Extraction Rule Creation by Text Snippet Examples

David W. Embley (Brigham Young University & FamilySearch) George Nagy (Rensselaer Polytechnic Institute)







Project Objectives

- Extraction Engines
 - Rules
 - NLP
 - Machine Learning
- Organization Pipeline
 - Curate
 - Import





- Rule Creation by Text Snippet Examples
 - (Hopefully) usable by non-experts
 - (Hopefully) rapid development
 - (Hopefully) high quality results

Pattern Examples

PARISH OF KILBARCHAN.

REGISTER OF MARRIAGES, 1649-1772, AND BAPTISMS, 1651-1769.

The dates, unless otherwise stated, are those of Baptisms. p. means proclaimed; m. married.

Adam, James, and Jannet Bannatyne, in Hair Lawis, 1676 James, 15 Dec. 1672. Robert, 15 Oct. 1676. Margaret, 6 April 1679. Adam, James, in Kilbarchan, and Jane Lyle p. 2 Aug. 1746 William, born 23 June 1747. James, born 19 June 1749. Mary, born 5 Jan. 1752. Janet, 24 Nov. 1754. William, born 10 Dec. 1755. lames, born 24 Oct. 1758. Isobel, born 12 Nov. 1761. p. 13 July 1751 Adam, James, and Janet Aitken Adam, James, in Coalboag, and Margaret Aitken Mathew, born 29 June 1752. Janet, born 26 Nov. 1754. Elizabeth, born 15 April 1756. William, 24 Oct. 1760. Adam, John, in Penall loan, 25 April 1651. obn, 30 May 1652. Isobell, 23 Mar. 1655. Adam, John, par. of Lochwinnoch Marion, 24 Jan. 1662. Adam, John, and Jean Reid John, 14 Nov. 1673. Adam, John, par., and Agnes Andro in Killellan, in Clavens 1691 m. Killellan 23 Jan. 1679 Janet, 8 Nov. 1601. Mary, 2 April 1693. Adam, John, and Janet King, in Kilmacolm John, June 1683. Adam, John, in Kirktoun, and Margaret How William, 24 Nov. 1689. Jean, 13 Mar. 1692. Margaret, 8 June 1694. Adam, John, in Barr, and Agnes Cochran William, 5 May 1703. Elizabeth, 24 June 1705. Robert, 3 Oct. 1709.

ABERNATHY, ELMER d 4 April 1924 252 Bellevernon Ave BD Greenville Cem 6 Apr 1924 b 5 Oct 1863 age 60-5-29 pd by ELLEN ABERNATHY

- ACCETTE, FRANK d 16 Oct 1942 Friday 3:15p.m. Greenville Dke Co OH BD Oct 1942 Abbottsville Cem Dke Co OH b 20 April 1897 Montreal Canada age 45-5-26 f JOSEPH ACCETTE m ACNES QUEIRLLON waiter in restaurant sp & informant LENA ACCETTE 405 Central Ave physician Dr Mills religion Catholic War record: enlisted 21 Feb 1918 disch 19 Aug 1919 World War I Canadian Expeditionary Force Army 2nd Depot Batt C.O. Reg. Services Catholic Church clergy Father Gnau
- ADAMS, ADAM DANIEL d 30 Aug 1931 Miami Valley Hosp Dayton OH BD Castine Cem 2 Sept 1931 b 18 May 1872 Preston Co WV age 59-3-12 f COLEMAN ADAMS Barber Co WV m RACHAEL BOWMAN Barber Co WV married farmer
- ADAMS, ANNA E. 2215 Rustic Road Dayton OH d 24 Aug 1942 Monday 2:45a.m. Dayton Montgomery Co OH BD 26 Aug 1942 Franklin Cem OH b 7 Feb 1856 Franklin OH age 86-6-17 f DAVID ADAMS single housekeeper informant Mrs LOUIS MEYERS 2215 Rustic Road Dayton OH physician Dr Sacks clergy Rev Jones Dayton OH services Baptist Church in Franklin OH
- ADKINS, HESTER d 6 Nov 1925 Weaver's Station BD Fort Jefferson Cem 8 Nov 1925 age 84-3-9 chg to RILEY ADKINS, pd by JAMES A. ADKINS
- ADKINS, JAMES ALEXANDER d 2 Sept 1944 Wayne Hosp Greenville OH BD 4 Sept 1944 Fort Jefferson Cem Dke Co OH b 19 June 1872 Vandalia IL age 72-2-13 f RILEY ADKINS Dke Co OH m HESTER McCOOL retired rural mail carrier sp CORA ADKINS 65 years sisters Mrs MARY VIETS Dayton & Mrs CLATE RIEGLE Fort Jefferson
- AIKEY, JACOB CLARENCE d 2 Oct 1937 N.W. of Pikeville 1½ mile BD Oakland Cem 5 Oct 1937 b 14 Dec 1855 Union Co PA age 81-9-18 f THOMAS AIKEY Maine m ALVINA KATHERMAN married retired farmer sp LYDIA
- AIKEY, LYDIA ANN d 27 Aug 1925 7½ mile N.E. of Greenville BD Oakland Cem 30 Aug 1925 age 60-9-23 chg to JACOB AIKEY
- ALBRIGHT, ADAM C. d 28 June 1920 Piqua OH hosp BD Abbottsville Cem 1 July 1920 age 72-7-27
- ALBRIGHT, CARL ROLAND d 21 June 1917 VanBuren Twp BD Abbottsville Cem 23 June 1917 b VanBuren Twp age 11-2-10 f ALLEN ALBRIGHT m ANNA WEAVER
- ALBRIGHT, CATHARINE d 10 May 1930 4 mile S.W. BD Greenville Mausoleum 13 May 1930 age 94-5-20 pd by DAYTON & CHAS ALBRIGHT
- ALBRIGHT, ESTHER R. d 1 Jan 1946 113 Sherman St Dayton OH BD Abbottsville Cem Dke Co OH 3 Jan 1946 b 22 July 1863 Butler Co OH age 82-6-9 f THOMAS BENTON MORRIS Butler Co OH m ANGELINE HARROD Hamilton Co OH housekeeper widow sp WINFIELD S. ALBRIGHT 1 daughter Mrs HENRY RANCH 4 sons HENDER-SON of Greenville WILBUR of Greenville GEO of Dayton ELBERT of Dayton 12 grandchildren 2 brothers ARTHUR MORRIS Venice OH & SAM MORRIS Harrison OH 2 sisters Miss ELLA MORRIS Greenville OH & Mrs ADA HARP Tulsa OK

