Bryn Mawr College
Scholarship, Research, and Creative Work at Bryn Mawr College

Women's History in the Digital World | Women's History in the Digital World 2013

Mar 23rd, 12:05 PM - 3:45 PM

Mining Hymns: Exploring Gendered Patterns in Religious Language

Jeri Wieringa
George Mason University

Follow this and additional works at: https://repository.brynmawr.edu/greenfield_conference
Let us know how access to this document benefits you.

Wieringa, Jeri, "Mining Hymns: Exploring Gendered Patterns in Religious Language" (2013). *Women's History in the Digital World*. 41.
https://repository.brynmawr.edu/greenfield_conference/papers/saturday/41

This paper is posted at Scholarship, Research, and Creative Work at Bryn Mawr College.
https://repository.brynmawr.edu/greenfield_conference/papers/saturday/41

For more information, please contact repository@brynmawr.edu.

Mining Hymns

Exploring Gendered Patterns in Religious Language

@jeriwieringa

19th Century Religious Language

Nearer, my God, to Thee
- Sarah Adams, 1841

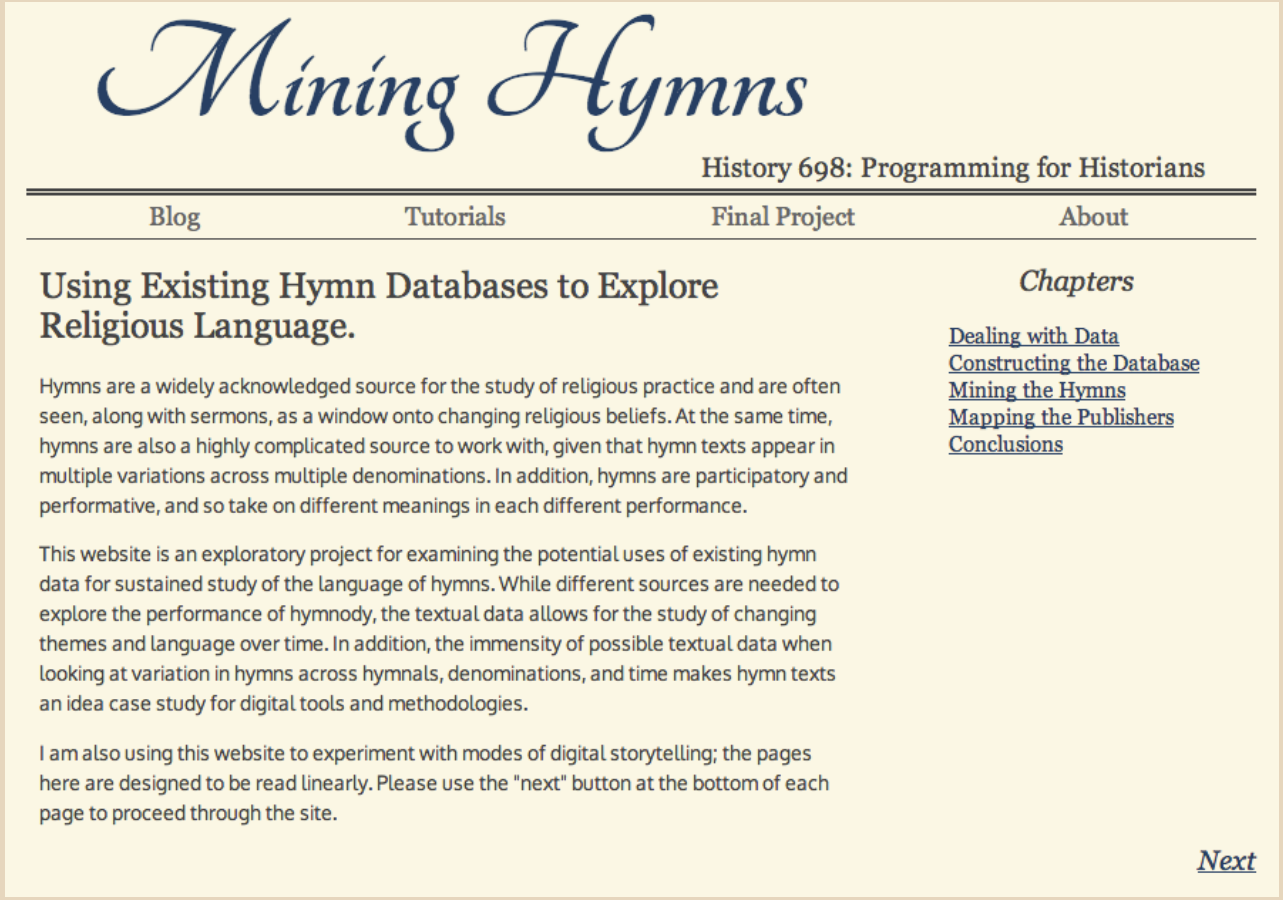
Nearer, my God, to thee, nearer to thee!
E'en though it be a cross that raiseth me,
still all my song shall be,
nearer, my God, to thee;
nearer, my God, to thee, nearer to thee!

Though like the wanderer, the sun gone down,
darkness be over me, my rest a stone;
yet in my dreams I'd be
nearer, my God, to thee;
nearer, my God, to thee, nearer to thee!

.

Crown Him with Many Crowns
- Matthew Bridges, 1851

Crown Him with many crowns,
The Lamb upon His throne;
Hark how the heav'nly anthem drowns
All music but its own.
Awake, my soul, and sing
Of Him who died for thee,
And hail Him as thy matchless King
Through all eternity.



