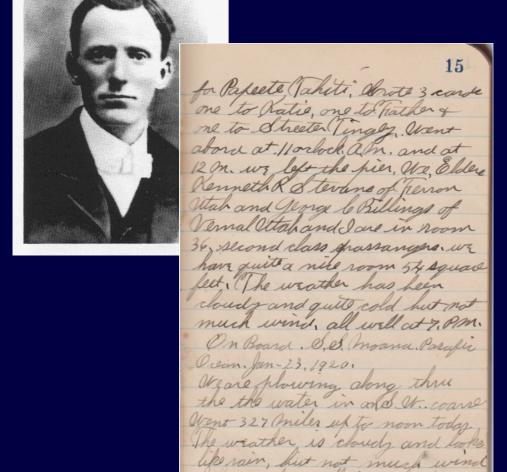
# Record Linkage and Tagging for the BYU Historic Journals Project (journals.byu.edu)

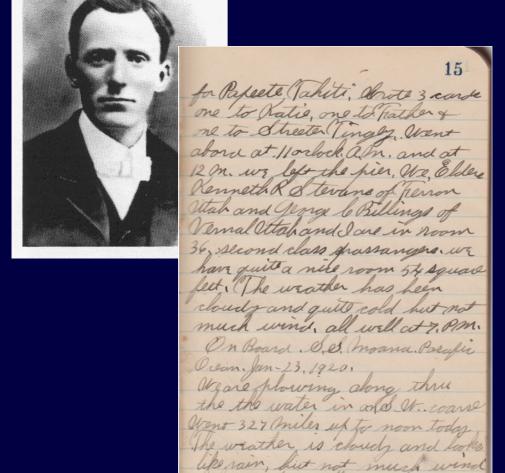
Douglas J. Kennard and Dr. William A. Barrett (BYU Computer Science Department)



keep from being sick to day but am feeling better tonight.

Staged in room most of day

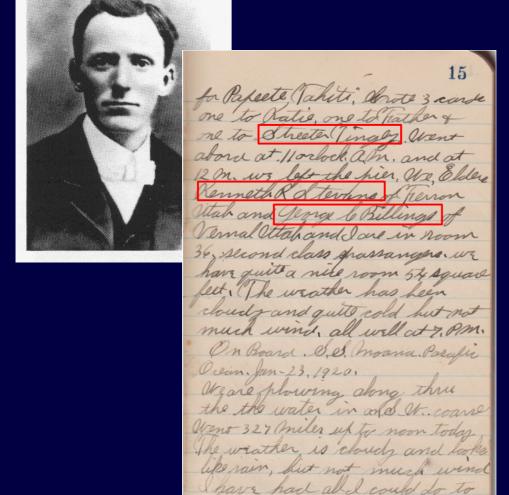
- How would you know?
- Where is it?
- What does it say?



keep from being sick to day but am feeling better tonight.

- Who might care?

900 living descendants! (my great-grandfather)



keep from being sick to day but am feeling better tonight. Staged in room most of day - Who might care?

900 living descendants! (my great-grandfather)

Descendants of people he wrote <u>about!</u>

### My questions:

Which of my ancestors wrote a journal?

Did anyone else write about my ancestors?

How can we share (within limits) diaries?

## journals.byu.edu











Http://journals.byu.edu/











## HISTORIC JOURNALS

- EXPERIENCE THE PAST -

BETA

### (Not logged in) Log In / Create Account

Journals
People & Connections
Rosters
Collaborators (friends)

Messages News Notifications Synchronize Ancestors Synchronize Rosters Your Ancestors Your Profile

Main Page Related Links Tutorials



Introduction Video (click to view)

#### Get to know your Ancestors

Your ancestors are more than just names, dates, and places. Find out who they really are - how they lived, what they thought, and how they felt - by reading their own personal diaries, letters, and biographies. Read what their friends, relatives, and acquaintances wrote about them, too. Do it quickly and easily here on the Historic Journals website for free. Help other people do the same thing.

#### Discover forgotten History

Much of what we know about historical people comes from what other people wrote about them. Our approach to linking journal entries (and other materials) to who they are about makes it easier than ever to discover even the most obscure references to the people you are interested in.