THE ELY ANCESTRY. 419 SEVENTH GENERATION.

241213. Mary Eliza Warner, b. 1826, dau. of Samuel Selden Warner and Azubah Tully; m. 1850, Joel M. Gloyd (who was connected with Chief Justice Waite's family).

243311. Abigail Huntington Lathrop (widow), Boonton, N. J., b. 1810, dau. of Mary Ely and Gerard Lathrop; m. 1835, Donald McKenzie, West Indies, who was b. 1812, d. 1839.

(The widow is unable to give the names of her husband's parents.) Their children :

1. Mary Ely, b. 1836, d. 1859. 2. Gerard Lathrop, b. 1838.

243312. William Gerard Lathrop, Boonton, N. J., b. 1812, d. 1882, son of Mary Ely and Gerard Lathrop; m. 1837, Charlotte Brackett Jennings, New York City, who was b. 1818, dau. of Nathan Tilestone Jennings and Maria Miller. Their children:

Maria Jennings, b. 1838, d. 1840.
 William Gerard, b. 1840.
 Donald McKenzie, b. 1840, d. 1843.
 Anna Margaretta, b. 1843.
 Anna Catherine, b. 1845.

243314. Charles Christopher Lathrop, N. Y. City, b. 1817, d. 1865, son of Mary Ely and Gerard Lathrop; m. 1856, Mary Augusta Andruss, 992 Broad St., Newark, N. J., who was b. 1825, dau. of Judge Caleb Halstead Andruss and Emma Sutherland Goble. Mrs. Lathrop died at her home, 992 Broad St., Newark, N. J., Friday morning, Nov. 4, 1898. The funeral services were held at her residence on Monday, Nov. 7, 1898, at half-past two o'clock P. M. Their children:

Charles Halstead, b. 1857, d. 1861.
 William Gerard, b. 1858, d. 1861.
 Theodore Andruss, b. 1860.
 Emma Goble, b. 1862.

Miss Emma Goble Lathrop, official historian of the New York Chapter of the Daughters of the American Revolution, is one of the youngest members to hold office, but one whose intelligence and capability qualify her for such distinction. Miss Lathrop is not without experience; in her present home and native city, Newark, N. J., she has filled the positions of secretary and treasurer to the Girls' Friendly Society for nine years, secretary and president of the Woman's Auxiliary of Trinity Church Parish, treasurer of the St. Catherine's Guild of St. Barnabas Hospital, and manager of several of Newark's charitable institutions which her grandparents were instrumental in founding. Miss Lathrop traces her lineage back through many generations of famous progenitors on both sides. Her maternal ancestors were among the early settlers of New Jersey, among them John Ogden, who received patent in 1664 for the purchase of Elizabethtown, and who in 1673 was

Pattern Examples – Large (layout <u>components)</u>

PARISH OF KILBARCHAN.

REGISTER OF MARRIAGES, 1649-1772, AND BAPTISMS, 1651-1769.

The dates, unless otherwise stated, are those of Baptisms. p. means proclaimed; m. married.

Adam, James, and Jannet Bannatyne, in Hair Lawis, 1676 James, 15 Dec. 1672. Robert, 15 Oct. 1676.

Adam, James, in Kilbarchan, and Jane Lyle p. 2 Aug. 1746
William, born 23 June 1747.
James, born 19 June 1749.
Mary, born 5 Jan. 1752.
Janet, 24 Nov. 1754.
Tomes how at Oct 1755.
Isobel born 12 Nov 1261
Adam, James, and Janet Aitken p. 13 July 1751
Adam, James, in Coalboag, and Margaret Aitken
Mathew, born 20 June 1752.
Janet, born 26 Nov. 1754.
Elizabeth, born 15 April 1756.
William, 24 Oct. 1760.
Adam, John, in Penall
Joan, 25 April 1651.
John, 30 May 1652.
Isobell, 23 Mar. 1055.
Adam, John, par. of Lochwinnoch
Adam John and Jean Paid
John La Nov 1672
Adam John par and Agnes Andro in Killellan, in Clavens 1601
m. Killellan 23 Jan. 1679
Janet, 8 Nov. 1691.
Mary, 2 April 1693.
Adam, John, and Janet King, in Kilmacolm
John, June 1683.
Adam, John, in Kirktoun, and Margaret How
William, 24 Nov. 1689.
Jean, 13 Mar. 1692.
Margaret, 8 June 1694.
Adam, John, in Barr, and Agnes Cochran
William, 5 May 1703.
Robert a Oct 1700
A A

ABERNATHY, ELMER d 4 April 1924 252 Bellevernon Ave BD Greenville Cem 6 Apr 1924 b 5 Oct 1863 age 60-5-29 pd by ELLEN ABERNATHY