http://jeriwieringa.com/dev/intro.php

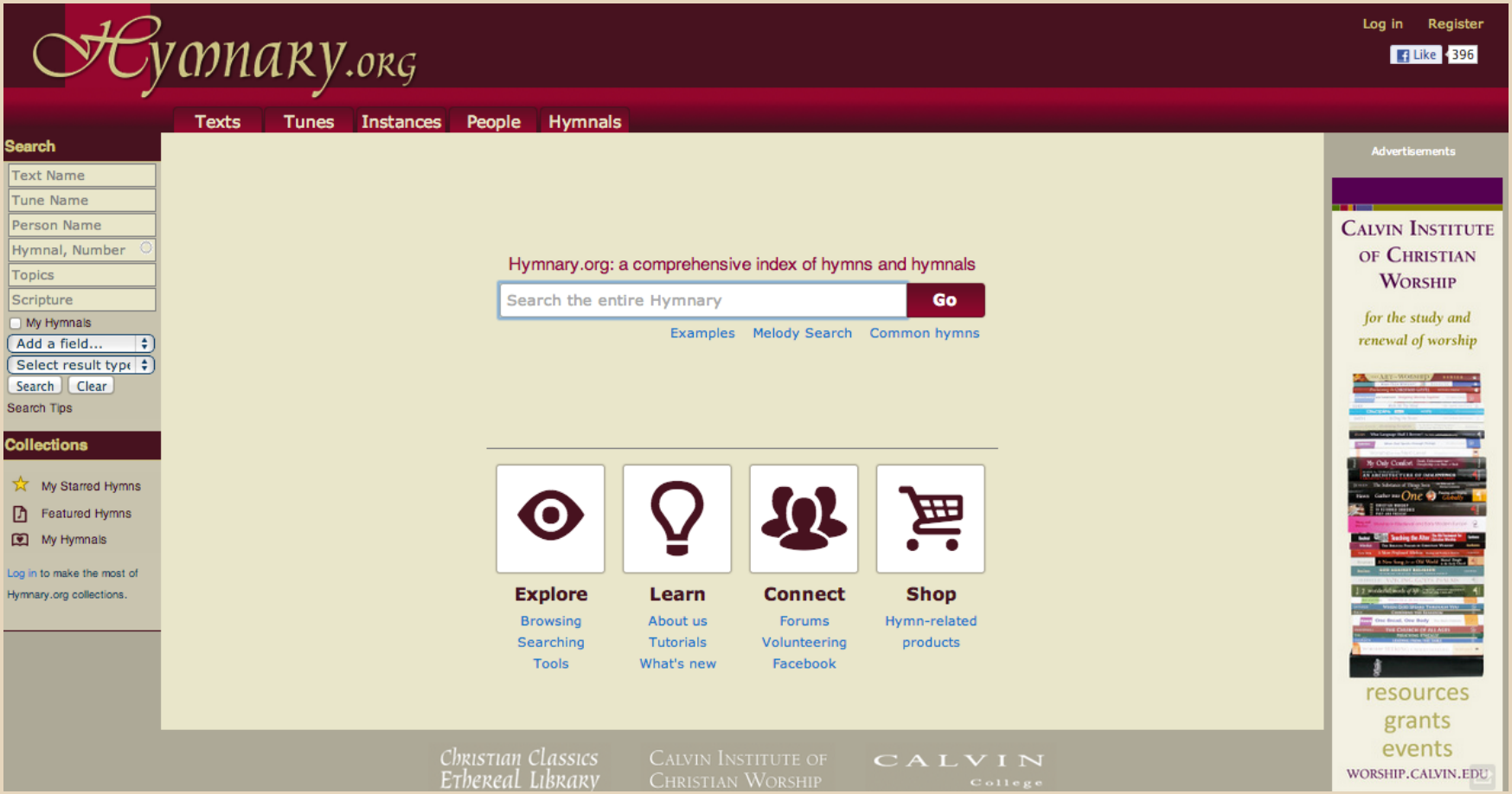
@jeriwieringa

Hist. 698: Digital Praxis

Syllabus:
http://fredgibbs.net/courses/digital-history-techne/

Course Blog with all of our tutorials:
http://www.fredgibbs.net/clio3workspace/blog/

