Project Investigator Name: Hikaru Nakano
Department: Cataloging and Metadata
Email address: hiknaka@uflib.ufl.edu Phone: 273-2727

Additional applicants, please give name and email for each:
Nancy Poehlmann: nanpoeh@uflib.ufl.edu
Alena Aissing: alena@uflib.ufl.edu

Title of grant application project:
Cyrillic character inclusion in University of Florida bibliographic records

Project abstract (no more than 100 words):
This project intends to include Cyrillic characters particularly used in Russian language in University of Florida bibliographic records. Currently, approximately 10,000 records have only Romanized fields. As many studies have shown, searchability for foreign language materials is enhanced when these records include the original language, and this have been proven by the UF library's efforts in Chinese and Japanese character inclusion, as well as national efforts in JACKPHY languages (Japanese, Arabic, Chinese, Korean, Persian, Hebrew and Yiddish).

Funds requested (Limit of $5,000):
$5,000

Describe how the 10% mandatory cost share will be met (be specific):
In-Kind support
- Provision of staff time
- Access to technology equipment: PC and printer

Please list the library resources to be used in this project and the name of the person authorizing the intended use and date authorized. If you need more room, continue on a separate page.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Authorizing Individual</th>
<th>Date Authorized</th>
</tr>
</thead>
<tbody>
<tr>
<td>One workstation for students' use (Cat. and Metadata Dept.)</td>
<td>Betsy Simpson</td>
<td>1/15/2009</td>
</tr>
</tbody>
</table>

I have my Department Head's approval to submit this application.  _____ Yes  _____ No

Signature: ________________ Date: 1/15/09
a. Description of Project

This project will add Russian (Cyrillic) characters to 10,000 existing MARC records, using Latin2Cyrillic, a transliteration macro developed by Joel Hahn, Niles Public Library District in Niles, Illinois to be used with the OCLC cataloging interface. The vernacular characters will be added to OCLC Worldcat and UF’s local catalog, as well as to the state-wide Endeca catalog.

b. Statement of importance

Tatiana Barr at Yale, and previously Head of Humanities and Special Collections at UF, wrote “In the field of Slavic studies, some of the holdings of American libraries are not duplicated anywhere and there is a growing demand for them all over the world.” A study by Alena Aissing, Slavic Studies Librarian at UF, showed the limitation of Russian records that only contain Romanized fields, and emphasized the importance and necessity of adding Cyrillic vernacular scripts to Cyrillic records (Aissing, 1995).

JACKPHY languages (Japanese, Arabic, Chinese, Korean, Persian, Hebrew and Yiddish) have historically been given priority for input. The Library of Congress has provided cataloging in machine-readable form for JACKPHY materials for more than 20 years, followed by many other cataloging agencies. Therefore, JACKPHY records can be easily downloaded into an institution’s local system by copy catalogers or students.

By contrast, Cyrillic characters have only been added into OCLC since OCLC Connexion was released in 2004, due to the technological limitations of earlier interfaces. Romanized transliteration is often imprecise and can render an online record irretrievable. Ingrid Kleespies, Slavic Languages professor at UF, states: “Although there is a standard system of transliteration, I find that there are glitches that slow down my research: older catalogue records were entered under a different transliteration system, meaning that sometimes I have to try spelling titles and names several different ways before I can find the record.” The availability of Cyrillic scripts eliminates this problem by allowing users to search and retrieve records in either Romanized transliteration or Cyrillic script. This improvement will deliver easier and timelier access to a significant part of the UF library collections.

As of Jan. 2009, approximately 10,000 Russian monographic records at the UF libraries do not contain vernacular scripts. ALEPH’s Unicode support also means that UF can display many more vernacular scripts than before, which is a major requirement for bibliographic records from the entire world not only to be viewed but also to be indexed and retrievable. If this project is successful, more than 30,000 fields including author, title, contents information, etc. in 10,000 records will be enhanced (see p.4).