- ACCETTE, FRANK d 16 Oct 1942 Friday 3:15p.m. Greenville Dke Co OH BD Oct 1942 Abbottsville Cem Dke Co OH b 20 April 1897 Montreal Canada age 45-5-26 f JOSEPH ACCETTE m AGNES QUEIRLLON waiter in restaurant sp & informant LENA ACCETTE 405 Central Ave physician Dr Mills religion Catholic War record: enlisted 21 Feb 1918 disch 19 Aug 1919 World War I Canadian Expeditionary Force Army 2nd Depot Batt C.O. Reg. Services Catholic Church clergy Father Gnau
- ADAMS, ADAM DANIEL d 30 Aug 1931 Miami Valley Hosp Dayton OH BD Castine Cem 2 Sept 1931 b 18 May 1872 Preston Co WV age 59-3-12 f COLEMAN ADAMS Barber Co WV m RACHAEL BOWMAN Barber Co WV married farmer
- ADAMS, ANNA E. 2215 Rustic Road Dayton OH d 24 Aug 1942 Monday 2:45a.m. Dayton Montgomery Co OH BD 26 Aug 1942 Franklin Cem OH b 7 Feb 1856 Franklin OH age 86-6-17 f DAVID ADAMS single housekeeper informant Mrs LOUIS MEYERS 2215 Rustic Road Dayton OH physician Dr Sacks clergy Rev Jones Dayton OH services Baptist Church in Franklin OH
- ADKINS, HESTER d 6 Nov 1925 Weaver's Station BD Fort Jefferson Cem 8 Nov 1925 age 84-3-9 chg to RILEY ADKINS, pd by JAMES A. ADKINS
- ADKINS, JAMES ALEXANDER d 2 Sept 1944 Wayne Hosp Greenville OH BD 4 Sept 1944 Fort Jefferson Cem Dke Co OH b 19 June 1872 Vandalia IL age 72-2-13 f RILEY ADKINS Dke Co OH m HESTER McCOOL retired rural mail carrier sp CORA ADKINS 65 years sisters Mrs MARY VIETS Dayton & Mrs CLATE RIEGLE Fort Jefferson
- AIKEY, JACOB CLARENCE d 2 Oct 1937 N.W. of Pikeville 1½ mile BD Oakland Cem 5 Oct 1937 b 14 Dec 1855 Union Co PA age 81-9-18 f THOMAS AIKEY Maine m ALVINA KATHERMAN married retired farmer sp LYDIA
- AIKEY, LYDIA ANN d 27 Aug 1925 7½ mile N.E. of Greenville BD Oakland Cem 30 Aug 1925 age 60-9-23 chg to JACOB AIKEY
- ALBRIGHT, ADAM C. d 28 June 1920 Piqua OH hosp BD Abbottsville Cem 1 July 1920 age 72-7-27
- ALBRIGHT, CARL ROLAND d 21 June 1917 VanBuren Twp BD Abbottsville Cem 23 June 1917 b VanBuren Twp age 11-2-10 f ALLEN ALBRIGHT m ANNA WEAVER

ALBRIGHT, CATHARINE d 10 May 1930 4 mile S.W. BD Greenville Mausoleum 13 May 1930 age 94-5-20 pd by DAYTON & CHAS ALBRIGHT

ALBRIGHT, ESTHER R. d 1 Jan 1946 113 Sherman St Dayton OH BD Abbottsville Cem Dke Co OH 3 Jan 1946 b 22 July 1863 Butler Co OH age 82-6-9 f THOMAS BENTON MORRIS Butler Co OH m ANGELINE HARROD Hamilton Co OH housekeeper widow sp WINFIELD S. ALBRIGHT 1 daughter Mrs HENRY RANCH 4 sons HENDER-SON of Greenville WILBUR of Greenville GEO of Dayton ELBERT of Dayton 12 grandchildren 2 brothers ARTHUR MORRIS Venice OH & SAM MORRIS Harrison OH 2 sisters Miss ELLA MORRIS Greenville OH & Mrs ADA HARP Tulsa OK

THE ELY ANCESTRY. SEVENTH GENERATION.

419

241213. Mary Eliza Warner, b. 1826, dau. of Samuel Selden Warner and Azubah Tully; m. 1850, Joel M. Gloyd (who was connected with Chief Justice Waite's family).

243311. Abigail Huntington Lathrop (widow), Boonton, N. J., b. 1810, dau. of Mary Ely and Gerard Lathrop; m. 1835, Donald McKenzie, West Indies, who was b. 1812, d. 1839.

(The widow is unable to give the names of her husband's parents.) Their children :

1. Mary Ely, b. 1836, d. 1859. 2. Gerard Lathrop, b. 1838.

243312. William Gerard Lathrop, Boonton, N. J., b. 1812, d. 1882, son of Mary Ely and Gerard Lathrop; m. 1837, Charlotte Brackett Jennings, New York City, who was b. 1818, dau. of Nathan Tilestone Jennings and Maria Miller. Their children:

Maria Jennings, b. 1838, d. 1840.
 William Gerard, b. 1840.
 Donald McKenzie, b. 1840, d. 1843.
 Anna Margaretta, b. 1843.
 Anna Catherine, b. 1845.

243314. Charles Christopher Lathrop, N. Y. City, b. 1817, d. 1865, son of Mary Ely and Gerard Lathrop; m. 1856, Mary Augusta Andruss, 992 Broad St., Newark, N. J., who was b. 1825, dau. of Judge Caleb Halstead Andruss and Emma Sutherland Goble. Mrs. Lathrop died at her home, 992 Broad St., Newark, N. J., Friday morning, Nov. 4, 1898. The funeral services were held at her residence on Monday, Nov. 7, 1898, at half-past two o'clock P. M. Their children:

Charles Halstead, b. 1857, d. 1861.
 William Gerard, b. 1858, d. 1861.
 Theodore Andruss, b. 1860.
 Emma Goble, b. 1862.