Programming Historian:
http://programminghistorian.org/

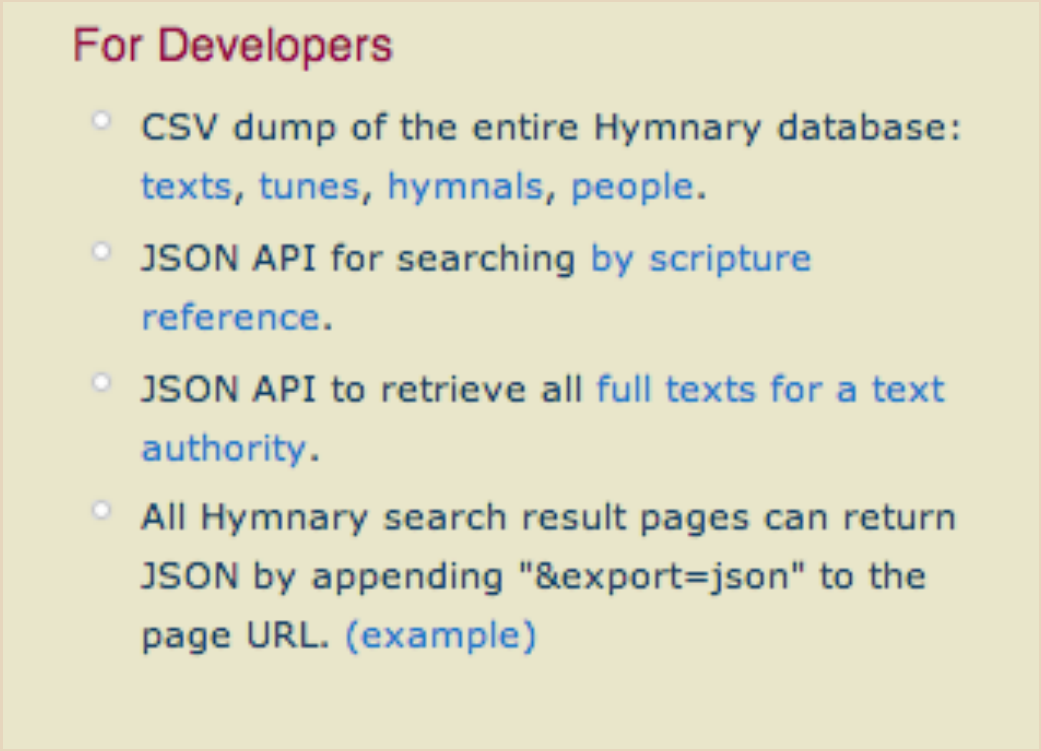


http://hymnary.org

@jeriwieringa

Sources

http://hymnary.org/explore

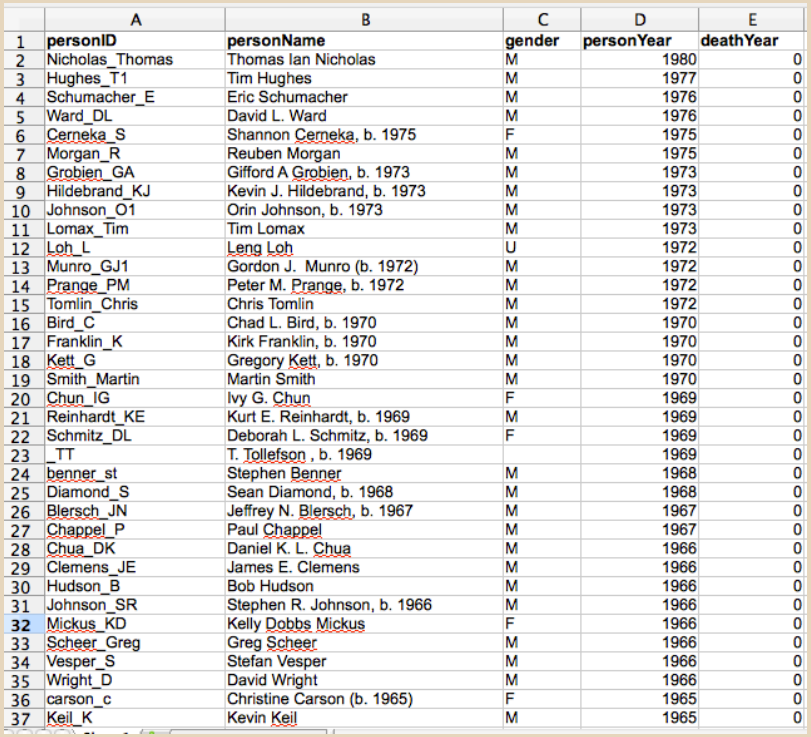


For Developers

- CSV dump of the entire Hymnary database: texts, tunes, hymnals, people.
- JSON API for searching by scripture reference.
- JSON API to retrieve all full texts for a text authority.
- All Hymnary search result pages can return JSON by appending "&export=json" to the page URL. (example)

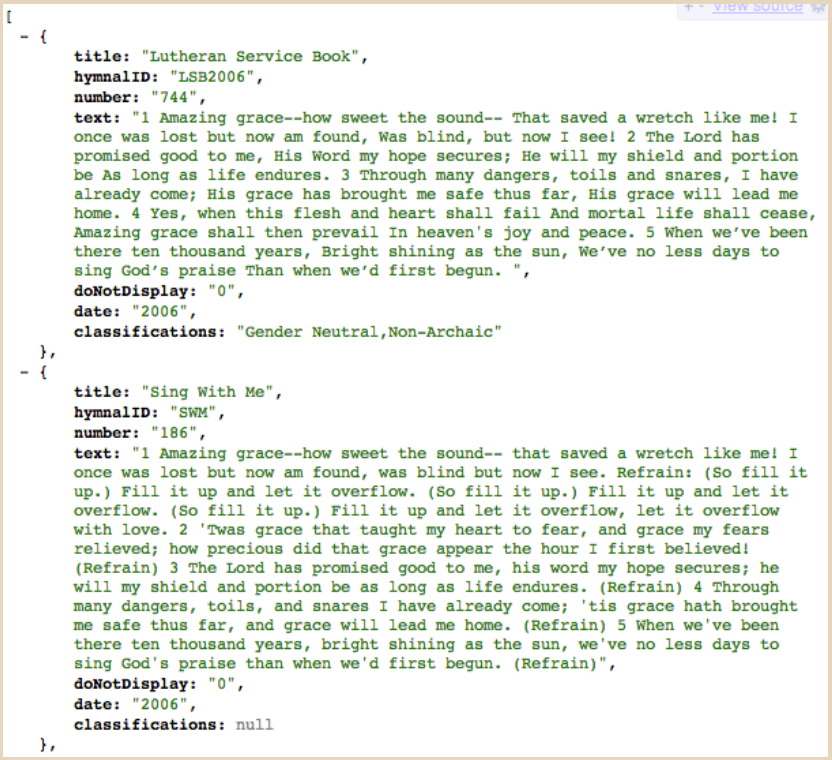
Sources

CSV Files



	A	B	C	D	E
1	personID	personName	gender	personYear	deathYear
2	Nicholas_Thomas	Thomas Ian Nicholas	M	1980	0
3	Hughes_T1	Tim Hughes	M	1977	0
4	Schumacher_E	Eric Schumacher	M	1976	0
5	Ward_DL	David L. Ward	M	1976	0
6	Cerneka_S	Shannon Cerneka, b. 1975	F	1975	0
7	Morgan_R	Reuben Morgan	M	1975	0
8	Grobien_GA	Gifford A Grobien, b. 1973	M	1973	0
9	Hildebrand_KJ	Kevin J. Hildebrand, b. 1973	M	1973	0
10	Johnson_O1	Orin Johnson, b. 1973	M	1973	0
11	Lomax_Tim	Tim Lomax	M	1973	0
12	Loh_L	Leng Loh	U	1972	0
13	Munro_GJ1	Gordon J. Munro (b. 1972)	M	1972	0
14	Prange_PM	Peter M. Prange, b. 1972	M	1972	0
15	Tomlin_Chris	Chris Tomlin	M	1972	0
16	Bird_C	Chad L. Bird, b. 1970	M	1970	0
17	Franklin_K	Kirk Franklin, b. 1970	M	1970	0
18	Kett_G	Gregory Kett, b. 1970	M	1970	0
19	Smith_Martin	Martin Smith	M	1970	0
20	Chun_IG	Ivy G. Chun	F	1969	0
21	Reinhardt_KE	Kurt E. Reinhardt, b. 1969	M	1969	0
22	Schmitz_DL	Deborah L. Schmitz, b. 1969	F	1969	0
23	_TT	T. Tollefson , b. 1969		1969	0
24	benner_st	Stephen Benner	M	1968	0
25	Diamond_S	Sean Diamond, b. 1968	M	1968	0
26	Blersch_JN	Jeffrey N. Blersch, b. 1967	M	1967	0
27	Chappel_P	Paul Chappel	M	1967	0
28	Chua_DK	Daniel K. L. Chua	M	1966	0
29	Clemens_JE	James E. Clemens	M	1966	0
30	Hudson_B	Bob Hudson	M	1966	0
31	Johnson_SR	Stephen R. Johnson, b. 1966	M	1966	0
32	Mickus_KD	Kelly Dobbs Mickus	F	1966	0
33	Scheer_Greg	Greg Scheer	M	1966	0
34	Vesper_S	Stefan Vesper	M	1966	0
35	Wright_D	David Wright	M	1966	0
36	carson_c	Christine Carson (b. 1965)	F	1965	0
37	Keil_K	Kevin Keil	M	1965	0