Click here to find out how to create an account

#### Conditions of Use | Privacy Policy | Contact Us

## journals.byu.edu











🖶 http://journals.byu.edu/











## HISTORIC JOURNALS

- EXPERIENCE THE PAST -



### (Not logged in) Log In / Create Account

Journals
People & Connections
Rosters
Collaborators (friends)

Messages News Notifications Synchronize Ancestors Synchronize Rosters Your Ancestors Your Profile

Main Page Related Links Tutorials



Introduction Video (click to view)

#### Get to know your Ancestors

Your ancestors are more than just names, dates, and places. Find out who they really are - how they lived, what they thought, and how they felt - by reading their own personal diaries, letters, and biographies. Read what their friends, relatives, and acquaintances wrote about them, too. Do it quickly and easily here on the Historic Journals website for free. Help other people do the same thing.

#### Discover forgotten History

Much of what we know about historical people comes from what other people wrote about them. Our approach to linking journal entries (and other materials) to who they are about makes it easier than ever to discover even the most obscure references to the people you are interested in.

Click here to find out how to create an account

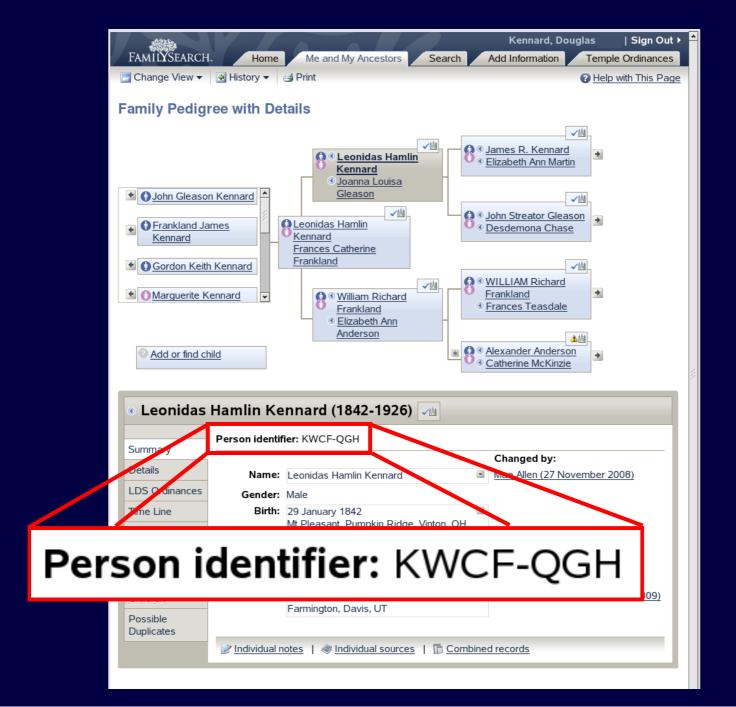
#### Conditions of Use | Privacy Policy | Contact Us

## **Previously Described Details**

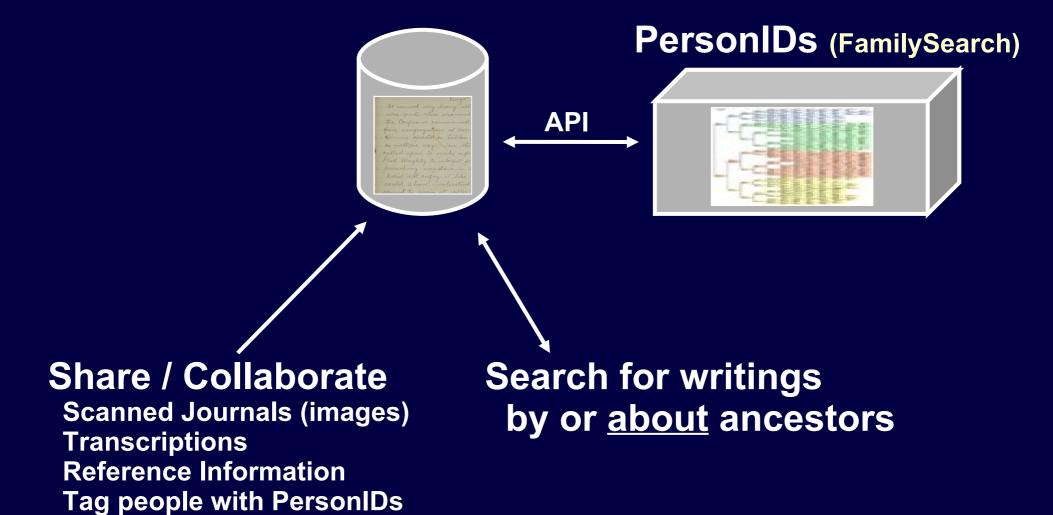
JCDL 2009 (Joint Conf. on Digital Libraries)