Miss Emma Goble Lathrop, official historian of the New York Chapter of the Daughters of the American Revolution, is one of the youngest members to hold office, but one whose intelligence and capability qualify her for such distinction. Miss Lathrop is not without experience; in her present home and native city, Newark, N. J., she has filled the positions of secretary and treasurer to the Girls' Friendly Society for nine years, secretary and president of the Woman's Auxiliary of Trinity Church Parish, treasurer of the St. Catherine's Guild of St. Barnabas Hospital, and manager of several of Newark's charitable institutions which her grandparents were instrumental in founding. Miss Lathrop traces her lineage back through many generations of famous progenitors on both sides. Her maternal ancestors were among the early settlers of New Jersey, among them John Ogden, who received patent in 1664 for the purchase of Elizabethtown, and who in 1673 was

Pattern Examples – Intermediate



5. Anna Catherine, b. 1845.

Pattern Examples – Small (text

Adam, James, in Kilbarchan, and Jane Lyle p. 2 Aug. 1 William, born 23 June 1747. James, born 19 June 1749. Mary, born 5 Jan. 1752. Janet, 24 Nov. 1754. William, born 10 Dec. 1755. James, born 24 Oct. 1758. James, born 12 Nov. 1761.

ALBRIGHT, ESTHER R. d 1 Jan 1946 113 Sherman St Dayton OH BD Abbottsville Cem Dke Co OH 3 Jan 1946 b 22 July 1863 Butler Co OH age 82-6-9 f THOMAS BENTON MORRIS Butler Co OH m ANGELINE HARROD Hamilton Co OH housekeeper widow sp WINFIELD S. ALBRIGHT 1 daughter Mrs HENRY RANCH 4 sons HENDER-SON of Greenville WILBUR of Greenville GEO of Dayton ELBERT of Dayton 12 grandchildren 2 brothers ARTHUR MORRIS Venice OH & SAM MORRIS Harrison OH 2 sisters Miss ELLA MORRIS Greenville OH & Mrs ADA HARP Tulsa OK

> 243312. William Gerard Lathrop, Boonton, N. J., b. 1812, d. 1882, son of Mary Ely and Gerard Lathrop; m. 1837, Charlotte Brackett Jennings, New York City, who was b. 1818, dau. of Nathan Tilestone Jennings and Maria Miller. Their children:

Maria Jennings, b. 1838, d. 1840.
 William Gerard, b. 1840.
 Donald McKenzie, b. 1840, d. 1843.
 Anna Margaretta, b. 1843.
 Anna Catherine, b. 1845.

Adam, James, in Kilbarchan, and Jane Lyle	I	p.	2 Aug.	1746
William, born 23 June 1747.				
James, born 19 June 1749.				
Mary, born 5 Jan. 1752.				
Janet, 24 Nov. 1754.				
William, born 10 Dec. 1755.				
James, born 24 Oct. 1758.				
Isobel, born 12 Nov. 1761.				

Couple record Name: ^ Adam, James, SpouseName: and Jane Lyle MarriageDate: p. 2 Aug. 1746 \$

Name: ^ Cap , Cap , SpouseName: and Cap Cap MarriageDate: p. Num Cap . Num \$





Name: ^ Cap , born Name: ^ Cap , Num

Adam, James, in Kilbarchan, and Jane Lyle William, born 23 June 1747.		p. 2 Aug. 1746
James, born 19 June 1749.		
Tanet, 24 Nov. 1754.		
James, born 24 Oct. 1755.		
Isobel, born 12 Nov. 1761.		- 1 1



Parent1: ^ Cap , Cap ,



Person record

Name: ^James, born Name: ^Janet, 24 ChristeningDate: , 24 Nov. 1754. \$ BirthDate: born 24 Oct. 1758. \$

Name: ^ Cap , born Name: ^ Cap , Num Couple record

Name: ^ Adam, James, SpouseName: and Jane Lyle MarriageDate: p. 2 Aug. 1746 \$

Name: ^ Cap , Cap , SpouseName: and Cap Cap MarriageDate: p. Num Cap . Num \$



Parent1: ^ Cap , Cap ,

...

Step1: Specify the Records

COMET Home Form Builder Logout	
Person Information Add a Form	
Person	Couple
Name * >	Name * >
GenderDesignator	SpouseName MarriageDate MarriagePlace
BirthDate	
BirthPlace	
ChristeningDate	Family
ChristeningPlace	Parent1 *
DeathDate	Parent2
DeathPlace	Child
BurialDate	
BurialPlace	

Step 2: Create Rules

				P	erson	p. means proclaimed; m. married.
Name *	GenderDesignator	BirthDate	BirthPlace	ChristeningDate	Ch	Adam, James, and Jannet Bannatyne, in Hair Lawis, 1676
nes				15 Dec. 1672		Robert, 15 Oct. 1676.
					- II	Margaret, 6 April 1679.
					- 11	William, born 23 June 1747.
						James, born 19 June 1749. Mary, born 5 Jan. 1752.
						Janet, 24 Nov. 1754.
James. 15	5 Dec. 1672. ELINE	Run Save	2			ames, born 24 Oct. 1755.
						Isobel, born 12 Nov. 1761.
						Adam, James, and Janet Altken p. 13 July 1751 Adam, James, in Coalboag, and Margaret Aitken
						Mathew, born 29 June 1752.
						Elizabeth, born 15 April 1756.
						William, 24 Oct. 1760.
						Joan, 25 April 1651.
						John, 30 May 1652.
						Adam, John, par. of Lochwinnoch
						Marion, 24 Jan. 1662.
						John, 14 Nov. 1673.
						Adam, John, par., and Agnes Andro in Killellan, in Clavens 1691 m. Killellan 23 Jan. 1670
						Janet, 8 Nov. 1691.
						Mary, 2 April 1693. Adam John and Japet King in Kilmacolm
						John, June 1683.
						Adam, John, in Kirktoun, and Margaret How William 24 Nov. 1680
						Jean, 13 Mar. 1692.
						Margaret, 8 June 1694. Adam: John, in Barr, and Agnes Cochran
						William, 5 May 1703.
						Elizabeth, 24 June 1705.
					-	A A