JSON Objects




```
[
 - {
      title: "Lutheran Service Book",
      hymnalID: "LSB2006",
      number: "744",
      text: "1 Amazing grace--how sweet the sound-- That saved a wretch like me! I
      once was lost but now am found, Was blind, but now I see! 2 The Lord has
      promised good to me, His Word my hope secures; He will my shield and portion
      be As long as life endures. 3 Through many dangers, toils and snares, I have
      already come; His grace has brought me safe thus far, His grace will lead me
      home. 4 Yes, when this flesh and heart shall fail And mortal life shall cease,
      Amazing grace shall then prevail In heaven's joy and peace. 5 When we've been
      there ten thousand years, Bright shining as the sun, We've no less days to
      sing God's praise Than when we'd first begun. ",
      doNotDisplay: "0",
      date: "2006",
      classifications: "Gender Neutral,Non-Archaic"
  },
 - {
      title: "Sing With Me",
      hymnalID: "SWM",
      number: "186",
      text: "1 Amazing grace--how sweet the sound-- that saved a wretch like me! I
      once was lost but now am found, was blind but now I see. Refrain: (So fill it
      up.) Fill it up and let it overflow. (So fill it up.) Fill it up and let it
      overflow. (So fill it up.) Fill it up and let it overflow, let it overflow
      with love. 2 'Twas grace that taught my heart to fear, and grace my fears
      relieved; how precious did that grace appear the hour I first believed!
      (Refrain) 3 The Lord has promised good to me, his word my hope secures; he
      will my shield and portion be as long as life endures. (Refrain) 4 Through
      many dangers, toils, and snares I have already come; 'tis grace hath brought
      me safe thus far, and grace will lead me home. (Refrain) 5 When we've been
      there ten thousand years, bright shining as the sun, we've no less days to
      sing God's praise than when we'd first begun. (Refrain)",
      doNotDisplay: "0",
      date: "2006",
      classifications: null
  },
```

Organizing the Data

- Normalize names between "People" and "Texts" CSV files.
- Isolate the Hymns associated with each selected Author.
- Pull down the 2000(+) JSON arrays and save as individual files.
- Filter out blank files.
- Split out each individual item in the array.
- Upload data in MySQL database.

All Records

Jump to the End

First Line	Title	Author	Gender	
Around the throne of God in heaven	Around the Throne of God in Heaven	Anne H. Shepherd	F	Details
The precious seed of weeping	The precious seed of weeping	Catherine Winkworth	F	Details
Ah Jesus, the merit	Ah Jesus, the merit	Catherine Winkworth	F	Details
Thou sore-oppress'd	Thou sore-oppress'd	Catherine Winkworth	F	Details
Of all the thoughts of God that are		Elizabeth Barrett Browning	F	Details
The only Son from heaven	The Only Son from Heaven	Elizabeth Cruciger	F	Details
O land of rest, for thee I sigh!	We'll Work Till Jesus Comes	Elizabeth Mills	F	Details
Oh what will be the day when won at last	Things to Come	Emma Frances Bevan	F	Details
I thirst, O, grant the waters pure		Frederica Bremer	F	Details
Cheek grow pale, but heart be vigorous		Frederica Bremer	F	Details
Lord Jesus, come; for here		Harriet Martineau	F	Details
Sweet is the prayer, whose holy stream		Harriet Martineau	F	Details
Father, who the light this day	Father, Who the Light This Day	J. A. Elliot	F	Details
There is no sorrow, Lord, too light	There is no sorrow, Lord, too light	Jane Crewdson	F	Details
Christ the Lord is ris'n today	Christ the Lord Is Risen Today; Alleluia	Jane Elizabeth Leeson	F	Details
Savior, teach me, day by day	Savior, Teach Me, Day by Day	Jane Elizabeth Leeson	F	Details
A little child may know		Jane Elizabeth Leeson	F	Details
Gracious Savior, gentle Shepherd	Gracious Savior, Gentle Shepherd	Jane Elizabeth Leeson	F	Details
How blessed, from the bonds of sin		Jane L. Borthwick	F	Details
Take the name of Jesus with you	Precious Name	Lydia Baxter	F	Details
Asleep in Jesus! blessed sleep!	Asleep in Jesus! blessed sleep!	Margaret Mackay	F	Details
Through the love of God our Savior	Through the Love of God Our Savior	Mary Peters	F	Details

http://jeriwieringa.com/dev/view.php

@jeriwieringa

Hymn: Around the throne of God in heaven

Title: Around the Throne of God in Heaven

First Line: Around the throne of God in heaven

Hymnal: The Hymnal: published by the Authority of the General Assembly of the Presbyterian Church in the U.S.A.

Denomination: Presbyterian Church (U.S.A.); Northern Presbyterian

Publisher: Presbyterian Board of Publication and Sabbath School Work

Publication Date: 1895

Author(s)

