Data Description

The Program Era Project (PEP) has assembled three related datasets of information on the Iowa Writers’ Workshop (IWW). The first two datasets include the thesis and matriculation records of the IWW from the first creative thesis submitted in 1941 until Spring 2014. These data were initially derived from University of Iowa Library catalog records of MFA theses. Additional data (primarily genre and advisor) were filled in by checking bound copies of theses held in the University Archives.

Also included are small number (~1.5% of total) of BA records from other institutions for early faculty and additional degrees earned at Iowa by IWW grads and faculty (e.g., several IWW graduates also have Nonfiction Writing Program MFAs).

Filename: people.csv (3115 rows)

- **person_id**: numeric id of the individual in the Program Era database
- **full_name**: name as listed in University of Iowa records in "Lastname, Firstname Middlename" format
- **given_name**: (first name)
- **middle_name**
- **family_name**: (last name)
- **pen_name**: preferred pen name of the writer, if different from full name (not exhaustive)
- **gender**: as provisionally labeled by the PEP team based on name; intended for study of broad patterns over time and not as definitive statements on any individual identity. values 'Male', 'Female', 'Gender Nonconfor', 'Unknown'
- **country**: country of origin

Filename: graduations.csv (3051 rows)

- **record_id**: numeric id of the graduation record in the Program Era database
- **graduate_name**: full name of the graduating student
- **graduate_id**: foreign key referring to person_id of the graduating student in people.csv
- **institution**: name of degree-granting institution
- **advisor_name**: full name of the student's thesis advisor
- **advisor_id**: foreign key referring to person_id of the thesis advisor in people.csv
- **year**: year of matriculation (range 1920-2014)
- **program**: program (e.g. Writers' Workshop) or major when no specific program applies (e.g. English)
- **degree**: e.g. "MFA"
- **thesis_title**: Title of thesis where applicable
- **genre**: literary genre of the thesis, e.g. "Poetry", "Fiction", "Literary Translation"

In order to assess demographic trends, we have also assigned gender information for authors in our people.csv dataset. This data was collected by researching biographical information available on authors included in our database. When no biographical information on an author was available, the PEP team attempted to infer (when possible) the author’s most likely gender based on name. While there are instances where an individual author’s gender identity information may not be discernable, our data does allow for a reliable aggregate examination of gender trends in the IWW over time.
Our people.csv dataset offers some initial glimpses into the makeup of the IWW. 96.4% of IWW-affiliated authors come from the United States, .9% from Canada, and 2.7% from other countries. 57.9% of the authors are listed as male, 38.3% as female, 3.9% as unknown and <1% as gender nonconforming.

In addition to the biographical and institutional data described above, we have provided a collection of metadata and publication data connected to a corpus of volumes written by authors affiliated with the IWW. This corpus, provided by the HathiTrust Research Center (HTRC), is comprised of 1228 volumes with 240 unique individuals listed as main authors. Because we knew it would be impossible to assemble a comprehensive corpus of every author in our institutional dataset and because we knew that many authors who attended the IWW never published work, when building our corpus in collaboration with HTRC we used two selection criteria. First, we searched HTRC’s collections for works by any author listed as director of a IWW graduate’s thesis. Secondly, given that we intend our database to drive a public-facing web service, we searched for authors from Wikipedia’s “List of Iowa Writers’ Workshop people,” as this list constituted an example of an externally and publicly defined list of notable IWW authors. We also felt these criteria would maximize our success rate at finding published works for our author list.

When we compare our data in the people and graduations datasets with our corpus, we find 75.6% of works are identified as male-authored volumes and 24.4% female-authored volumes. The publication dates for these works range from 1913 (consisting of books by the earliest creative thesis directors) to 2011.

The volume publication data and metadata provided follow the model established by Ted Underwood’s dataset submitted to the Post45 Collective. Along with author and title information, we have listed each volume's publisher as well as publication date and location (from MARC metadata). We have also provided Library of Congress and OCLC catalog information for the volumes.

Filename: pephtrccorpusdata.csv

- **htid**: HathiTrust item identifier number
- **person_id**: as in people.csv, for linking
- **author1-normalized**: main author name (normalized by PEP team). We have manually added this field to ensure consistency for the searching and counting of lead author names. In some cases, HTRC author listings featured birth and death dates or punctuation that made names inconsistent. We have also included the original author fields as listed by HTRC for reference.
- **author1**: main author (as listed in HTRC record)
- **author2**: secondary author or attribution
- **author3**: tertiary author or attribution
- **author4**: quaternary author or attribution
- **title**: volume title
- **imprint**: publishing imprint
- **date**: publication date
- **place**: publication place, drawn from MARC metadata, abbreviated, see the MARC Code List for Countries to interpret abbreviations (https://www.loc.gov/marc/countries/)
- **oclc**: OCLC (WorldCat) Control Number (where available)
- **lccn**: Library of Congress Control Number (where available)

Our data will be useful to any literary historian or analyst interested in the significance of creative writing programs generally and the Iowa Writers’ Workshop specifically. For instance, someone studying literary networks could trace connections between and amongst writers, editors, and agents; someone studying the
significance of region and place could follow authorial itineraries in relation to literary subject matter; or someone studying literary influence could examine relations between mentors and mentees. Many other uses would be possible, especially with other data sets available for comparison.

Our IRB board has determined that none of our information violates the privacy or confidentiality of anybody listed in the database, although we are willing to partially or totally remove information about anyone upon their request. Other than gender, we have not tagged any of our entries with demographic data that might be considered sensitive or subjective.