Step 2: Create Rules

				P	erson	The dates, unless otherwise stated, are those of Baptisms. p. means proclaimed; m. married.
Name *	GenderDesignator	BirthDate	BirthPlace	ChristeningDate	Cł	Adam, James, and Jannet Bannatyne, in Hair Lawis, 1676
James				15 Dec. 1672		Robert, 15 Dec. 1672. Margaret, 6 April 1670.
William		23 June 1747				Adam, James, in Kilbarchan, and Jane Lyle p. 2 Aug. 1746 William, born 23 June 1747. James, born 10 June 1740.
					- 11	Mary, born 5 Jan. 1752. Janet, 24 Nov. 1754.
born 23	3 June 1747. ELINE	Run Save				James, born 12 Oct. 1758. Isobel, born 12 Nov. 1761.
						Adam, James, and Janet Aitken p. 13 July 1751 Adam, James, in Coalboag, and Margaret Aitken Mathew, born 20 June 1752.
					- 1	Janet, born 26 Nov. 1754. Elizabeth, born 15 April 1756. William 24 Oct. 1756.
					- 15	Adam, John, in Penall Joan, 25 April 1651.
						John, 30 May 1652. Isobell, 23 Mar. 1655. Adam John per of Lochwinnoch
					- 10	Marion, 24 Jan. 1662. Adam, John, and Jean Reid
						Adam, John, 14 Nov. 1673. Adam, John, par., and Agnes Andro in Killellan, in Clavens 1691 m. Killellan 23 Jan. 1679
						Janet, 8 Nov. 1691. Mary, 2 April 1693.
						Adam, John, and Janet King, in Kilmacoim John, June 1683. Adam, John, in Kirktoun, and Margaret How
					- 12	William, 24 Nov. 1689. Jean, 13 Mar. 1692.
						Adam, John, in Barr, and Agnes Cochran William, 5 May 1703.
						Elizabeth, 24 June 1705. Robert, 3 Oct. 1709.

Step 2: Create Rules (check rule set)

				Pe	rson -
Name *	GenderDesignator	BirthDate	BirthPlace	ChristeningDate	Cł
James				15 Dec. 1672	
Robert				15 Oct. 1676	
Margaret					
William		23 June 1747			
James		19 June 1749			
Mary					
Janet				24 Nov. 1754	
William					
James					
Isobel] []			
Mathew		29 June 1752			
Janet) [
		15 April 1756			
William				24 Oct. 760	
Joan					
John					
Isobell				23 Mar. 1655	

Step 3: Process Candidate Rules

1523	3 Nam	. 1753 Brown, William, in Kilbarchan, and Sarah > Make Dismiss
48	Name	Feb. 1759. Brune, William Jeane, > Make Dismiss
19	Name	Oct. 1752. Napier and William, born 8 Feb > Make Dismiss
18	Name	Robert, in Hilhead James (daughter), 8 June > Make Dismiss

John, 16 Feb. 1690.	
Brown, John, and Jean Cumming, in Kirktoun	m. 25 Dec. 1601
Jean, 3 Sept. 1693.	
Brown, John, and Jean Wilson, in Meiklefultoun	
Margaret, 28 July 1698.	
Brown, John, in Kilbarchan, and Jane M'Dowall	p. 13 Nov. 1761
Jane, born 3 Mar. 1704.	and a substants
Brown, Matthew, par. or Killelian, and Elizabeth Kinble,	par. p.21 June 1705
broune, Robert, in Finnead	
Bronne William in Fultours, 1676 in Little Foltours	
James, 18 May 1656.	
Agnes, 15 Ian, 1658.	
Robert, 20 Jan. 1650.	
Katherine, 26 Feb, 1660,	
Agnes, his daughter, gotten in adultery	with Jeane Adam,
presented by Robert Adam, in Clipping	gs, 18 May 1655.
Broune, William, and Margaret Henderson, par. of I	Paisley,
in Mains of Houstoun	m. 10 Nov. 1676
William, 26 Aug. 1677.	
Margaret, 16 April 1680.	
Marie, 22 June 1689	
Brown, William, and Janet Cochrane, in Little Fultoun	m. 1 June 1682
Robert, 6 June 1690.	
Alexander, 18 May 1693.	
Alexander, 27 May 1694.	
Elizabeth, 27 Sept. 1696.	
Hugh, 20 Jan. 1698.	
Brown, William, elder, in Littlefultoun	
Alexander, 3 Jan. 1692.	
Brown, William, in Kirktown, and Isobel Allason	
Anna, 10 Nov. 1695.	
William, 26 Dec. 1697.	
Martha, 2 Feb. 1701.	
Brown William nan and Agnes MiCibb in Glassow	11. 25 Aug. 1752
Brown, William in Killmenhan and Sarah Douglas	ht = 3 1 + 4 E = 1 3 3
Mary born 11 Mar 1765.	
Bruce John at Craigends and Elizabeth Martin	
William born 17 Feb. 1750.	
Brune, William	
Jeane, I July 1653.	