Author Name: Anne H. Shepherd

Born: 1809

Died: 1857

Gender: F

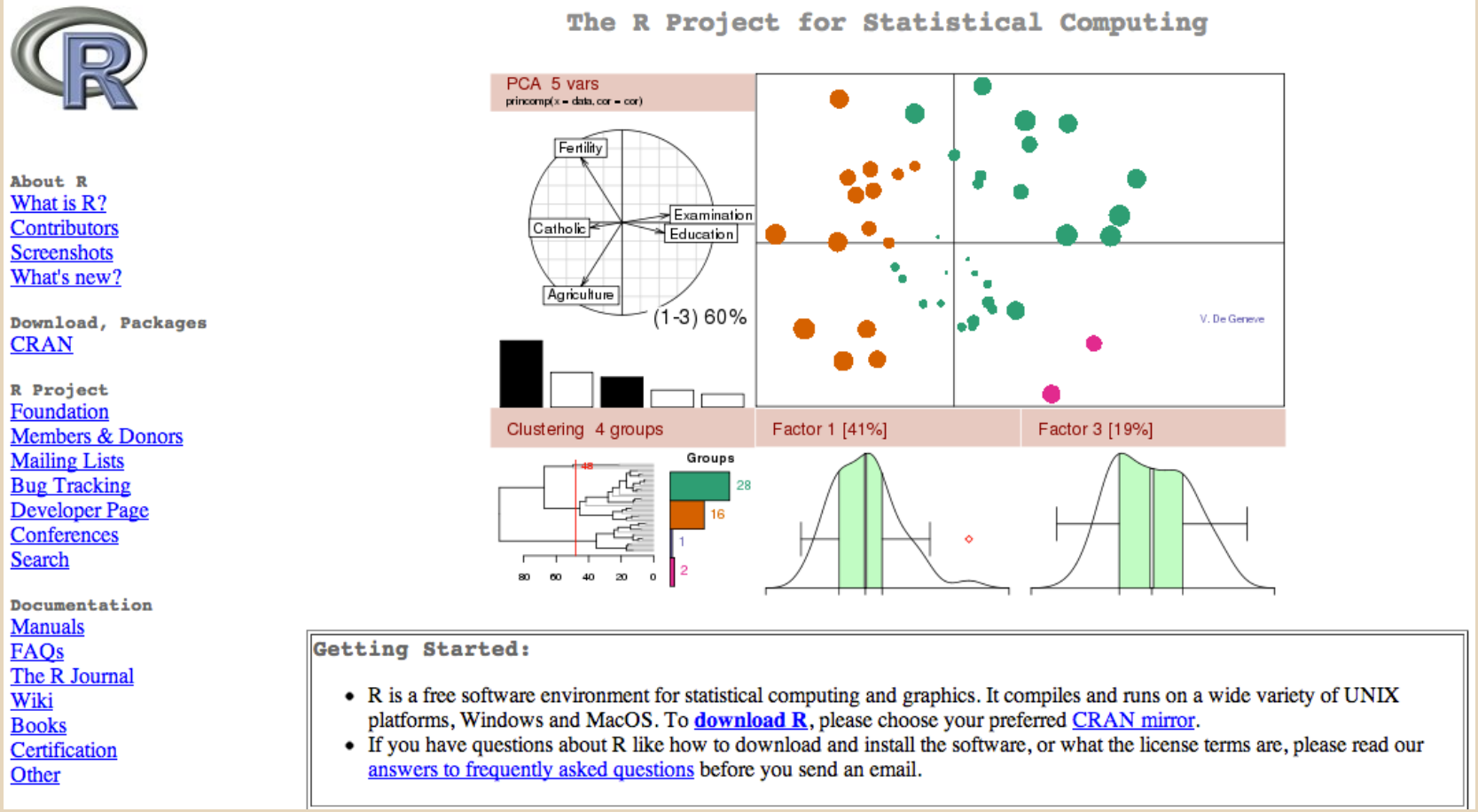
Full Text

1 Around the throne of God in heaven Thousands of children stand, Children whose sins are all forgiven, A holy, happy band, Singing, 'Glory, glory, Glory be to God on high.' 2 In flowing robes of spotless white See every one arrayed; Dwelling in everlasting light And joys that never fade, Singing, 'Glory, glory, Glory be to God on high.' 3 What brought them to that world above, That heaven so bright and fair, Where all is peace, and joy, and love; How came those children there, Singing, 'Glory, glory, Glory be to God on high?' 4 Because the Saviour shed His blood To wash away their sin; Bathed in that pure and precious flood, Behold them white and clean, Singing, 'Glory, glory, Glory be to God on high.' 5 On earth they sought the Saviour's grace, On earth they loved His Name; So now they see His blessed face, And stand before the Lamb, Singing, 'Glory, glory, Glory be to God on high.' Amen.

Return to All Records

http://jeriwieringa.com/dev/view.php

@jeriwieringa



http://www.r-project.org/

@jeriwieringa

Methods: tm Library

- Developed by Ingo Feinerer as part of his dissertation.

- Text Mining requires a number of "preprocessing" steps. tm enables these to be done quickly and efficiently.

citation: Ingo Feinerer, Kurt Hornik, David Meyer. "Text Mining Infrastructure in R" in *Journal of Statistical Software*, 25.5 (March, 2008) http://www.jstatsoft.org/v25/i05/.

Method: MALLET

- MALLET (MAchine Learning for LanguagE Toolkit) is a package that uses Java (and a variety of statistical algorithms) to generate topic models.

- Tutorial on using MALLET available at the Programming Historian. (http://programminghistorian.org/lessons/topic-modeling-and-mallet)

citation: McCallum, Andrew Kachites. "MALLET: A Machine Learning for Language Toolkit." http://mallet.cs.umass.edu. 2002.

Gendered Language: Frequent Words

Words in Common

HYMNS BY FEMALE AUTHORS

"alleluia" "christ" "day" "earth" "glory" "god" "hath" "heart" "heavn" "holy" "home" "hope" "jesus" "joy" "life" "light" "lord" "love" "name" "nearer" "oer" "peace" "precious" "rest" "sweet"

HYMNS BY MALE AUTHORS

"art" "bright" "christ" "day" "death" "earth" "father" "glory" "god" "grace" "heart" "heaven" "holy" "jesus" "king" "life" "light" "lord" "love" "mind" "peace" "praise" "sing" "son" "soul" "spirit" "word" "world"

Gendered Language: Frequent Words

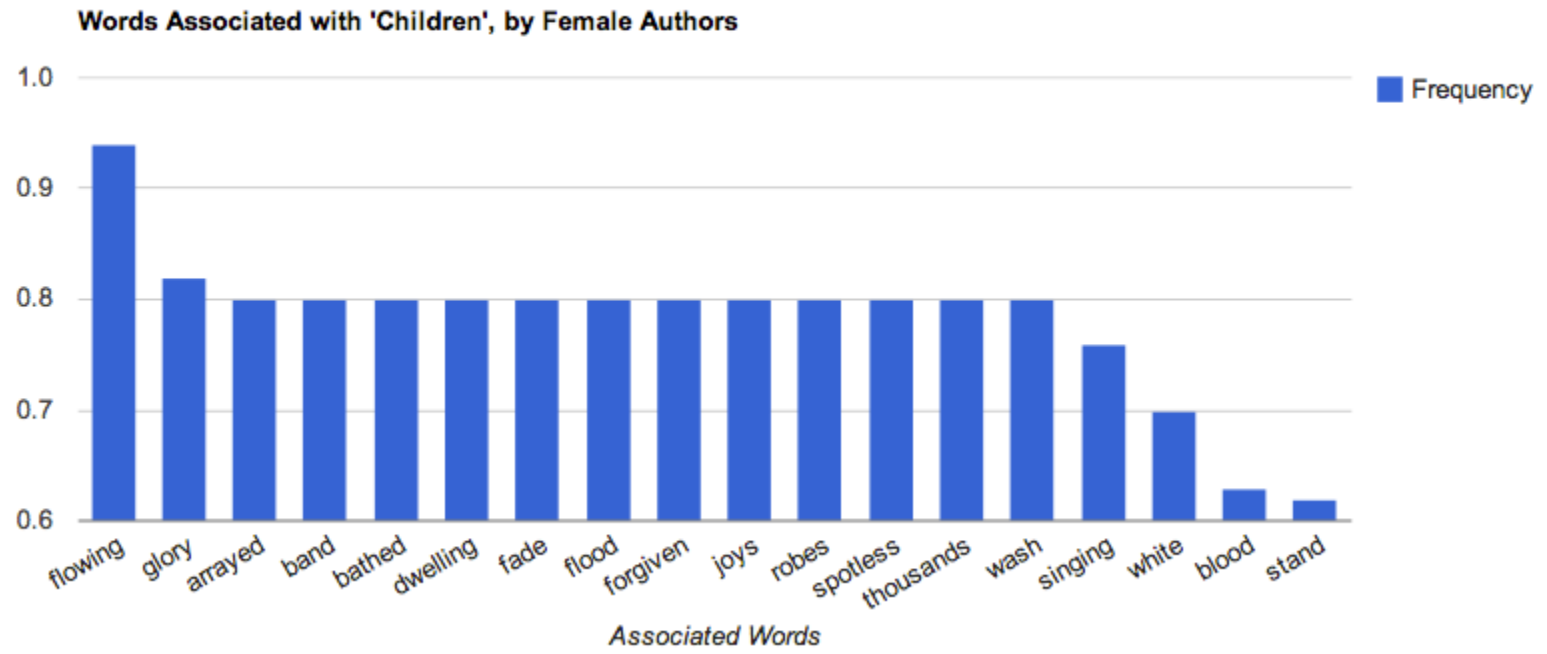
Variations

HYMNS BY FEMALE AUTHORS

"alleluia" "christ" "day" "earth" "glory" "god" "hath" "heart" "heavn" "holy" "home" "hope" "jesus" "joy" "life" "light" "lord" "love" "name" "nearer" "oer" "peace" "precious" "rest" "sweet"