Cyrillic script addition will be a major contribution to the advancement of East European scholarship and research, but it contributes to other fields as well. According to Madden (2008), “[m]any Judaica materials are in Cyrillic script as are 50+ languages, mostly Eastern European and Central Asian.” Concretely, more than 300 students who
take language courses and content course, and not to mention, 11 faculty members will benefit from the improvement of OPAC (Kleespies, 2008). In addition, the availability of Cyrillic scripts will foster the sharing of Russian-language bibliographic data with all SUS libraries, by utilizing Union MARC in Endeca. On a national level, as a PCC (Program for Cooperative Cataloging) institution, UF has pledged to create records that will maximize their utility to all OCLC member libraries. Tatiana Barr indicates in her letter that several institutions, including Stanford, have begun adding these fields in 2008. The enhancement of these 10,000 OCLC records will be a tremendous contribution to OCLC member libraries worldwide.

c. Description of resources needed
This proposal requests funding for stipends for two OPS students for two semesters to compete quality control (QC) that will measure the attributes and performance of Russian bibliographical records after implementing the project. The macro must be checked after de-transliteration by human eyes with language expertise. Recently, Cataloging and Metadata dept. hired two OPS students (one regular student and one Federal Work Study student), who have near-native familiarity with the language, at a wage of $9/hr. The wage was determined based on the Student Assistant Pay Rate Policy. The students check the problems with de-transliteration such as corrupted characters, incorrect hyphens or spacing, incorrect character choice, and correct them as necessary. When the records are correct, the students update the OCLC records as well as the Aleph records. These tasks require skills beyond the usual student-level work.

For the pilot project, the Cat/Met department set aside $1,350 from the FY08-09 student OPS budget for 240 hours (($9/hr and $2.25/hr)* 8 hr/wk*15 wk). Based on the students’ actual performance, we expect that they will revise 9 records / hour on average. At this rate 2,160 records will be revised during the pilot project. We are requesting continued funding for these same two students who will revise the remaining approximately 8,000, records. For the full project, total working time will be 888 (8,000/9; 444 each) hours, and it will cost approximately $5,000 (444*9+444*2.25=4,995)

d. Action plan

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Activity</th>
<th>Responsible Party</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot project</td>
<td>Jan. 2009 Hire OPS students</td>
<td>Hikaru Nakano</td>
</tr>
<tr>
<td></td>
<td>Jan. 2009 Training</td>
<td>Randy Jewett /</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hikaru Nakano</td>
</tr>
<tr>
<td></td>
<td>Spring 2009 Pretest</td>
<td>Alena Aissing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nancy Poehlmann</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hikaru Nakano</td>
</tr>
<tr>
<td></td>
<td>Apr. 2009 Pilot project ends (240 hrs.)</td>
<td>Jason Fleming</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Randy Jewett</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Michael Jay</td>
</tr>
<tr>
<td>Full project</td>
<td>Apr. 2009 Actual project starts</td>
<td>Jason Fleming</td>
</tr>
<tr>
<td></td>
<td>-Update list</td>
<td>Randy Jewett</td>
</tr>
<tr>
<td></td>
<td>-Training and advice throughout the project</td>
<td>Michael Jay</td>
</tr>
<tr>
<td></td>
<td>-Statistic Macro</td>
<td></td>
</tr>
<tr>
<td>Jul. 2009</td>
<td>Presentation at ALA Annual meeting in Chicago (proposal accepted)</td>
<td>Nancy Poehlmann, Alena Aissing, Hikaru Nakano</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Fall 2009</td>
<td>Pretest</td>
<td>Alena Aissing</td>
</tr>
<tr>
<td>Dec. 2009</td>
<td>OPS students finish adding characters</td>
<td></td>
</tr>
<tr>
<td>Jan. 2010</td>
<td>Posttest</td>
<td>Alena Aissing Nancy Poehlmann Hikaru Nakano</td>
</tr>
</tbody>
</table>

e. **Is project collection specific?**
There are no collection issues; all materials are owned by UF libraries, and no copyright issues are involved.