Step 3: Process Candidate Rules

				Person		
Name *	GenderDesignator	BirthDate	BirthPlace	ChristeningDate	Ch	
James						
					- 1	
	SLINE James (daughter) 8	Run Save				
					- 1	
					- 1	
					_	
					- 1	
					- 1	
					- 1	
					- 1	
					- 1	

ean, 3 Sept. 1603. Brown, John, and Jean Wilson, in Meiklefultoun Margaret, 28 July 1698. Brown, John, in Kilbarchan, and Jane M'Dowall p. 13 Nov. 1761 Jane, born 3 Mar. 1764. Brown, Matthew, par. of Killellan, and Elizabeth Kibble, par. p. 21 June 1765 Broune, Robert, in Hilhead James (daughter), 8 June 1655. Broune, William, in Fultoune, 1656 in Little Fultoune -, son, 6 May 1655. James, 18 May 1656. Agnes, 15 Jan. 1658. Robert, 30 Jan. 1650. Katherine, 26 Feb. 1660. Agnes, his daughter, gotten in adultery with Jeane Adam, presented by Robert Adam, in Clippings, 18 May 1655. Broune, William, and Margaret Henderson, par. of Paisley, in Mains of Houstoun m. 10 Nov. 1676 William, 26 Aug. 1677. Margaret, 16 April 1680. Marie, 22 June 1689. Brown, William, and Janet Cochrane, in Little Fultoun 10. 1 June 1682 Robert, 6 June 1690. Alexander, 18 May 1693. Alexander, 27 May 1694. Elizabeth, 27 Sept. 1696. Hugh, 20 Jan. 1698. Brown, William, elder, in Littlefultoun Alexander, 3 Jan. 1692. Brown, William, in Kirktown, and Isobel Allason Anna, 10 Nov. 1695. William, 26 Dec. 1697. Martha, 2 Feb. 1701. May, 23 May 1706. Brown, William, par., and Agnes M'Gibb, in Glasgow p. 25 Aug. 1753 Brown, William, in Kilbarchan, and Sarah Douglas Mary, born 11 Mar. 1765. Bruce, John, at Craigends, and Elizabeth Martin William, born 17 Feb. 1759. Brune, William Jeane, 1 July 1653.

Step 3: Process Candidate Rules

	John, born 20 Sept. 1765. Inglis, John, in Kilbarchan, and Janet Lang, 1768 in Muir of Waterstoun p. 7 May 1763 Janet, born 13 Aug. 1763. Janet, born 30 May 1766. John, born 2 Oct. 1768. Inglis, John, in Quarreltoun, Abbey par. of Paisley Elizabeth, born 7 June 1768. Inglis, William, and Anne Colquboun Patrick, 4 Jan. 1745. Alexander, 20 Mar. 1748. Inglis, William, and Lillias Campbell William, born 17 April 1762.
	Inglis, William, in Kilbarchan, and Janet Cameron p. 10 Sept. 1768 Janet, born 25 Nov. 1769. Inglis, William, in Abbey par. of Paisley, and Barbara Orr
19NameOct. 1752. Napier and William, born 8 Feb>MakeDismiss	Jack, Matthew, par., and Agnes Whyte, par. of Houstown p. 30 Mar. 1764 Jackson, Andrew, in Mains of Blackstown, and Agnes Wilson James, born 7 June 1750. James, born 22 Oct. 1752.
	Napier and William, born 6 Feb. 1755. Napier, born 6 May 1758. Jackson, James, and Margaret Love m. 8 Dec. 1659 Jackson, James, in Clochodrick Issabell, 10 Sept. 1660. Jackson, Jaces, in Auchnames, and Mary Love T6 April 1675.
	Jackson, Robert, and Elizabeth Riddell Patrick, t Aug. 1701. Jackson, Robert, in Hutthead, and Isobel King James, 20 April 1705. Jackson, Robert, and Janet Speir p. 13 Aug. 1763 Jackson, William, in Cartside, and Marion Hill William, 3 Jan. 1706. Jameson, Hugh, and Janet Ewen, in Miniabroch, in Lochwinnoch pat-
	Jamieson, James, in Corbar, and Mary Jackson Robert, born 24 Oct. 1767. Jameson, John, and Jonet Brydon, in Lochwinnoch Agnes, 16 Mar. 1673. Jameson, John, and Issabell Barclay, in Lochwinnoch Hugh, 9 Feb. 1679.

GreenQQ (current implementation)

- Green: tools that improve with use
- Q1: Quick
 - Quick to learn to use
 - Quick to execute
- Q2: Quality
 - Quality rules
 - Quality results
- GreenQQ characterization: record-based NER

Demo (input doc's)

Houstown William lawful in to 28 Houstown in Rillarchan and Stuart al in Selectand & Matha Stevenson Man amer Stevenvon & Barbara Rit Mithen Elizabeth, lawful D." to John Allen & Samet Helson his spouse was baptored Der other Reid John, hasful con to Jahn herd and Margt. Storie use was baptored Deer 12the Stevenson Marion lawful Dr. to Tames Stevenson and Tanet Marthur his spouse was haptired Der 12 th Hamilton Absander; lawful in to John Hamilton Gardener to Blackstown was baptined Der. 20the Thorgrafo Margaret, lawful Dith. Hear. Nighet his Spoure was bapling Der 25 Aithen John, lawful don to James Aithen in Sandholes & Mary Henderoon his spower was bapt. Dei 26th Marshell Janet lawful

<Page number="0" width="1432.0" height="2743.0">

<Line top="563.28699" bottom="639.28253" left="181.42943" right="1213.43237">

<Word top="563.28699" bottom="639.28253" left="181.42943" right="468.42950" value ="PARISH">

<Letter top="563.28699" bottom="639.28253" left="181.42943" right="223.26042" offset="0" value="P" />
<Letter top="563.28699" bottom="639.28253" left="227.15288" right="281.47299" offset="1" value="A" />
<Letter top="563.28699" bottom="639.28253" left="285.36545" right="335.54761" offset="2" value="R" />
<Letter top="563.28699" bottom="639.28253" left="339.44006" right="364.49353" offset="3" value="I" />
<Letter top="563.28699" bottom="639.28253" left="368.38596" right="364.49353" offset="4" value="S" />
<Letter top="563.28699" bottom="639.28253" left="368.38596" right="410.21695" offset="4" value="S" />