HYMNS BY MALE AUTHORS

"art" "bright" "christ" "day" "death" "earth" "father" "glory" "god" "grace" "heart" "heaven" "holy" "jesus" "king" "life" "light" "lord" "love" "mind" "peace" "praise" "sing" "son" "soul" "spirit" "word" "world"

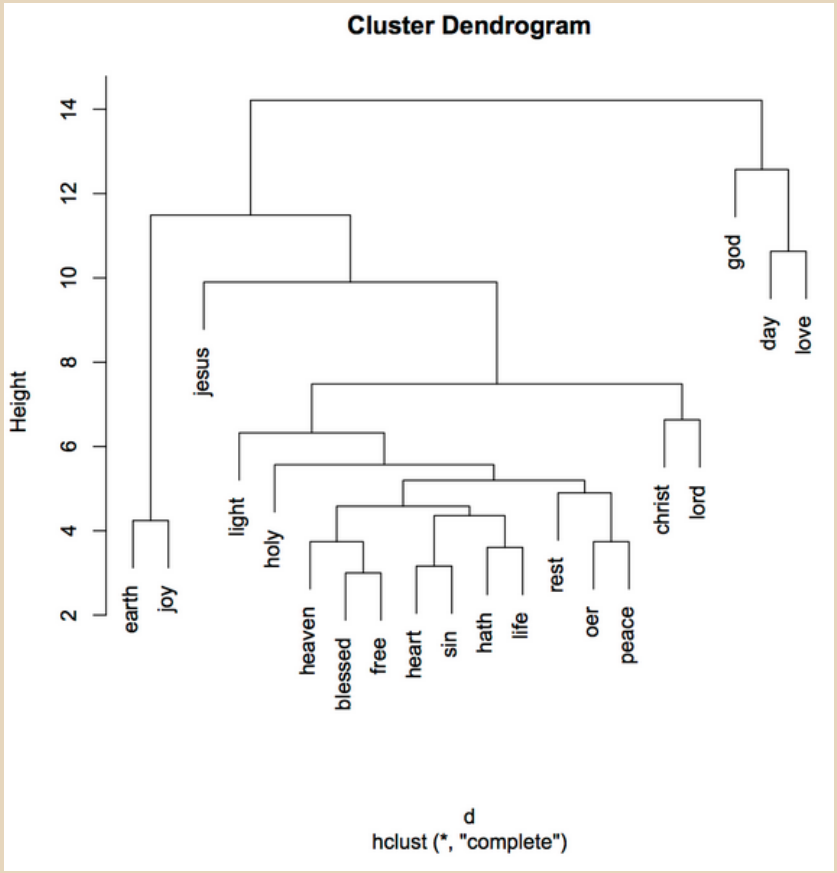


Word Associations: "Children"

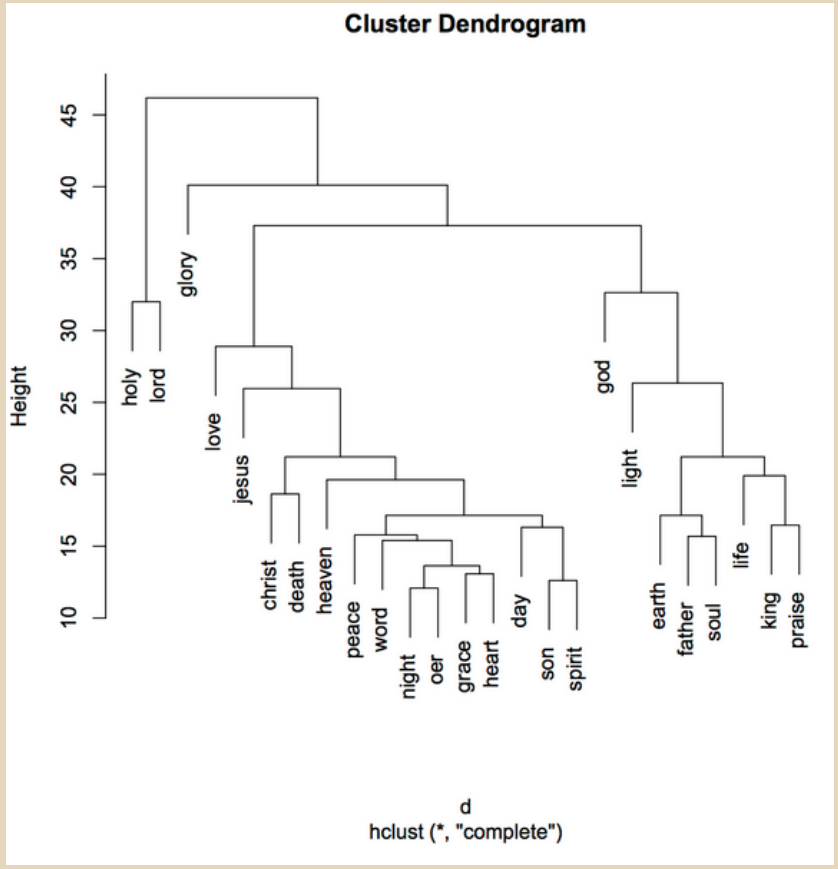
@jeriwieringa

Gendered Language: Across Hymns

by Female Authors



by Male Authors



Topic Modeling: Comparison

Female Authors

thy thou thee love lord home death sin blessed art light holy rest meet happy hast thine calm blest

joy earth precious heav sweet hope jesus star journey sorrow holy shining comfort er king feet ev father receive

er love peace life heavenly light part deep living hearts sweetly praise happy ah worship teach tis safely dead

glory god singing love behold high blood saviour sing face grace bright white heaven throne free dying lesson saints

ransomed full curse reap promise er time saileth silence wailing sees fainter deeper fight age fresh green men irradiate

nearer thee god sun song mercy moon sky wait praise forgot stars cleaving wing woes bethel griefs stony beckon

day gloom tomb christ angel spirit pray light make hand rough sea left grown mist followwe sweetest harps immortal

jesus asleep power place fear weep wakes bless high days hosannas bread vainest dayspring mark crossv laid gethsemane remembering

tears small solace wise gleam drink flowed washed gainsay lure flattering trust dust ere prompts seed goal struggles forever

children stream easy shepherd cleanse savior holy thine thy streamed join saints offerings sing lisp wear whate approve guide

