

# DynaTable: A Wiki Extension for Structured Data

Carrie Arnold

Todd Fleming

David Largent

Chris Lürer

Ball State University

Computer Science Department

Muncie, Indiana, USA

{cmarnold2,tjfleming,dllargent,clueer}@bsu.edu

## ABSTRACT

DynaTable is an extension for the MediaWiki software that provides support for structured data. While tables and lists are popular features in many wikis, they are currently unsupported as first-class entities. As a consequence, creating and editing tables is a manual, error-prone task. DynaTable allows wiki editors to create tables that can be displayed on and transcluded to multiple wiki pages, and can be partially displayed based on different criteria..

## Categories and Subject Descriptors

H.3.5 [Information Storage and Retrieval] On-line Information Services – *Web-based services*; H.5.4 [Information Interfaces and Presentation] Hypertext/Hypermedia – *navigation, user issues*.

## General Terms

Design, Human Factors.

## Keywords

Wikis, structured data, tables, MediaWiki.

## 1. INTRODUCTION

Wikis are a popular and useful method of dynamically collaborating on open text [8]. The data on wiki pages most often takes the form of unstructured text, but is also often contained in tables and lists, which are essentially tables with a single column. We call these types of data *structured data*, since they are organized according to a well-defined schema, such as the rows and columns in a table. Typically, each row or column has a well-defined meaning. For example, a table could use rows to represent countries and columns to represent their area, their population, and so on. Tables could include “countries by area”, “countries by population”, and so on.

In many cases it is desirable for user-created tables to be transcluded (transparently included) across wiki pages in differing representations. In the example of lists of countries, whenever one of the data values about a country changes, it is likely that several of these tables need to be updated manually, and there are

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WikiSym '09, October 25–27, 2009, Orlando, Florida, U.S.A.

Copyright © 2009 ACM 978-1-60558-730-1/09/10...\$10.00.

likely to be inconsistencies between the tables. Explicit support for structured data that allows these data to be automatically sorted, filtered, and transcluded would address these issues

We chose MediaWiki [2] [5] as the basis for our work because of its large user base and its extension mechanism. MediaWiki has support for lists and tables, but no explicit support for structured data. Templates [6] [7] are often used to display and format structured data, for example in Wikipedia infoboxes. However, templates are not specifically designed for structured data, and are too complex to create for most wiki users.

The concept of first-class entity tables in wikis was originally proposed by the WikiData project, a modification involving multiple changes to the underlying software [3]. This was never fully realized, but two extensions to the MediaWiki software have been under development to implement this concept. The DataTable extension [1] is a small extension allowing for data to be declared on a wiki page, then retrieved and displayed on other pages with a special tag. It does not address contradictory table definitions, nor create a user friendly way of editing table data or schemas.

WikiDB [4] is a MediaWiki extension that creates a new namespace in the MediaWiki software for the declaration of tables, then stores that information in three database tables in the MediaWiki database. This code remains under development, but has been released under a Creative Commons license. However, WikiDB allows for the declaration of structured data on any wikipage, and does not eliminate the need for users to search wiki pages to find data declarations or change them. WikiDB stores data inside the MediaWiki database in a way that makes it difficult to retrieve the data and for administrators to make changes.

## 2. REQUIREMENTS FOR STRUCTURED DATA SUPPORT IN WIKIS

Users should be able to define both the data schema and the data itself, using easily located special namespaces. The table should be created and the data inserted after they are defined in wiki markup, then retrieved and displayed. The user should also be given the option to display the data anywhere, and as many times as desired, on the wiki. Record duplication should be handled gracefully using smart insertion/updates, always checking against the database schema, which would also allow the system to ignore any columns that did not already exist in the table.

Tables should be dynamically updated (specifically in terms of adding and deleting columns) after they were created by redefining the schema on the table definition page.

Company name	Logo	Founded	Location	Revenue
Intel	Image:Intel logo.jpg	USA		
SD Squared		January 2009		\$27.42
Apple	Image:Apple logo.png	April 2005		\$7.42
Microsoft		1492	Seattle	\$8

Code:

```
<echo table="Companies" THColor="orange"
ERColor="yellow"></echo>
```

**Figure 1. Table example.** A DynaTable wiki table shown formatted and as markup. The data inside the table are transcluded from a separate page.