</Word>

<Word top="563.28699" bottom="639.28253" left="521.24884" right="622.90009" value ="0F">

<Letter top="563.28699" bottom="639.28253" left="521.24884" right="575.56897" offset="7" value="0" />
<Letter top="563.28699" bottom="639.28253" left="581.06909" right="622.90009" offset="8" value="F" />
</Word>

<Word top="563.28699" bottom="639.28253" left="681.06116" right="1192.58398" value ="KILBARCHAN">

<Letter top="563.28699" bottom="639.28253" left="681.06116" right="735.38135" offset="10" value="K" />
<Letter top="563.28699" bottom="639.28253" left="737.42059" right="762.47406" offset="11" value="I" />
<Letter top="563.28699" bottom="639.28253" left="764.51331" right="810.48224" offset="12" value="L" />
<Letter top="563.28699" bottom="639.28253" left="764.51331" right="810.48224" offset="12" value="L" />
<Letter top="563.28609" bottom="639.28253" left="764.51331" right="810.48224" offset="12" value="L" />
<Letter top="563.28609" bottom="639.28253" left="764.51331" right="810.48224" offset="12" value="L" />
<Letter top="563.28609" bottom="639.28253" left="912.59155" right="964.28609" bottom="910" />
<Letter top="563.28609" bottom="639.28253" left="912.59155" right="910.28609" rig

PARISH OF KILBARCHAN.

REGISTER OF MARRIAGES, 1649-1772, AND BAPTISMS, 1651-1769.

The dates, unless otherwise stated, are those of Baptisms. p. means proclaimed; m. married.

Adam, James, and Jannet Bannatyne, in Hair Lawis, 1676 James, 15 Dec. 1672. Robert, 15 Oct. 1676. Margaret, 6 April 1679. Adam, James, in Kilbarchan, and Jane Lyle p. 2 Aug. 1746 William, born 23 June 1747. James, born 19 June 1749. Mary, born 5 [an. 1752. Janet, 24 Nov. 1754. William, born 10 Dec. 1755. lames, born 24 Oct. 1758. Isobel, born 12 Nov. 1761. Adam, James, and Janet Aitken p. 13 July 1751 Adam, James, in Coalboag, and Margaret Aitken Mathew, born 20 June 1752.

> PARISH OF KILBARCHAN. REGISTER OF MARRIAGES, 1649-1772, AND BAPTISMS, 1651-1769. The dates, unless otherwise stated, are those of Baptisms. p. means proclaimed; m. married. Adam, James, and Jannet Bannatyne, in Hair Lavvis, 1676 James, 15 Dec. 1672. Robert, 15 Oct. 1676. Margaret, 6 April 1679. Adam, James, in Kilbarchan, and Jane Lyle p. 2 Aug. 1746 William, born 23 June 1747. James, born 19 June 1749. Mary, born 5 Jan. 1752. Janet, 24 Nov. 1754. William, born 10 Dec. 1755. James, born 24 Oct. 1758. Isobel, born 12 Nov. 1761. Adam, James, and Janet Aitken p. 13 July 1751 Adam, James, in Coalboaq, and Margaret Aitken Mathew horn 29 June 1752

Demo (I/O)

Input

Documents library

KilbarchanParish

Name

- KilbarchanParish.001.txt
- KilbarchanParish.002.txt
- KilbarchanParish.003.txt
- KilbarchanParish.004.txt
- KilbarchanParish.005.txt
- KilbarchanParish.006.txt
- •••
- KilbarchanParish.142.txt
- 📄 KilbarchanParish.143.txt

В C D Е . Α 1 2 LIT born 3 4 EX Name SLINE William, born William 5 EX BirthDate born 23 June 1747. ELINE 23 June 1747 6 7 InKperson_demo0.0 (+)∃ ₹

Output

FamilyMain Elapsed time = 2.64 seconds
>>>

Templates:

1,1118,['Name'],"['SLINE', 'William', ',', 'born']" 2,959,['BirthDate'],"['born', '23', 'June', '1747', '.', 'ELINE']"

Robert, born 2 Dec. 1761. Aitken, Margaret, in Boydsyard Thomas, born 23-24 May 1750.

Records



Demo (candidate rule generation)



SLINE Elizabeth (natural), 29

Name

Initial Experimental Resu

Adam, James, in Kilbarchan, and Jane Lyle William, born 23 June 1747. James, born 19 June 1749. Mary, born 5 Jan. 1752. Janet, 24 Nov. 1754. William, born 10 Dec. 1755. James, born 24 Oct. 1758. Isobel, born 12 Nov. 1761.