FHTW 2009 (fht.byu.edu)

## new.familysearch.org

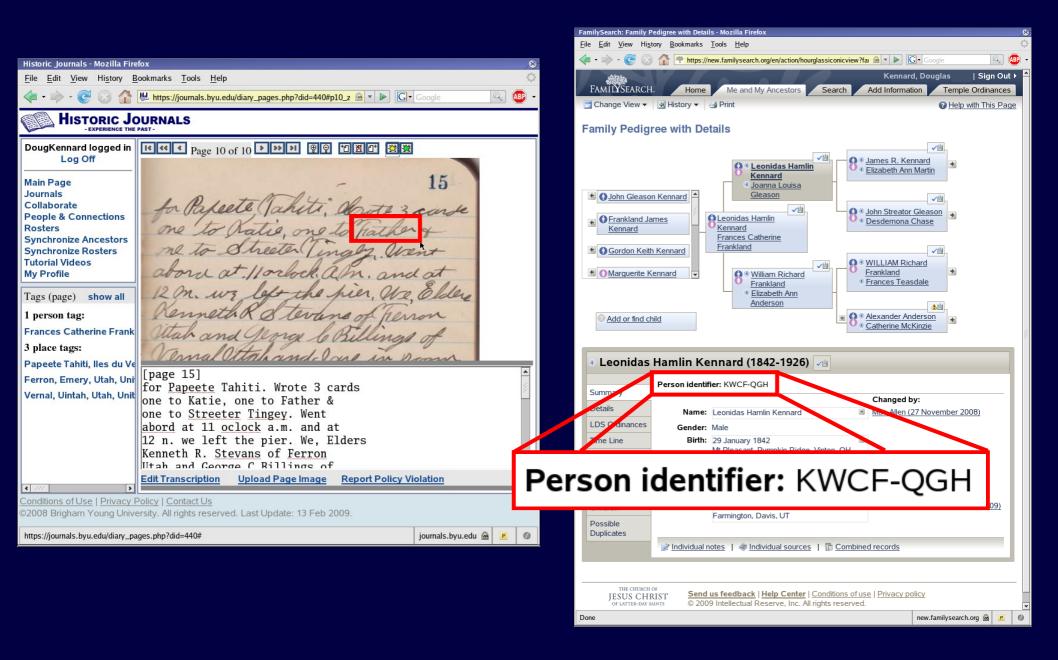


## **BYU Historic Journals**

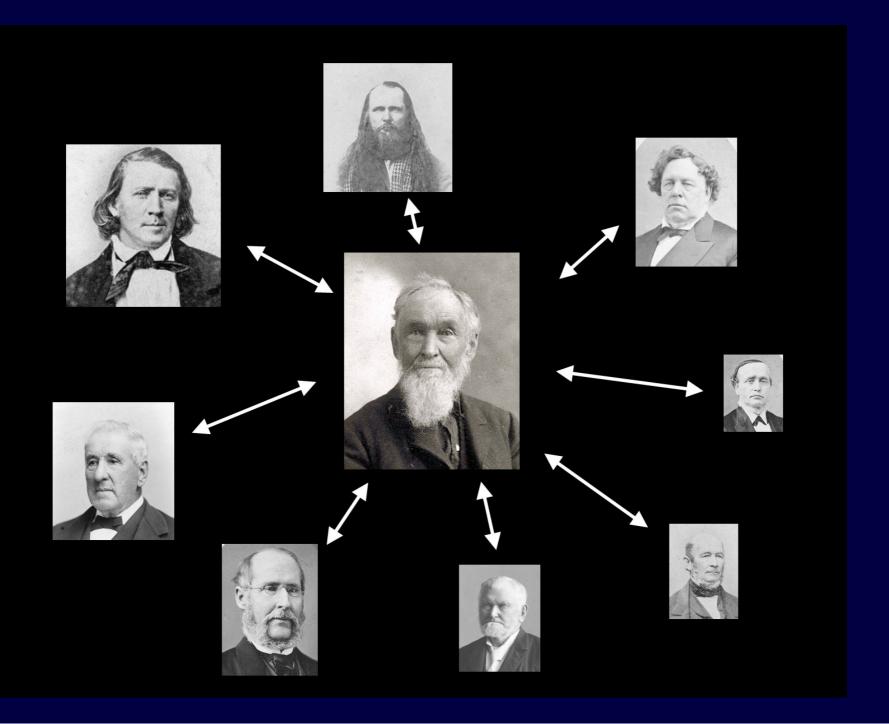


## Tagging who is written about

(similar to tagging on social networks, but PIDs)



