1 Nexus Data Codebook

1.1 About the Data

1.1.1 Data Inclusion

The pilot dataset spans a fifteen-year period from 1998 through 2013. Every program in the dataset meets the following criteria.

- *US-based:* Programs had to be based in the U.S. and geared toward U.S. librarians.
- *Leadership focused:* Program curriculums demonstrably trained attendees in leadership skills, not management skills. As described below, these are separate skillsets. Some of the programs included in the dataset do contain a management-training component; all of the programs included contain leadership-training components.
- *Library focused:* Programs were included only if they focused primarily on librarians.
- *Run between 1998-2013:* Programs included actively hosted cohorts during some portion of this designated time span.

The dataset was used to record both ongoing and concluded activities.

1.1.2 Information Sources

The dataset was compiled through manual data mining of published reports, articles, websites, and promotional materials. For historical data, the team made extensive use of the Internet Archive's WayBack Machine to locate preserved copies of defunct, updated, and relocated websites.

In January 2014, after a preliminary dataset had been gathered, the research team submitted an email survey to the contact people associated with each program. The email contained a brief summary of the project and a full listing of their program's data entry, and it invited their feedback, corrections, and validation regarding the data we had recorded. Entries that were verified and/or adjusted through this review (26%) are marked as such in the dataset's "completion" field.

1.1.3 Dataset Characteristics

While each program is unique, conducting a broad analysis required classifying programs according to shared features. Because overly broad or overly specific categorizations can frustrate analysis, an effort was made to strike the correct balance. This investigation applied the following categories:

1. *Program Type:* The programs represent a wide range of designs and models. Some programs included features from multiple categories; such programs were assigned only to the descriptive category with which they shared the most characteristics.
    a. Residential – Focus on a week-long, on-site experience; may include longer project.
    b. Fellowship – Meetings spread over longer period, nine months to three years.
    c. Workshop – one-day event.
    d. Virtual – Webinar only program.
2. *Sector:* Many programs targeted specific types of libraries. These have been grouped as:
    a. Academic – University libraries.

    b. Archives – Special collection libraries and archives.
    c. Public – Municipally funded libraries, including K-12 libraries.
    d. Special – Industry or professional linked libraries.
    e. General – All libraries.
3. *Audience:* Many programs targeted audiences based on their career stage or job title. The classifications varied from program to program (i.e., what one program might call "mid-career" another might call "early career" depending upon the way each defined its years of experience. The project team normalized these groups as:
    a. Early career – not managers, first five years.
    b. Mid-career – may or may not be manager, more than five years.
    c. Senior management – e.g. directors.
    d. All – all librarians.

Every effort was made to gather granular, year-to-year data to better understand the changes in leadership programs. However, this data was not always available for programs. For ease of data manipulation and analysis, locations were divided into cities and states, and dates were divided into years, months, and days. Where information on a specific program could not be located, such as a city or a day, the corresponding field was marked "null". Programs that lacked both a clear presentation date and location were excluded from the "year-to-year" analysis table.

As of the publication of this white paper, the dataset is being released as version 1.0. However, the project expects to find more leadership programs and will release updated versions of the dataset as required. At the time of publication, the team discovered a style of leadership program that had not appeared in other literature, programs developed for and run by single university campuses. Programs that bear such a defined internal focus are more difficult to discover and account for; however, these will be added to the dataset as they are known.

1.2 Codebook

1.2.1 Summary

This data was collected as part of an effort to survey the state of library leadership training in the United States between 1998 and 2013. Data was collected from websites, reports, articles, listservs, and surveys of leadership program hosts. Variables included program features, cost, facilitators, locations, and attendees.

1.2.2 Details

Extent: Two tables

Format: TSV

Date: 2014/03/31

Version: 1.0

1.2.3 Dataset components

Name	Purpose	Extent
codebook	Documenting the dataset	
programs	General descriptions of each program recorded	42 columns, 79 rows
yeartoyearlocations	Location, dates, and attendance numbers for each program where possible.	10 columns, 318 rows

1.2.4 Fields

Field Name	Variable	Values or Explanation
programs		
completion	Tracking research state of a given program	Web - all web info available harvested emailed - survey emailed to program director survey results - survey response added
key_field	Program key for linking tables	4 digit integer
program_title	Most commonly used program title	Program name (abbreviation)
program_type	Classification of program type	Workshop - 1 day event Residential - Focus on 2-5 day-long school, can include optional longer project Fellowship - Meetings spread over longer period, 9 months to 3 years Virtual - Webinar only program Symposium - A conference style event with a focus on networking
description	Description of program harvested from promotional materials	
website	URL of program web site	Includes outdated/dead links; must use

		Internet Archive's Wayback Machine for these
organizational_lead	Organization(s) that ran the program	organization 1, subunit (abbreviation of smallest unit); organization 2
organizational_partners	Organization(s) that supported the program with resources or logistics	organization 1, subunit (abbreviation of smallest unit); organization 2
state_offered	Primary U.S. states where the program was hosted	state 1, state 2
lectures	Are lectures part of the program?	Binary
guest_speakers	Are guest speakers beyond the instructors part of the program?	Binary
case_studies	Are case studies of leadership scenarios part of the program?	Binary
field_trips	Does the class visit other institutions?	Binary
discussion	Must participants discuss leadership material?	Binary
projects	Must participants complete projects at their institutions?	Binary
mentors	Do participants have mentors during the program?	Binary
webinars	Are any lessons offered only online?	Binary
networking	Are participants explicitly encouraged to network?	Binary
sector	Type of memory institution	Academic - University libraries Public - Municipally funded libraries, including K-12 libraries Archives - Special collection libraries and archives Special - Industry or professional linked libraries Other - Category not listed above All - All
audience	Career stage of intended audience	Senior management - e.g. directors Mid-career - may or may not be manager, more than 5 years Early career - not managers, first 5 years All - any librarian at any career stage
restrictions	Further requirements for intended audience	Member Based - Attendee must be member of host organization State Based - Attendee must live in defined geographic area Minority Based - Attendee must be a member of targeted demographic
attendance	Total attendance	integer, total since 1998
attendance_complete	Is the attendance figure a lower	binary

	bound	
year_began	First year program was offered	yyyy, year of the first class
year_ended	Last year program was offered	yyyy, Ongoing - program has not ended
length_in_person	Number of days of in-person classes, mentoring, visits, etc	in days
length_total	Total number of days from start to completion	in days
funding_sources	Organizations that support the program financially	organization 1, subunit (abbreviation of smallest unit); organization 2
cost	Highest participant cost	in dollars
discount_cost	Most discounted participant cost	in dollars
discount_reason	Reasons for discount_cost if it exists	free text
student_evaluation	Method of incorporating program evaluations by students	free text
external_evaluation	Method of having a third party evaluate program	free text
contact_name	Current program director	name 1, name 2
founders_facillitators	People or groups that were influential in founding the program	name 1, name 2
key_players	Noting prolific library leadership program contributors	name 1, name 2
analysis_ignore	Whether program falls within project scope	Binary
year_to_year_locations		
key_field	Program key for linking tables	4 digit integer
location_city	City of specific instance of program	string
location_state	State of specific instance of program	string
start_year	Start year of specific instance of program	yyyy
start_month	Start month of specific instance of program	mm
start_day	Start day of specific instance of program	dd
end_year	End year of specific instance of program	yyyy
end_month	End month of specific instance of program	mm
end_day	End day of specific instance of program	dd
attendee_amount	Number of participants of specific instance of program	integer