f. **Measurement of success**
During the pilot project, Hikaru Nakano, working with Alena Aissing, will conduct a survey in two Russian language classes (one at a beginning level, and another at an advanced level) to assess user satisfaction and ease of access to Russian bibliographical information in the OPAC. The survey results will be compared to another set of surveys performed, again at two different levels, at the end of the grant period. The expectation is that there will be a measurable improvement of scholars’ and students’ ability to retrieve information from the enhanced bibliographical data. Furthermore, once students are used to retrieving Russian materials by typing vernacular scripts, this improved information literacy skill will take them to further exploration of Russian databases.

Successful completion of this project will lead to inclusion of Cyrillic characters in other languages using this script, as well as inclusion of other scripts not included in the JACKPHY priority languages. Faculty input will be sought for guidance as to which languages to pursue. Hikaru Nakano, Alena Aissing and Nancy Poehlmann will present the project to a national audience during the ALA annual meeting in Chicago in June, 2009. Future plans include an article to be published in a peer-reviewed library science journal to encourage other libraries to pursue vernacular character inclusion.

g. **Budget**
See page 5

[Reference]


Kleespies, Ingrid (personal communication, Nov. 25, 2008)

Madden, Emily (personal communication, Nov. 06, 2008)
Before and After Cyrillization

BEFORE

Aleksey, Vasiliy Pavlovich, 1932-
Title: V poslednom gryaznom bliz' - poyest' - Bumazhnyy pechat
Description: 666 p. : ill. ; 21 cm.
ISBN: 5966610112
9785966610117
Subjects, general: Russians - United States - Fiction
Americanization - Fiction
Subjects, geographic: United States - Fiction
Other author(s), etc: Aleksey, Vasiliy Pavlovich, 1932-
Other title(s): Bumazhnyy pechat
Format: BK

AFTER

Aleksey, Vasiliy Pavlovich, 1932-
Title: V poslednom gryaznom bliz' - poyest' - Bumazhnyy pechat
Description: 655 p. : ill. ; 21 cm.
ISBN: 5968810112
9785968810117
Subjects, general: Russians - United States - Fiction
Americanization - Fiction
Subjects, geographic: United States - Fiction
Other author(s), etc: Aleksey, Vasiliy Pavlovich, 1932-
Other title(s): Bumazhnyy pechat
Format: BK
Material type: <Book> <Fiction>
**Mini Grant Budget Form**

Please add lines to table as needed. If you need help completing this form, please contact Bess de Farber, PH# 273-2626.

### 1. Salaries and Wages (no fringe benefits required)

<table>
<thead>
<tr>
<th>Name of Person</th>
<th>Salary times % of effort</th>
<th>Grant Funds</th>
<th>Cost Share</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Jay</td>
<td>$20.313 per hour * 10 = 203.13</td>
<td>$0.00</td>
<td>$203.13</td>
<td>$203.13</td>
</tr>
<tr>
<td>Jason Fleming</td>
<td>$267.47 per hour * 5 = 133.735</td>
<td>$0.00</td>
<td>$133.74</td>
<td>$133.74</td>
</tr>
<tr>
<td>Randy Jewett</td>
<td>$16.53 per hour * 10 = 165.3</td>
<td>$0.00</td>
<td>$165.3</td>
<td>$165.3</td>
</tr>
<tr>
<td>OPS student(s)</td>
<td>$9 per hour * 444 hrs * (9+2.25)</td>
<td>$5,000.00</td>
<td>$</td>
<td>$5,000.00</td>
</tr>
</tbody>
</table>

**SUBTOTAL** | $5,000.00 | $502.17 | $5,502.17 |

### 2. Equipment

<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity times Cost</th>
<th>Grant Funds</th>
<th>Cost Share</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
</tbody>
</table>