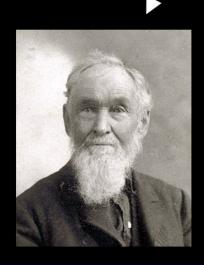
## **Historical Social Network**



## Rosters (implicit connections)



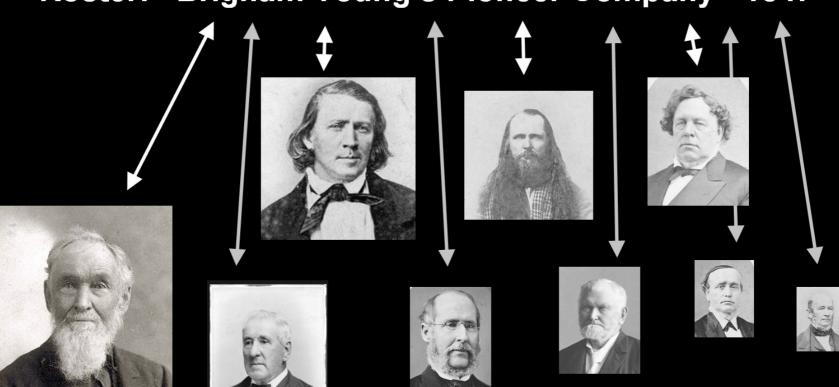
**Roster: Brigham Young's Pioneer Company - 1847** 



## Rosters (implicit connections)



Roster: Brigham Young's Pioneer Company - 1847



## Rosters (implicit connections)

**Military** (ex: Captain Stout's Army, Revolutionary War)

**Community** (ex: Bastrop, TX, USA)

Church (ex: Bannockburn Baptist Church, Snowville, AK)

**Team** (ex: 1980 US Olympic Hockey Team)

Class (ex: Davis High School, Mr. Smith English class, Fall 1983)

Work (ex: Austin, TX, Joe's Grocery Store, 1980-1983)

Other (ex: Brigham Young's pioneer company, 1847)

## Record Linkage

Manual Tagging / Crowd-Sourcing

**Davis Bitton "Guide to Mormon Diaries..."** 

**BYU HBLL - Overland Trails** 

**Mormon Missionary Diaries** 

## Automatic / Semi-automatic Record Linkage

F. Esshom "Pioneers and Prominent Men of Utah" (photos)

**Ids.org** "Mormon Overland Travel" pioneer DB

## Guide to Mormon Diaries and Autobiographies (Manual Tagging)

to emigrate, 1852.891. Gill, David R.

Diaries (1887–1889) Diaries (1887–1889)

Holograph. 6 vols. 16-18 cm. HDC [LDS Church Archives] (Ms d 1486)

British Mission, 1887-89.

#### 892. Ginn, John L, captain Autobiography (1857–1858)

Autobiography (1857–1858) Autobiography (1857–1858)

Photocopy of typescript. 147 pp. Princeton [Firestone Library, Princeton U]. Typescript. 184 pp. UHi [Ut St Hist Soc] (Mun A77)

"Personal recollections." Departure from Independence, Missouri, 1887. Member of Magraw's Pacific Wagon Road Expedition. Joined Tenth Infantry at Fort Laramie. Amusing account of Momon raid on mules. Resigned and thew self on mercy of Mormons. To Salt Lake. Rocco with "Mother Taylor." View of Mormonism by a scoffer and unbeliever. Aik en parts. Complaints about polygamy. To California, where author spread charges against stormons. Details of linvestigations.

Not all first-hand observation. One section goes back to tell the assory of Mormonism, see its fraud. Written about 1903. Comment by Charles Kelly in 199 suggests that the accord while generally reliable, drew upon "an old man's imagin dron." Repetition. Utah War is gone over a second time, with quotations from govern new reports. Details on Mormon-11 and cooperage.

