

Response from the  
**SIG on Libraries**  
to the grant call for TEI special interest groups for 2010

Submitted by Kevin Hawkins on behalf of the SIG on Libraries

SIG name	SIG on Libraries
Contact	Kevin Hawkins ( <a href="mailto:kevin.s.hawkins@ultraslavonic.info">kevin.s.hawkins@ultraslavonic.info</a> )
Amount requested	\$500–\$5500 (see below)
Date for final report	November 30, 2010

At the meeting of the SIG on Libraries in November 2007, it was noted that the *TEI Text Encoding in Libraries Guidelines for Best Encoding Practices* requires substantial revision in order to make it fully compliant with P5 and might require adjustment due to the advent of TEI Tite. Work on this revision began in earnest after a meeting in April 2008, when various members of the SIG on Libraries undertook a complete rewriting of this document, now called the *Best Practices for TEI in Libraries* (canonical location:

<http://purl.oclc.org/NET/teiinlibraries>). In addition to updating for P5, there were corrigible errors to fix and other aspects that needed rethinking to give more explicit guidance to libraries on use of the TEI. The revised prose of this document is nearly finalized, with a few minor changes and the completion of ODD specifications for the various levels of encoding remaining before publicizing this major revision.

The working group has received numerous requests to develop stylesheets for converting between MARC records (the most common metadata format in libraries) and TEI headers according to the *Best Practices*, and requests for stylesheets that convert between TEI headers and various library metadata formats come up periodically on TEI-L. While encoding practice (both choice of fields and content use within them) in headers and even MARC records varies between institutions, we believe that development of robust stylesheets will solidify the use of the *Best Practices* and encourage use of TEI in libraries.

Michael Sperberg-McQueen of Black Mesa Technologies LLC volunteered in October 2009 to begin development of the stylesheets, spending 22.3 billable (but pro-bono) hours on this between October and December 2009. The results of his work are freely available (licensed under GNU GPL) in a package called Thutmose I (documentation and links to the stylesheets at <http://blackmesatech.com/2009/10/thutmose/L1/userdoc.xml>). Thutmose I translates from MARCXML (a more contemporary variant of MARC, for which tools to convert to and from traditional MARC are readily available) to TEI headers.

The SIG on Libraries proposes a grant to support further development of Thutmose by Black Mesa Technologies. While there is much work that could be done, Michael suggests that given the following amounts of grant funding, he could accomplish the following:

Item	Cost	Description
A	\$500	Add support for as many MARC fields not yet mapped to appropriate elements in the TEI header as possible.
B	\$500	Improve support for the <code>tei-from-source</code> and <code>tei-from-digital</code> mapping variants (which are not as well supported as the <code>source</code> variant).
C	\$750	Improve functionality of the <code>beautification</code> parameter.
D	\$750	Create a web form for the tool allowing a human editor to select for each field whether to use the beautified or original version of the field.
E	\$2000	Add support for <i>all</i> MARC fields mentioned in the <i>Best Practices</i> and for all three mapping variants ( <code>source</code> , <code>tei-from-source</code> , and <code>tei-from-digital</code> ).
F	\$2000	Support for TEI-to-MARC conversion. (Only MARC-to-TEI is currently supported.)

Thus the minimum grant that would be useful to us is \$500, and the maximum is \$5500.<sup>1</sup> If we receive any amount in that range, the working group revising the *Best Practices* will choose which options offered by Michael best fit our priorities given the amount of funding.

---

<sup>1</sup> Because A and B are included in E.