

## **IN Harmony: Sheet Music from Indiana Email Content Analysis Evaluation Plan**

Indiana University (IU), the Indiana State Library (ISL), the Indiana State Museum (ISM), and the Indiana Historical Society (IHS) have received a grant from the Institute of Museum and Library Services (IMLS) to provide electronic access to Indiana-related sheet music from each of the institutions' collections. Approximately 10,000 pieces of sheet music will be available online as a result of this grant activity. One of the primary goals for this project is to provide robust, consistent browse and search access across all participants' collections or within a particular collection. As a result, cataloging guidelines and tools for sheet music description will be developed to aide the project partners in a) cataloging sheet music not yet described in their respective collections; and b) map existing cataloging records to a format that will facilitate cross-collection searching.

### **Background**

---

The Lilly Library, one of the contributors to the IN Harmony project, is home to several sheet music collections including the Starr Sheet Music Collection and the Sam DeVincent Collection of American Sheet Music. The Lilly receives hundreds of email inquiries regarding sheet music a year, mostly from people not affiliated with Indiana University. Analysis of those email inquiries will help us understand the sheet music needs of the population-at-large, especially those relying on librarians as opposed to search engines for their information needs.

The email content analysis is part of a broader, second phase of usability studies that aim to answer the following questions:

- What is the nature of the sheet music request – academic, personal interest, etc.?
- What are the requesters search parameters?
- Are the requesters looking for known or unknown items?
- Are the requesters interested in musical content or cover art?
  - Are their approaches or query descriptors different depending on whether they are looking for musical content or cover art?
- How do requesters define sheet music?

Phase one of the usability studies involved in-depth analysis of query logs generated from the IU Sheet Music and UCLA Sheet Music Consortium websites<sup>1</sup>.

---

<sup>1</sup> For final report of the logs study, see:  
<http://www.dlib.indiana.edu/projects/inharmony/projectDoc/usability/logs/index.shtml>

## **Purpose of Study**

---

On a weekly basis, the Lilly Library receives email inquiries regarding sheet music either via their “Ask A Question” form<sup>2</sup> or their email, [liblilly@indiana.edu](mailto:liblilly@indiana.edu). Most of these requests are posed by people outside of Indiana University (IU). Audiences of the *IN Harmony* website include users not affiliated with any of the project partner institutions. In order to understand the sheet music needs of the population-at-large, content analysis of the email requests will be performed<sup>3</sup>.

Furthermore, we hope to balance what we learned from our analysis of the log data generated by two sheet music collection websites (Query Logs Analysis Study, #04-9626) with sheet music-related queries posed to librarians.

Email content analysis will help us determine guidelines and recommendations for cataloging of sheet music. They will also inform searching and browsing requirements for the end-user *IN Harmony* website.

## **Methodology: Data Gathering and Analysis**

---

### **Data Gathering**

The email messages will come from the Indiana University Lilly Library. Only sheet music-related requests sent to [liblilly@indiana.edu](mailto:liblilly@indiana.edu) will be analyzed. Rebecca Cape, reference librarian, sheet music specialist and moderator of the liblilly email account, will collect nearly a year’s worth, from August 2004 to mid-May 2005, of sheet music emails for analysis.

The email messages will be stripped of all personal identification information such as name, email or physical addresses, etc. by Rebecca Cape. Only reference questions and other contextual information such as need, intended use, etc. will be assessed by researchers.

Our goal is to analyze approximately 100 messages collected within the above-stated time frame.

### **Data Analysis**

Content analysis is the methodology that will be used to categorize and analyze the email messages. A preliminary analytic coding scheme has been developed and will be refined as data is coded. This scheme includes three dimensions: content (e.g. nature of inquiry), search and retrieval strategy (e.g. what/where/how of search and retrieval) and profile (e.g. teacher). Each dimension contains a set of categories. Below is a preliminary listing of categories that will be modified during analysis:

#### **Content: What type of information is the user requesting?**

- Information need (lyrics, music to perform, etc.)
- Type of inquiry (based on lyrics, title, etc.)
- Known v. unknown item request

---

<sup>2</sup> <http://www.indiana.edu/~liblilly/ask.shtml>; The form sends a structured message to the liblilly email account.

<sup>3</sup> Content analysis is a methodology used to categorize and analyze in textual data such as email messages.

- Nature of inquiry (academic, personal interest, etc.)
- Interest in musical content or cover art
- Copyright questions

**Search & Retrieval Strategy: What is the discovery approach taken by the user? How does the user expect to gain access to the content?**

- Resources consulted (e.g. sheet music website, OAI record, OPAC, film, etc.)
- Query term used (if posed to an electronic system)
- Copy request (print, digital, etc.) and how (mail, fax, download, email, etc.)

**Profile: Who are the users in terms of profession and why are they looking for sheet music?**

- Academic, research or scholarly use
- Personal use (event such as wedding, birthday, etc.)
- Professional affiliation (teacher)

Coding rules will be further developed once a sample of the email messages is consulted. Below is a list of basic coding rules we will follow:

- Each question is coded separately even if part of the same email.
- Assign each question a unique identifier for later reference.
- Each question should be analyzed in light of all three dimensions: content, search & retrieval strategy and profile.
- Two passes of coding will take place, first by the researcher and second by a domain expert.
- Inductively derive additional categories or dimensions.
- If category is unknown, label as “unknown” until further inspection.
- If clarification is needed regarding a request, contact Rebecca Cape at the Lilly Library. Provide Rebecca with exact message so she can search it within her email records/client.

Microsoft’s Excel spreadsheet program will be primarily used for analysis. Data can be easily sorted, coded and displayed in various views including graphical representations.