**SUBTOTAL** | $0.00 | $0.00 | $0.00 |

### 3. Supplies

<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity times Cost</th>
<th>Grant Funds</th>
<th>Cost Share</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
</tbody>
</table>

**SUBTOTAL** | $0.00 | $0.00 | $0.00 |

### 4. Travel

<table>
<thead>
<tr>
<th>From/To</th>
<th># of people/# of days</th>
<th>Grant Funds</th>
<th>Cost Share</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
</tbody>
</table>

**SUBTOTAL** | $0.00 | $0.00 | $0.00 |

### 5. Other (services vended, etc.)

<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity times cost</th>
<th>Grant Funds</th>
<th>Cost Share</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
</tr>
</tbody>
</table>

**SUBTOTAL** | $0.00 | $0.00 | $0.00 |

### Total Direct Costs (add subtotals of items 1 through 5)

<table>
<thead>
<tr>
<th>Grant Funds</th>
<th>Cost Share</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$5,000.00</td>
<td>$502.17</td>
<td>$5,502.17</td>
</tr>
</tbody>
</table>
From: Gorham, Michael S [mailto:mgorham@UFL.EDU]
Sent: Monday, September 22, 2008 2:24 PM
To: Nancy Poehlmann
Subject: RE: Vernacular language in catalog records

Dear Nancy,

Adding Cyrillic would greatly increase the speed and accuracy with which catalog searches are conducted. The existing system of transliteration is imprecise (co-existing with at least 2 other systems commonly used by scholars) and arcane for poor Russian-language students have invested enough time as it is in learning Cyrillic. Unfortunately, maintaining the current "Latin-only" policy will indirectly encourage students and potential scholars of the next generation to rely only on Google and Wikipedia for their data sources and believe they are getting the same (if not a greater) quality product as academic library interfaces can provide.

Hope this helps (and feel free to edit)!

Michael Gorham

________________________
Michael S. Gorham
Associate Professor of Russian Studies
Department of World Languages and Cultures University of Florida

301 Pugh Hall
P.O. Box 115565
Gainesville, FL 32611-5565

Tel: 352-392-2101 x206
Fax: 352-392-1067
http://www.clas.ufl.edu/users/mgorham

From: Nancy Poehlmann [nanpoeh@uflib.ufl.edu]
Sent: Thursday, September 18, 2008 11:38 AM
To: Gorham, Michael S
Subject: Vernacular language in catalog records

Dear Dr. Gorham,

Would you be willing to write a few sentences for me in a return email regarding the importance of being able to search the library catalog in Cyrillic characters? I am writing a grant proposal to add Cyrillic to our records, and need to convince the granting agency that scholars would find this addition helpful. I have anecdotal evidence, but need actual quotations. Thank you for your consideration.

Regards,
Nancy Poehlmann

Nancy Mitchell Poehlmann, MA, MLS
Head, Humanities and Special Collections Cataloging Unit P. O. Box 117007
University of Florida Gainesville, FL 32611-7007
(352) 273-2728 (vox)
(352) 392-7365 (fax)
Appendix 2

From: Kleespies, Ingrid [mailto:iakl@ufl.edu]
Sent: Tuesday, November 25, 2008 10:52 AM
To: Nancy Poehlmann
Subject: RE: Vernacular language in catalog records

Dear Nancy,

I’m very sorry not to have responded to you until now! I’m afraid this email got buried in my inbox – it’s been a very, very busy semester. Do you still need a brief statement about the importance of being able to search in Cyrillic? I’ve written a few words below in the hope that I may still be of help.

Again, I greatly apologize for being so slow to respond.

Best,
Ingrid
The proposal to add Cyrillic to the library records at UF is extremely important to me as both a scholar and a teacher. I am a scholar of nineteenth and even eighteenth century texts. Although there is a standard system of transliteration, I find that there are glitches that slow down my research: older catalogue records were entered under a different transliteration system, meaning that sometimes I have to try spelling titles and names several different ways before I can find the record. Cyrillic in the catalogue would be a very valuable tool for our students, who are often at beginning stages of research with Russian materials. They are not familiar with the Library of Congress transliteration system, so frequently miss many useful sources. The addition of Cyrillic would greatly streamline our research process.