#### 893. Gleason, John Streator

Diaries (1857-1863)

Diaries (1857-1863)

Microfilm of Holograph. 162 pp. HDC [LDS Church Apraives] (Ms f 182 ftem 1)

Mission in New York State, April 1887 to July 15. 8. Visited home. Montoe, New York Buptized his eighty-one-year-old futher. Vision sister in Ohio Mother sister in Michigan Return to Utuh, 1888. Mission to England 1890–63.

#### 894. Gledhill, Sarah S. Moulding, 1858-

Autobiography (1858-c.1890)

Autobiography (1858-c.1890)

Typescript. 12 pp. UHi [Ut St Hist Soc] Offin A581); USU [Merrill Libr, Ut State] (Mormon Diaries bx 2)

Rambling recollections of chik® s.d., except for last paragraph. Born at Philadelphia, 1858. To Salt Lake, 1865. Description recovered wagon. Settled in Ogden. Living conditions. Married Adam Gledhill, 1888.

895. Glines, James H., 1822–1905 Autobiography (1822–1904) Autobiography (1822–1904) Reference book of 2,894 known diaries

Alphabetical (last name)

Where to find the diary

Synopsis of content / bio info

## Guide to Mormon Diaries and Autobiographies (Manual Tagging)

to emigrate, 1852.891. Gill, David R. Dia ri es (1887-1889)

Diaries (1887-1889)

Holograph. 6 vols. 16-18 cm. HDC [LDS Church Archives] (Ms d 1486)

British Mission, 1887-89.

#### 892. Ginn, John I., captain Autobiography (1857-1858)

Autobiography (1857-1858)

Photocopy of types cript. 147 pp. Princeton [Firestone Library, Princeton U]. Typescript. 184 pp. UHi [Ut St Hist Soc] (Man A77)

"Personal recollections." Departure from Independence, Missouri, 1857. Member of Magraw's Pacific Wag on Road Expedition. Joined Tenth Infantry at Fort Laramie. Amusing account of Mormon raid on mules. Resigned and threw self on mercy of Mormons. To Salt Lake. Roomed with "Mother Taylor." View of Mormonism by a scoffer and unbeliever. Aiken party. Complaints about polygamy. To California, where author spread charges against Mormons. Details of investigations.

Not all first-hand observation. One section goes back to tell the history of Mormonism, seen as a fraud. Written about 1903. Comment by Charles Kelly in 1959 suggests that the account, while generally reliable, drew upon "an old man's imagination." Repetition. Utah War is gone over a second time, with quotations from government reports. Details on Mormon-Indian cooperation.

Gleason, John Streator rs (1857-1863)

f Holograph. 162 pp. HDC [LDS Church Archives] (Ms f 183 item 1)

Mission in New York State, April 1857 to July 1858. Visited home in Montoe, Ne Baptized his eigh one-year-old father. Visited sister in Ohio. Another sister in M Return to Utah, 18

94. Gledhill, Sarah uto bio graphy (18: tobiography (185) escript. 12 pp. U rmon Diaries bx

Ram ling recollection Salt Luke, 1865. Des Adam Gledhill, 188

895. Glives, James I Autobiography (18: Autobiography (182) 893. Gleason, John Streator **Diaries (1857–1863)** 

Diaries (1857–1863)

Reference book of 2,894 known diaries

Alphabetical (last name)

Where to find the diary

Synopsis of content / bio info

Microfilm of Holograph. 162 pp. HDC [LDS Church Archives] (Ms f 183 item 1)

Mission in New York State, April 1857 to July 1858. Visited home in Monroe, New Baptized his eighty-one-year-old father. Visited sister in Ohio. Another sister in M Return to Utah, 1858. Mission to England, 1860–63.

## Guide to Mormon Diaries and Autobiographies (Manual Tagging)

to emigrate, 1852.891. Gill, David R.

Dia ri es (1887-1889) Diaries (1887-1889)

Holograph. 6 vols. 16-18 cm. HDC [LDS Church Archives] (Ms d 1486)

British Mission, 1887-89.