Male Authors

thy lord god give thine hast word heart free grace blest spirit prayer make hear save soul throne turn

god holy father praise earth son king spirit heaven sing eternal bless world glorious raise christ wisdom born heavenly

lord love god amen peace sin mercy hath grace thee joy saviour make hearts high feet fear cease fill

light day night dark bright er word love truth rest voice true shine thought living live lead darkness souls

thou thee art thine light thy life didst divine power pure dost great strength thyself earthly glory day st

death christ life cross jesus die bear sin man great flesh faith died grave blood dear foes glorious victory

life earth world heart er good vain deep children men bound strife things pain flowing find ill sweet grief

love heaven home father faith hand woe place pain grief sight loved fears voice human twas death sheep fold

glory heavenly crown high lamb saints angels eternity people king power slain church throne sun kings voices praises sky

jesus soul mine blood sins thousand precious prepare blessed wash guilt lay freely savior comfort load treasure heart delight

Potential Vectors for Further Analysis

- Popularity of texts over time and tradition.
- Major topics over time and by tradition.
- Word Associations over time.

Challenges of Text-Mining Hymns

- Digitization of the sources is incomplete.
- Authorship is often hard to determine.
- Validity of results depends on the quality of the data.

Potential Results of Text-Mining Hymns

- See patterns over a wide range of texts.
- Move beyond the most popular hymn writers to see broader cultural patterns.
- Expose the variety in religious culture that is often glossed over.
- Improve our attribution of sources by identifying "style" of female authors.

Jeri Wieringa

March 23, 2013

Presentation for the Women's History in the Digital World Conference at Bryn Mawr

[Slide 1]

Mining Hymns: Gendered Patterns in Religious Language

What can we learn about religious culture through the study of hymns? Are there different patterns in male and female piety that the study of hymns can help us identify?

[Slide 2]

Take for example these two hymns from the middle of the nineteenth - century. The first, written by Sarah Adam, stresses the relationship between the singer and God, asking to be drawn closer, despite the suffering currently experienced or even the suffering required for such closeness to occur. The second, by Matthew Bridges, is a powerful text that draws on the language of Revelation and emphasizes the power and majesty of a God who is never directly named but is referred to as "His", as "Lamb", and as "King".

[slide 3]

"Mining Hymns" is an exploratory project to discover whether text-mining can be a useful tool for studying religious culture and for uncovering broader patterns in religious language. In framing the project, my hypothesis was that text-mining offers a way to handle the overwhelming quantity of information involved in studying hymns. Many hymns appear in multiple hymnals from many traditions, at times changing across the versions. Tracking all of these changes at scale manually is impossible. Text-mining offers a way to track these changes and to see patterns across an abundance of texts over time.

[slide 4]

This project comes out of a course on digital methods at George Mason University. For those interested in learning or teaching text-analysis or data manipulation, our course syllabus has links to many of the resources we relied on and our course blog houses the tutorials we created over the course of the semester. Also, the Programming Historian offers a number of very helpful introductory and intermediate tutorials, some of which I relied on to do the analysis I am presenting here.

[slide 5]

My hypothesis was that hymns offer a rich source of cultural data and that text-mining could help reveal useful patterns across collections of hymns. My area of study is the 19th century, so I was particularly interested in seeing what sort of patterns of religious piety apply text-mining to 19th century hymns could highlight.

In order to test this hypothesis, I needed a collection of hymn data to analyze.

Enter hymnary.org. Hymnary is an open database of North American hymnody and is a combination of the Dictionary of North American Hymnology and the efforts of the Hymn Society of the United States and Canada. As this is the self-proclaimed "most complete database" of hymns "on the planet", I wanted to evaluate the usefulness of the database for gender focused text-mining.

[slide 6]

An additional reason to use the Hymnary database is that they make the data relatively easily accessible for people interested in research. Between the CSV files of the data and the JSON arrays of the text, the database offered much promise of data ready to be used for analysis.

[slide 7]

This is the format of the data as Hymnary presents it. On the left you are looking at the “People” file and on the right is the first two entries of the JSON array for “Amazing Grace.”

As you may notice, the names in the “People” file are not entirely consistent. Some include birth year with the name, some middle initials, some not. Similar inconsistencies also appeared in the “Texts” file, which included the author names, not the person ID, and, of course, the inconsistencies were not themselves consistent across the files. This resulted in a lot of work using Google Refine (now Open Refine) and a fair amount of pattern noticing on my part to match up the people with all of the relevant hymns.

[slide 8]

My steps in gathering, normalizing, and storing the data included normalizing the names across the CSV files, identifying the hymns associated with the authors I chose (and for this project I took those authors born between 1800 and 1809 as they were most likely to be productive during the middle of the century and seemed as good of a criteria as any for a broad cross section of the available hymns), downloading the JSON arrays, cleaning up the empty files, splitting out all the individual instances to locate the oldest version available (to limit complexity for this experiment, I only used one version of the text - the oldest texts available from the database, which sadly wasn’t necessarily the 19th century version), and entering all of the data into a MySQL database.

During this process I moved from a collection of 175 authors and over 2100 hymns to a set of 100 authors and approximately 275 hymns with full texts.

After much sweat and a few tears, I had both a database and an interface that allows you to see the sources that underly the textual analysis.

[slide 9]

The full list is available on my website and can be sorted by the different columns.

[slide 10]

Selecting details will open the page I designed for displaying author information along with the hymnal that the text I analyzed is from and the full text itself.

Again, I choose these hymns based on the birth year of the author, rather than on denomination or on hymnal, because I am looking for broad patterns across traditions. As a result, the hymns surveyed come from a range of Protestant and Catholic hymnals. However, the hymnals transcribed do reflect the preferences of hymnary.org's audience and volunteers, who seem to favor mainline or established denominations and popular hymns.

[slide 11]

There are a number of computational tools that I used to analyze these texts. The first is R, which is a program for statistical analysis, as well as …

[slide 12]

The tm (or text-mining) library, which is downloaded and run through R for doing statistical analysis on texts. tm makes it much easier to go through a number of the preprocessing steps that text-mining requires in addition to generating common statistical reports.

The preprocessing steps that I performed on these texts included: removing punctuation, numbers, white spaces, converting to lower case and removing stop words using the standard list, as well as a custom list with words particular common in hymns, such as "Amen" and "Refrain."

[slide 13]

The final tool I used was MALLET, which is one of the more popular tools for topic modeling in the humanities. To guide my process with MALLET, I used the tutorial published on the

Programming Historian. MALLET has a lot of moving parts and until I understand better what those pieces are doing, my results here are strictly exploratory.