p. 2 Aug. 174

			K	ilbarchan A	ccuracy Res	ults (Pages	20, 41, 123)					
Person							Soft: Cor	rect = Corre	ct+Partial	Hard: Incor	rect = Partia	al+Incorrect
Class / Record	Nr. Rules	Nr. Found in Book	Nr. in Pages	Correct	Partial	Incorrect	Recall	Precision	F-score	Recall	Precision	F-Score
Name	2	3959	92	87	0	0	0.95	1.00	0.97	0.95	1.00	0.97
BirthDate	1	2823	24	18	0	0	0.75	1.00	0.86	0.75	1.00	0.86
ChristeningDate	1	7461	. 68	57	0	0	0.84	1.00	0.91	0.84	1.00	0.91
Field Values	4	14243	184	162	0	0	0.88	1.00	0.94	0.88	1.00	0.94
Record	1	3959	92	75	12	0	0.95	1.00	0.97	0.82	0.86	0.84
Couple				•	D 1		Soft: Cor	rect = Corre	ct+Partial	Hard: Incor	rect = Partia	al+Incorrect
Class / Record	Nr. Rules	Nr. Found in Book	Nr. in Pages	Correct	Partial	Incorrect	Recall	Precision	F-score	Recall	Precision	F-Score
Name	1	/908	6/	65	0	0	0.97	1.00	0.98	0.97	1.00	0.98
SpouseName	1	4324	55	49	0	0 0	0.89	1.00	0.94	0.89	1.00	0.94
MarriageDate	3	2610	28	24	0	0 0	0.86	1.00	0.92	0.86	1.00	0.92
MarriagePlace	1	67	3	1	0	0 0	0.33	1.00	0.50	0.33	1.00	0.50
Field Values	6	14909	153	139	0	0	0.91	1.00	0.95	0.91	1.00	0.95
Record	1	7908	67	55	9	2	0.96	0.97	0.96	0.82	0.83	0.83
Family							Soft: Cor	rect - Corre	ct⊥Dartial	Hard: Incor	roct – Parti	al±Incorrect
	Nr. Pulos	Nr. Found in Rool	Nr in Pages	Corroct	Partial	Incorroct	Bocall	Procision		Recall	Procision	
Parent1	NI. Kules	7569	69 69	63			0 91	1 00	0.95	0 91	1 00	0.95
Parent2	1	4316	58	52	0	0	0.90	1.00	0.95	0.90	1.00	0.95
Child	2	3959	91	85	0	0	0.93	1.00	0.97	0.93	1.00	0.97
Field Values	4	15844	218	200	0	0	0.92	1.00	0.96	0.92	1.00	0.96
Record	1	7569	72	61	1	. 4	0.86	0.94	0.90	0.85	0.92	0.88
6							C (1) C					
Summary	No. Dela		N D	C	D	1	Soft: Cor	rect = Corre	ct+Partial	Hard: Incor	rect = Partia	al+incorrect
Class / Record	Nr. Rules	Nr. Found in Book	Nr. In Pages	Correct	Partial	Incorrect	Recall	Precision	F-score	Recall	Precision	F-Score
Field Values	14	44996	555	501	0	0	0.90	1.00	0.95	0.90	1.00	0.95
Records	3	19436	231	191	22	6	0.92	0.9	0.95	0.83	0.8	0.85
Unique	10	29152										

Initial Experimental Results

ALBRIGHT, ESTHER R. d 1 Jan 1946 113 Sherman St Dayton OH BD Abbottsville Cem Dke Co OH 3 Jan 1946 b 22 July 1863 Butler Co OH age 82-6-9 f THOMAS BENTON MORRIS Butler Co OH m ANGELINE HARROD Hamilton Co OH housekeeper widow sp WINFIELD S. ALBRIGHT 1 daughter Mrs HENRY RANCH 4 sons HENDER-SON of Greenville WILBUR of Greenville GEO of Dayton ELBERT of Dayton 12 grandchildren 2 brothers ARTHUR MORRIS Venice OH & SAM MORRIS Harrison OH 2 sisters Miss ELLA MORRIS Greenville OH & Mrs ADA HARP Tulsa OK

Miller Accuracy Results (Pages 296, 318, 363)												
							Soft: Correct = Correct+Partial		ct+Partial	Hard: Incorrect = Partial+Incorrect		
Class / Recor	d Nr. Rules	Nr.Found in Book	Nr. In Pages	Correct	Partial	Incorrect	Recall	Precision	F-score	Recall	Precision	F-score
Name	2	15265	30	29	1	0	1.00	1.00	(1.00	0.97	0.97	0.97
DeathDate	1	12629	30	30	0	0	1.00	1.00	1.00	1.00	1.00	1.00
DeathPlace	3	5940	26	19	0	0	0.73	1.00	0.84	0.73	1.00	0.84
BurialPlace	2	4865	7	6	0	0	0.86	1.00	0.92	0.86	1.00	0.92
BurialDate	3	1545	26	7	0	0	0.27	1.00	0.42	0.27	1.00	0.42
BirthDate	1	9608	21	21	0	0	1.00	1.00	1.00	1.00	1.00	1.00
BirthPlace	4	3090	21	9	0	1	0.43	0.90	0.58	0.43	0.90	0.58
SpouseName	e 3	3802	15	10	1	1	0.73	0.92	0.81	0.67	0.83	0.74
Father	2	6006	21	13	7	0	0.95	1.00	0.98	0.62	0.65	0.63
Mother	2	5846	17	13	4	0	1.00	1.00	1.00	0.76	0.76	0.76
Field Values	23	68596	214	157	13	2	0.79	0.99	0.88	0.73	0.91	0.81
Record	1	15265	30	29	0	1	0.97	0.97	0.97	0.97	0.97	0.97

"Gotchas"

- Document applicability
- Record identifiers
- Overlapping records
- OCR errors
- Ambiguity
- Boundary-crossing patterns
- Application tailoring





Future Work (in progress)

- Build Interface
- Adjust Code to Resolve "Gotchas"
- Seize Opportunities
 - Improve candidate pattern identification
 - Assess and adjust for increased usability





Conclusion

- Rule creation by text snippet examples
- (Hopefully) objectives will be achieved



- Usable by non-experts (examples only; user-friendly interface)
- Quick development (click/copy rule development; candidate rule generation)
- Quality results (good precision and recall)

Conclusion

• Rule creation by text snippet examples



- (Hopefully) objectives will be achieved
 - Usable by non-experts (examples only; user-friendly interface)
 - Quick development (click/copy rule development; candidate rule generation)



sults (good precision and recall)