#### 892. Ginn, John I., captain Autobiography (1857-1858)

Autobiography (1857-1858)

Photocopy of types cript. 147 pp. Princeton [Firestone Library, Princeton U] Typescript. 184 pp. UHi [Ut St Hist Soc] (Man A77)

"Personal recollections." Departure from Independence, Missouri, 1857. Member of Magraw's Pacific Wag on Road Expedition. Joined Tenth Infantry at Fort Laramie. Amusing account of Mormon raid on mules. Resigned and threw self on mercy of Mormons. To Salt Lake. Roomed with "Mother Taylor." View of Mormonism by a scoffer and unbeliever. Aiken party. Complaints about polygamy. To California, where author spread charges against Mormons. Details of investigations.

Not all first-hand observation. One section goes back to tell the history of Mormonism, seen as a fraud. Written about 1903. Comment by Charles Kelly in 1959 suggests that the account, while generally reliable, drew upon "an old man's imagination." Repetition. Utah War is gone over a second time, with quotations from government reports. Details on Mormon-Indian cooperation.

Gleason, John Streator

rs (1857-1863)

f Holograph. 162 pp. HDC [LDS Church Archives] (Ms f 183 item 1)

York State, April 1857 to July 1858. Visited home in Monroe, Ne one-year-old father. Visited sister in Ohio. Another sister in M Return to Utah, 18

94. Gledhill, Sarah uto bio graphy (18: tobiography (185) escript. 12 pp. U rmon Diaries bx

Salt Luke, 1865. Des Adam Gledhill, 188

895. Glives, James I Autobiography (18: Autobiography (182)

Reference book of 2,894 known diaries

Alphabetical (last name)

Where to find the diary

Synopsis of content / bio info

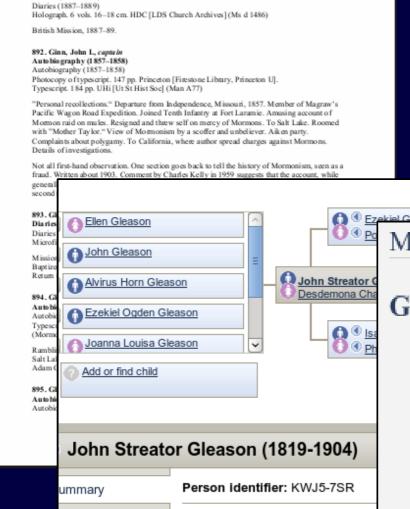
893. Gleason, John Streator **Diaries (1857–1863)** 

Diaries (1857–1863)

Microfilm of Holograph. 162 pp. HDC [LDS Church Archives] (Ms f 183 item 1)

Mission in New York State, April 1857 to July 1858. Visited home in Monroe, New Baptized his eighty-one-year-old father. Visited sister in Ohio. Another sister in M Return to Utah, 1858. Mission to England, 1860-63.

## Guide to Mormon Diaries and Autobiographies (Manual Tagging)



to emigrate, 1852.891, Gill, David R.

Diaries (1887-1889)

Manual search in new FamilySearch using:

**David Bitton's Guide** 

pioneer DB on lds.org

Mormon Pioneer Overland Travel, 1847–1868

Gleason, John Streator

Birth Date: 13 Jan. 1819

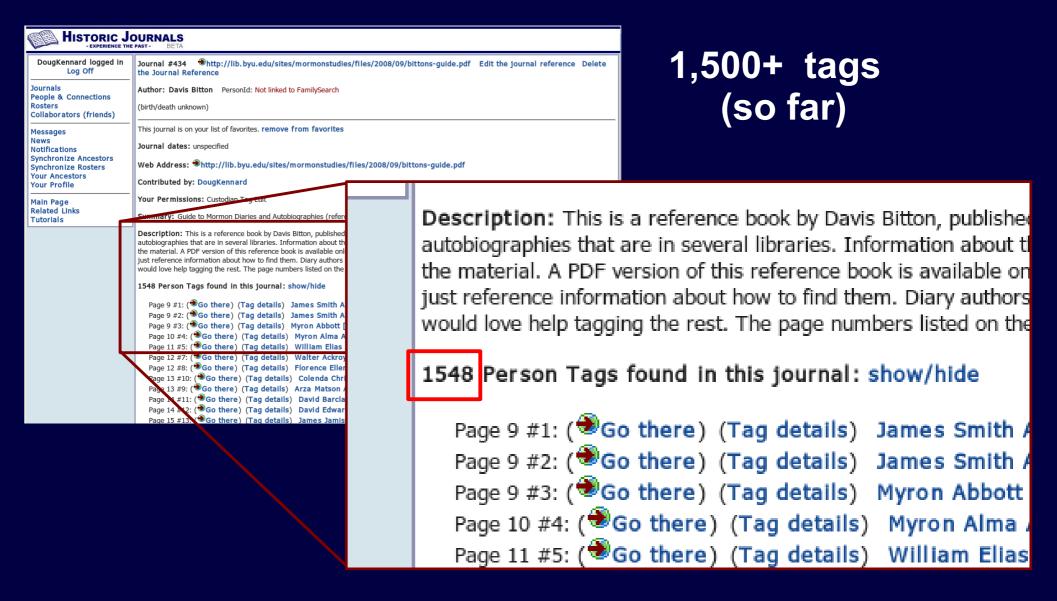
Death Date: 21 Dec. 1904

Gender: Male

Age: 28

Company: Brigham Young Pioneer Company (1847)

## Guide to Mormon Diaries and Autobiographies (Manual Tagging)



## Mormon Missionary Diaries / Overland Trails

433 Diaries

**PersonIDs of authors:** 

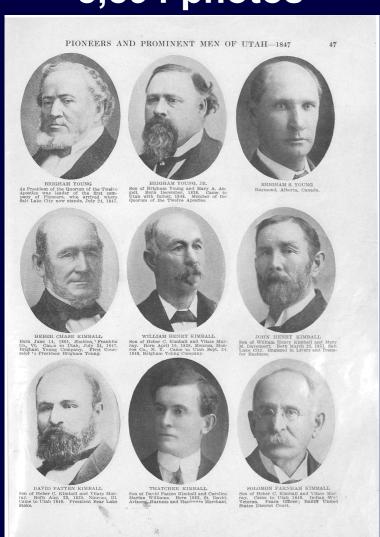
used online biographical info to search

PersonIDs of people talked about:

"Crowd-source" the tagging

Needed: bigger crowd to help tag

### **5,894 photos**



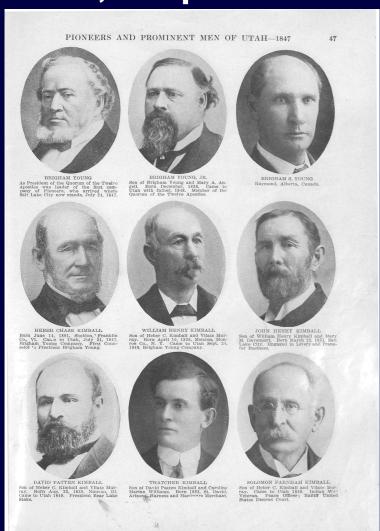
Start: PDF with (poor) OCR

Auto extract photos (3x3 grid)

Manually correct obvious crop/OCR errors

Auto parse: name, birth, parents' names

**5,894 photos** 



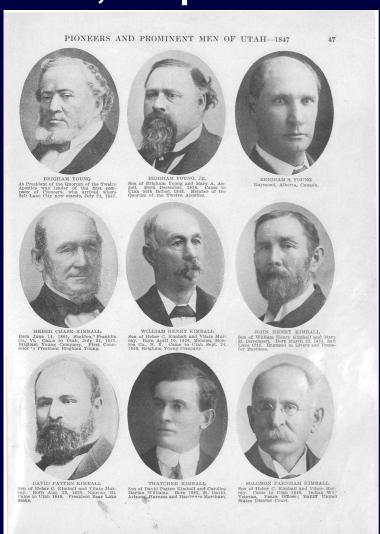
**Start: PDF with (poor) OCR** 

Auto extract photos (3x3 grid)

Manually correct obvious crop/OCR errors

Auto parse: name, birth, parents' names

### **5,894 photos**



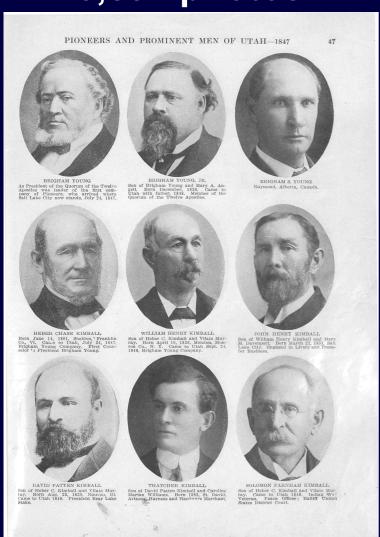
**Start: PDF with (poor) OCR** 

Auto extract photos (3x3 grid)