[slide 14]

The step of text-mining analysis I performed on my hymns was to split the hymns into sets based on the gender of the author and used R to determine the most frequently used words in each set. This slide shows the most frequently used words from both the female and male authors, with the words appearing in both list highlighted in blue. These words are fairly common in Christian religious language, and reflect a focus on conversion, heaven, and God.

[slide 15]

Looking at the words that are different, we can start to see an interesting pattern. Where the female authors frequently use language that is intimate, nurturing, and warm, the male authors use words that are formal, abstract, emphasize patriarchy and sound "theological" in tone.

"Feminization" is a term often used to describe nineteenth-century religion. It is a term tied to the concept of the cult of domesticity and also to the idea of the "sentimentalization" of religion. These words that female authors used frequently fit into that description, as they emphasize the home and use nurturing and sentimental adjectives such as "precious" and "sweet."

By contrast, the male authors, favor a very different depiction of piety. Emphasizing masculine concepts of God in "father" and "king" as well as more abstract aspects of the human person, such as the "soul" and "spirit", the religious language of the male authors reflects traditional theology much more than the language of the female authors. This is not surprising, given the near male monopoly on formal seminary training during the early 19th century. The differences point to variations in religious culture by gender and the need to investigate further concepts like feminization to allow for gendered variations in religious culture.

[slide 16]

None of the female authors directly refer to women in their hymns, but one place where I thought more direct reflections of gender could be teased out was in the use of "children." This word is not among the most frequently used by my female authors, and so I am not certain how popular of a subject it was in their hymns. However, the words associated with "children" reveal some of the unexpected patterns uncovered by doing text-mining research for gendered language.

With the pattern of intimate language revealed in the frequent words, and the framework of feminization and domesticity that is often associated with it, it would be reasonable to expect some attention to the piety of children among the female authors. What I found in looking at the words associated with "Children" was a concern with redemption and the cleansing of sins. Reading back into the hymns themselves, I found that in these descriptions, the women authors were often referring to themselves and their having been "washed" in the "blood" and "arrayed" in "robes" of "white". An association of children with salvation could be interpreted in terms of concern with the salvation of children. But more often, these women were describing themselves as "children" - children of God through this process of redemption. The motif of the child is not uncommon in Christian literature, but the pattern of women describing themselves as children calls for additional analysis as to prevalence over time and tradition, as well as a comparison study of its use among men.

[slide 17]

Another way of visualizing the relationship of words across a variety of texts is through the Dendrogram. What is being visualized here is the similarity between the words: the closer the words are to one another, the more "similar" the terms. The longer the "branch", the less similar the term is from the ones it is connected to. Also, the order from left to right in not significant apart from the connection to other branches - think of a hanging mobile (hat-tip to Wheaton College's tutorial on reading dendrograms)

Many of these clusterings are expected, which is good. From the women, we see "Christ" and "Lord" appearing together, from the men, "holy" and "Lord". Some of the clusterings, though,

are curious and spark questions. One such pairing is that of “earth” and “joy” by the female authors, a pairing that is quite dissimilar from the others as evidenced by the length of the vertical line. To investigate further, the uses of “earth” and “joy” would need to investigated in their contexts to see why the pairing and what sort of significance to assign it.

What draws my attention from the male authors is the split between the two clusters of words, the one on the right focuses on “God the Father” while the language of the one on the left gravitates with the language on the left focusing on the Son. Theology embedded in hymns revealed right before your eyes.

[slide 18]

Finally, a little topic modeling experiment using MALLET. What I am displaying here are the topics most heavily weighted across all of the texts. Another way to look at the results is to see which of the topics appear in each text, which is helpful for identifying the theme the topics are capturing. I used 25 topics to generate the models - I experimented with 50, but there were too many outliners when working with a corpus that is rather repetitive in its subject matter.

*** Note: After a number of most helpful conversations with Michelle Maurice, Cameron Blevins, and Bridget Baird, and seeing a few other well formed topics, these seem less coherent than I initially thought. How best to apply topic modeling to hymns that are poetic and form and repetitive in content is a puzzle that I am still working out. ***

Interpreting these results is something that I am still wrestling with and it something that the community as a whole is working develop best practices around. A couple of trends that I want to point out:

1. You can see the same difference in language in these topics that we noticed in the frequent words. The topics from the hymns by female authors reflect an intimate faith while the topics from the male authors reflect a patriarchal faith.

2. [Topic underlined in purple] some topics are very coherent and reflect well common themes found in hymns.

3. Some of the challenges of working with hymns (such as the split words) are apparent in the topics, though they also appear to have been well incorporated with their separated halves.

[slide 19]

These initial results did support my initial hypothesis that text mining can help uncover gendered patterns in religious language. In looking through the results, further research is needed to get a better sense of what the historical implications of these themes and trends might be. Some "vectors" to pursue are:

1. Popularity of texts over time and tradition, shown, in part, by which traditions include them in their hymnals and how many hymnals include them over time.

2. Topics over time and by tradition, which will involve tracking more metadata along with the texts .

3. Word associations over time to see if there are changes in the connotations in the religious language over time.

[slide 20]

This is not to say that applying text-mining to hymns doesn't comes with a number of challenges.

1. For one, the digitization of sources is incomplete. My dataset was much smaller than anticipated because of the preferences of those volunteering to transcribe hymnals for hymnary.org. Less popular hymns and hymnals from smaller denominations are under-represented.

2. Authorship is often hard to determine or is multiple. I divided the hymns based on knowing the gender of the author - this is often not possible.

3. Validity of the results depends on the quality of the data. This shifts the core of the work to the selection and preparation of the data to be analyzed. In terms of assessing text-mining projects, the majority of the attention must be directed at the data.

[slide 21]

That being said, this experiment in text-mining has made clear to me that benefits to be gained from text-mining outweigh these challenges. As I see them, these benefits include:

1. The ability to see patterns over a wide range of texts, which enables the researcher to either verify or problematize existing interpretations, such as the "feminization" of religion.

2. It becomes easier to move beyond the popular hymn writers, to move beyond Watts, Wesley and Julia Ward Howe, to see broader cultural patterns and changes in those patterns over time.

3. It also helps to expose the variety within religious culture that is often glossed over in the construction of a historical narrative. Along the same line, it helps expose the ways different religious movements are connected, sharing hymns, changing hymns, and in doing so, creating cultures that are in conversation with one another.

4. And, there is the potential for improving our ability to identify additional pieces by female authors. By identifying a set of characteristics typical of female authors, we may be able to predict the likelihood of a particular text having been written by a female author. For historical hymns, this may help us compensate for the problem of authorship noted previously.

Taken together, these and the ability to track changes in language over time, all support the claim that text-mining offers a useful lens for analyzing the construction of gender within religious communities and the contributions of gender to constructions of religious culture.

Thank you.