Once the structured data has been entered, an easy yet powerful method should be provided for defining how to display the data on a wiki page. This should include the ability to:

- sort the data;
- filter the data;
- specify what columns should be displayed;
- format the appearance of the table;
- transclude structured data in several pages, in the same way as can be done with images.

### 3. THE DYNATABLE APPROACH

We created a MediaWiki extension called DynaTable, which is inspired by WikiDB [4]. The main goal of DynaTable is to provide wiki users with a convenient way to define and display structured data within a wiki (see Figure 1).

The following are the features of DynaTable:

- It allows for the definition of structured data via wiki text in a separate wiki namespace. The structured data is not defined or created on a regular wiki page.
- When an entry is provided for structured data, that entry is placed in a database table. The data is also stored in wiki text in a separate wiki namespace. It is not stored in a regular wiki page in any way.
- Echoing (displaying) of structured data on a regular wiki page is accomplished solely with the use of a tag (and needed parameters) referencing the structured data. The data (and the meta-data) are dynamically retrieved from the database at the time of the wiki page parsing or rendering (see Figure 1).

A given set of structured data is given a name and is defined and its data entered on wiki pages in a separate namespace. This namespace is the only place where the tags used to define and enter data are allowed to be used. By limiting their use in this way, the definition and entering of structured data is done in a single place, making maintenance of the structured data much easier and more accurate.

When the user chooses to display some or all of a given set of structured data, a tag is used to indicate where the table should be displayed. Parameters allow for the sorting and filtering of the

data. Significant control over the appearance of the table (for example, background color, border size and cell spacing) is provided to the user by the use of additional tag parameters which make use of wiki templates. By default, all fields of the structured data are displayed, but the user has the option of specifying exactly what fields, and in what order the fields should be displayed. Additionally, a row counter can be easily added to the first column of the table.

### 4. CONCLUSIONS

We argued for the need for wikis that support structured data. To address this issue, we introduced DynaTable, a MediaWiki extension that adds support for explicit, dynamic tables. DynaTable allows wiki users to:

- create and edit tables in a dedicated table namespace;
- transparently include these tables in other wiki pages;
- edit tables and table data separately;
- filter tables by selecting a subset of columns and rows to display.

These features allow collaborative and open editing of data sets. While similar approaches exist, DynaTable is the first technology to realize these features in a wiki context.

### ACKNOWLEDGMENTS

We would like to thank Mark Clements, the author of WikiDB [4], for the permission to study and use his code.

The following people contributed to the design and development of the DynaTable project: Sirisha Bhagavatula, Nagateja Boppa, Soumik Chaudhuri, Mary Fleming, Naveen Irukulapati, Adam Kanuchok, You Li, Keshav Poudyal, Achyuth Pydimarri, Amitesh Ramname, Debduitta Roy, Logesh Sampath, and Ravi Tadavarti.

### REFERENCES

- [1] DataTable extension. <http://www.mediawiki.org/wiki/Extension:DataTable>.
- [2] MediaWiki. <http://www.mediawiki.org/wiki/MediaWiki>.
- [3] Wikidata. <http://www.mediawiki.org/wiki/Wikidata>.
- [4] WikiDB. <http://www.kennel17.co.uk/testwiki/WikiDB>.
- [5] Barrett, D.J. *MediaWiki*. O'Reilly, Sebastopol, 2008.
- [6] Di Iorio, A., Vitali, F. and Zacchiroli, S. Wiki content templating *Proceedings of the 17th International Conference on World Wide Web (WWW)*, ACM, Beijing, 2008, 615-624.
- [7] Haake, A., Lukosch, S. and Schümmer, T. Wiki-templates: adding structure support to wikis on demand *Proceedings of the 2005 international symposium on Wikis (WikiSym)*, ACM, San Diego, 2005, 41 - 51.
- [8] Leuf, B. and Cunningham, W. *The Wiki Way: Quick Collaboration on the Web*. Addison-Wesley, 2001.