Manually correct obvious crop/OCR errors

Auto parse: name, birth, parents' names

### **5,894 photos**



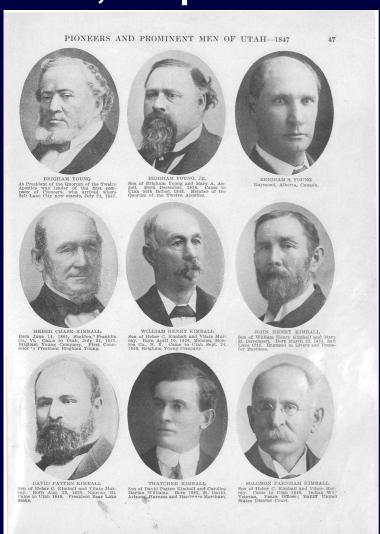
**Start: PDF with (poor) OCR** 

Auto extract photos (3x3 grid)

Manually correct obvious crop/OCR errors

Auto parse: name, birth, parents' names

### **5,894 photos**



**Start: PDF with (poor) OCR** 

Auto extract photos (3x3 grid)

Manually correct obvious crop/OCR errors

Auto parse: name, birth, parents' names

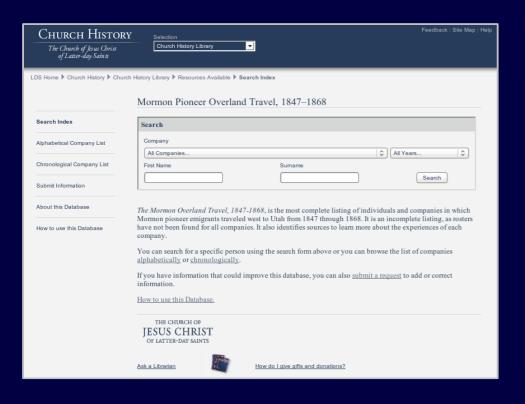
## Mormon Overland Travel pioneer database (on classic.lds.org)



Rosters
Trail Excerpts

Browse by company Search by person name (one at a time)

## Mormon Overland Travel pioneer database



Goal: index by PersonID, provide hyperlink back to the database

Code to automatically find the PersonIDs

Status: proof of concept, one pioneer company, no links on our site



**Brigham Young's Pioneer Company - 1847** 

Crawl the roster to get: Name, Birth, Death, URLs of Trail Excerpts

Use FamilySearch API to search

Store PersonIDs for "close" matches (up to 3)

Manually verify results

Observation: only use best match (not best 3)

**Results:** 

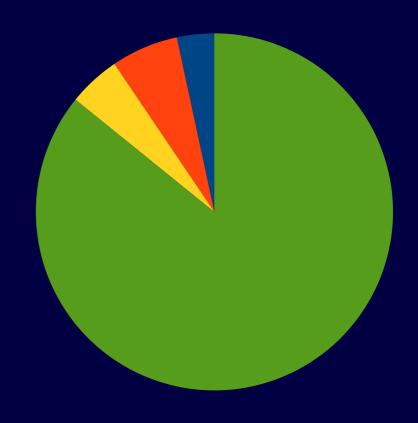
**Total People: 148** 

Correct: 127

Unsure: 7

Incorrect: 9

Not found: 5



Results very promising: auto-link entire database

**Unanswered Questions:** 

First pioneer company (better records?)

Only 3 women, 2 were wrong (maiden vs married?)

## **Future Work**

### Investigate auto-linking more document types

- Census
- Birth records
- Death records

## Semi-automatic tagging

- find names in diary
- compare to family names,
   other resources (rosters, city dir., census, news, etc.)
- ranked suggestions in tagging tool

## Thank You

**Incorrect: 9** 

- 3 sons of person (2 of which were juniors)
- 2 Females (maiden vs married name?)
- 2 completely wrong
- 1 James Cox instead of James Case, but had a James Case as an alternate name
- 1 different guy with same 1<sup>st</sup> / last name and born / died within a year of the same dates