*) and a *frontend* that makes the sources and the semantic services available to users. Specifically, we use the GitLab repository manager [GL] as the data store, the MMT API as the semantic service provider and Drupal as the frontend.

Figure 1 shows the detailed architecture.

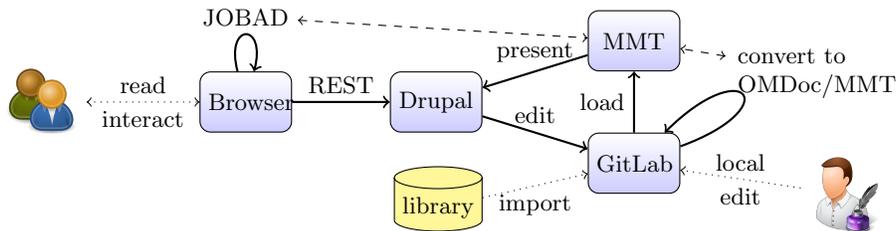


Fig. 1. The MathHub.info Architecture

In this setup, Drupal serves as a container management system<sup>1</sup> that supplies uniform theming, user management, discussion forums, etc. GitLab on the other hand, provides versioned storage of the content documents, and organizes them into repositories owned by users and groups. The advantage of this setup is that we can combine two methods for accessing the contents of **MathHub.info**: *i*) an online, web-based editing/interaction workflow for the casual user, in the spirit of the Planetary system and *ii*) (new) an offline editing/authoring workflow based on a GIT working copy. The latter is important for power authors and for bulk editing jobs. A user can fork or pull the relevant repositories from GitLab, edit them and submit them back to MathHub either via a pull request to the repository masters or a direct commit/push. As the content is usually highly networked and distributed across multiple GIT repositories, we have developed a command line tool `lmh` (local MathHub) that manages working copies across repository borders.

In the web-based system, semantic services (notation-based, presentation, definition lookup, relational navigation, dependency management, etc.) are provided by MMT and are made available to the user, primarily by dedicated JOBAD [GLR09] modules. Note that even though the active document functionalities and semantic editing support in MathHub.info are based on OMDoc/MMT representation of the content, the authors interact with the content in the source format. Both of these representations are versioned in GitLab and are converted into OMDoc/MMT by custom transformers. `lmh` also supports running these transformers locally and previewing HTML5 renderings of the generated OMDoc/MMT.

In order to deal with flexiformal mathematical content in OMDoc, we have also extended the MMT API, which was previously restricted to fully formal content. In the extended MMT API, each MMT service works whenever it is theoretically applicable (e.g. type checking when there exists type information, change management when there is dependency information, etc.).

### 3 Conclusion

MathHub.info is deployed at <http://mathhub.info> and has reached a state, where it can be used for initial experiments and resources, but has not been scaled much beyond 10 000 documents and a couple of dozens of users and repositories yet.

Specifically, we are currently hosting a test set of formal and informal mathematical content to develop and evaluate system functionality; concretely: *i*) the SMGloM termbase with ca. 1500 small  $\text{\LaTeX}$  files containing definitions of mathematical terminology and notation definitions. *ii*) ca. 6500 files with  $\text{\LaTeX}$ -encoded teaching materials (slides, course notes, problems, and solutions) in Computer Science, *iii*) the LATIN logic atlas with ca. 1000 meta-theories and logic morphisms, *iv*) the Mizar Mathematical Library of ca. 1000 articles with ca. 50.000

---

<sup>1</sup> Drupal and similar systems self-describe as content management systems, but they actually only manage the documents without changing their internal structure.

theorems, definitions, and proofs, and *v*) a part of the HOL Light Library with 22 theories and over 2800 declarations. Already now, it is unique in its class in that it gives a unified interface to multiple theorem prover libraries together with linguistic and educational resources. Now that the ground work has been laid, we anticipate the rapid integration of new semantic services, editing support and new content.

*Acknowledgements* This work has been partially funded by DFG under Grant KO 2428/13-1. The authors acknowledge that the MathHub system builds on a long series of experiments in system integration in the KWARC group and that the design and implementation would not have been possible without substantial discussions in the group.

## References

- [Cod+11] Mihai Codescu et al. “Project Abstract: Logic Atlas and Integrator (LATIN)”. In: *Intelligent Computer Mathematics*. Ed. by James Davenport et al. LNAI 6824. Springer Verlag, 2011, pp. 289–291. ISBN: 978-3-642-22672-4.
- [EUD] *EuDML – The European Digital Mathematics Library*. URL: <http://eudml.eu> (visited on 08/02/2011).
- [GL] *GitLab*. URL: <http://gitlab.org> (visited on 02/24/2014).
- [GLR09] Jana Giceva, Christoph Lange, and Florian Rabe. “Integrating Web Services into Active Mathematical Documents”. In: *MKM/Calculamus Proceedings*. Ed. by Jacques Carette et al. LNAI 5625. Springer Verlag, July 2009, pp. 279–293. ISBN: 978-3-642-02613-3. URL: <https://svn.omdoc.org/repos/jomdoc/doc/pubs/mkm09/jobad/jobad-server.pdf>.
- [Ian+13] Mihnea Iancu et al. “The Mizar Mathematical Library in OMDoc: Translation and Applications”. In: *Journal of Automated Reasoning* 50.2 (2013), pp. 191–202. DOI: 10.1007/s10817-012-9271-4.
- [Koh06] Michael Kohlhase. OMDoc – *An open markup format for mathematical documents [Version 1.2]*. LNAI 4180. Springer Verlag, Aug. 2006. URL: <http://omdoc.org/pubs/omdoc1.2.pdf>.
- [Koh12] Michael Kohlhase. “The Planetary Project: Towards eMath3.0”. In: *Intelligent Computer Mathematics*. Ed. by Johan Jeuring et al. LNAI 7362. Berlin and Heidelberg: Springer Verlag, 2012, pp. 448–452. ISBN: 978-3-642-31373-8. arXiv: 1206.5048 [cs.DL].
- [MizLib] *Mizar Mathematical Library*. URL: <http://www.mizar.org/library> (visited on 09/27/2012).
- [RK13] Florian Rabe and Michael Kohlhase. “A Scalable Module System”. In: *Information & Computation* 0.230 (2013), pp. 1–54. URL: <http://kwarc.info/frabe/Research/mmt.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/mcs10.pdf>.