From: Nancy Poehlmann [mailto:nanpoeh@uflib.ufl.edu]
Sent: Thursday, September 18, 2008 11:40 AM
To: Kleespies, Ingrid
Subject: Vernacular language in catalog records

Dear Dr. Kleespies,
Would you be willing to write a few sentences for me in a return email regarding the importance of being able to search the library catalog in Cyrillic characters? I am writing a grant proposal to add Cyrillic to our records, and need to convince the granting agency that scholars would find this addition helpful. I have anecdotal evidence, but need actual quotations. Thank you for your consideration.

Regards,
Nancy Poehlmann

Nancy Mitchell Poehlmann, MA, MLS
Head, Humanities and Special Collections Cataloging Unit
P. O. Box 117007
University of Florida
Gainesville, FL 32611-7007
(352) 273-2728 (vox)
(352) 392-7365 (fax)
npoehlmann@ufl.edu
Why Cyrillic?

For many decades the use of the romanization of vernacular scripts was considered acceptable by the Anglo-American cataloging community. However, with the growing internationalization of the field of librarianship, global access to the Web and library OPACS, as well as a recognition by such institutions like OCLC and RLG of the international community as part of their mandate, and new technology, the use of vernacular script in Anglo-American bibliographic records becomes not only an enhancement but a necessity if we are to participate in the globalization of access to our valuable resources. In the field of Slavic studies, some of the holdings of American libraries are not duplicated anywhere and there is a growing demand for them all over the world.

With the implementation of Unicode compliant OCLC Connexion 2.0 and Voyager 6.1, we now have the technical capability of creating, editing, importing, and exporting many non-Latin scripts. Our peer institutions are increasingly implementing systems that are UNICODE compliant, and are beginning to create parallel fields in Cyrillic, Greek, and Thai scripts, in addition to the JACKPHY scripts. By creating a policy and procedure supporting the programmatic acceptance and creation of these non-Latin scripts, we can ensure consistency in our own catalog, and also support national cooperative cataloging efforts in which our peers also participate.

Over the decades several institutions have taken the initiative. The University of Illinois with its famous Russian Research Center was one of the earliest to add vernacular script and it was only available to visiting scholars for a while. But increasingly more institutions have followed suit when their collections warranted it and they had the skills and technology to institute it. With the growing sophistication of browsers and search engines as well as the ability to key in Cyrillic, more Russian and other language scholars and users are searching using Cyrillic script, in a word, their own language. English may be known by many but more seriously affecting access is that not everyone, in fact, only scholars and librarians, know the variety of transliteration systems being used worldwide. To add Cyrillic makes the University of Florida’s substantial, 11,000 and growing, items accessible to the world in a way they have not been.

In an informal inquiry put on the Slavic Librarians’ Listserv, the following institutions reported putting Cyrillic into their Bibliographic or both Bibliographic and Authority records:

Library of Congress, University of Wisconsin-Madison (and they are doing it retrospectively as well for Russian and other Slavic languages in turn), University of Indiana, Columbia University, National University of Kyiv Mohyla Academy. Harvard University edits copy cataloguing records already containing Cyrillic fields, but does add Cyrillic to original records for its Harvard Theatre Collection at the Houghton. Others are: the Center for Jewish History and many of their partner organizations (under the umbrella of the Center) are now adding Cyrillic script (particularly YIVO Institute), Stanford University (which began in November 2008 to add Cyrillic to new records created in Connexion), the University of Maryland Libraries. Davis Library University of North Carolina at Chapel Hill, The Slavic and East European Division at the University of Michigan. Whereas many began in 2008, University of Pittsburgh has been doing it a number of years. Others include: University of Washington and the University of Chicago.

Tatiana Barr, Catalog Librarian/Team Leader, Yale University Library