

PAF Insight - Improvements in Record Matching

John Vilburn

Ohana Software LLC

Improvements in Record Matching

- Focus on overall user experience
- User interface
 - “What do these options mean?”
 - “I don’t want to wade through all of the false matches.”
 - “What is the difference between these records?”
- Matching methods
 - Presenting fewer false matches
 - Finding more valid matches

PAF Match/Merge

- Too many options
- No way to focus on the most likely matches
- Difficult to see the difference between records

PAF Match / Merge

– Too many options

Match/Merge Options [?] [X]

Matching

Match primarily on ---

- ☒ Names that sound alike (Soundex)
- ☐ Names with exact spelling
- ☐ Ancestral File Numbers (AFNs)
- ☐ Unique record serial numbers

Advanced...

Merge

- ☐ Automatically merge exact duplicates
- ☒ Combine notes
- ☒ Combine source citations
- ☒ Combine multimedia objects
- ☒ Confirm when Merge button pressed

OK Cancel Help

Advanced Matching Options [?] [X]

Years between birth dates

- ☐ Include individuals with no surname
- ☐ Include individuals with no birth date
- ☒ Consider Ancestral File Numbers (AFNs)
- ☐ Consider middle names
- ☒ Consider parents

OK Cancel Help

PAF Match/Merge

- Can't focus on likely matches

Merge Individuals [?] [X]

Individual Information | Marriage Information | Family Information

Primary Individual - 712 Find Edit

Betsey Adams

B: 16 Sep 1774 B, W, Massachusetts

C:

D: 9 Sep 1775 B, W, Massachusetts

B:

AFN: Sex: F LMD: 15 Dec 1999

Baptism: Child

Endowed: Child

SealPar:

Father: Benjamin Adams

Mother: Eunice Hale

Spouse:

Spouse:

Spouse:

Duplicate Individual - 713 Find Edit

☐ Betsy Adams

B: ☐ 21 Aug 1776 ☐ B, W, Massachusetts

C: ☐ ☐

D: ☐ ☐

B: ☐ ☐

AFN: ☐ Sex: ☐ F LMD: 15 Dec 1999

Baptism: ☐

Endowed: ☐

SealPar: ☐

☐ Father: Benjamin Adams

Mother: Eunice Hale

Spouse: Cyrus Ayres

Spouse:

Spouse:

Select All Clear All

Merge < Switch > Next Match Previous Options... Close Help

PAF Match/Merge

– Hard to see differences

Merge Individuals [?] [X]

Individual Information | Marriage Information | Family Information

Primary Individual - 712	Duplicate Individual - 713
<input type="button" value="Find"/> <input type="button" value="Edit"/>	<input type="button" value="Find"/> <input type="button" value="Edit"/>
Betsey Adams	<input type="checkbox"/> Betsy Adams
B: 16 Sep 1774 B, W, Massachusetts	B: <input type="checkbox"/> 21 Aug 1776 <input type="checkbox"/> B, W, Massachusetts
C:	C: <input type="checkbox"/>
D: 9 Sep 1775 B, W, Massachusetts	D: <input type="checkbox"/>
B:	B: <input type="checkbox"/>
AFN: Sex: F LMD: 15 Dec 1999	AFN: <input type="checkbox"/> Sex: <input type="checkbox"/> F LMD: 15 Dec 1999
<hr/>	
Baptism: Child	Baptism: <input type="checkbox"/>
Endowed: Child	Endowed: <input type="checkbox"/>
SealPar:	SealPar: <input type="checkbox"/>
<hr/>	
Father: Benjamin Adams	<input type="checkbox"/> Father: Benjamin Adams
Mother: Eunice Hale	Mother: Eunice Hale
<hr/>	
Spouse:	Spouse: Cyrus Ayres
Spouse:	Spouse:
Spouse:	Spouse:
<div><input type="button" value="Select All"/> <input type="button" value="Clear All"/></div>	

PAF Insight - Merge

- No options
 - Designed for genealogists, not techies
- Matches prioritized – Most probable first
- Differences highlighted

PAF Insight - Merge

– Matches prioritized

Barnett.paf

List: Normal

Previous Next Merge Switch Research Not A Match

Status	Prim. RIN	Primary Name	Dup. RIN	Duplicate Name	% Match
	307	Sarah /UNKNOWN/	308	Sarah /UNKNOWN/	98
	455	Thomas /Bartlett/	1044	Thomas /Bartlett/	98
	2202	Charles /Bruce/	2208	Charles /Bruce/	98
	2200	Ephraim /Bruce/	2203	Ephraim /Bruce/	98
	2341	Benjamin /Deverix/	2348	Benjamin /Deverix/	98
	2340	Hannah /Deverix/	2346	Hannah /Deverix/	98
	2347	Lois /Deverix/	2349	Lois /Deverix/	98
	2338	Ralph /Deverix/	2342	Ralph /Deverix/	98
	2339	Ruth /Deverix/	2344	Ruth /Deverix/	98
	1759	Henry /Farrar/	1760	Henry /Farrar/	98
	2212	Abraham /Howe/	2219	Abraham /Howe/	98
	2078	Alice Hay /Hurdman/	2079	Alice /Hurdman/	98
	1801	Frances /Ollyver/	1802	Frances /Oliver/	98
	1806	Francis /Ollyver/	1808	Francis /Oliver/	98
	1811	Gilbert /Ollyver/	1812	Gilbert /Oliver/	98
	1796	Mary /Ollyver/	1797	Mary /Ollyver/	98
	454	Joan /Bartlett/	1043	Joane /Bartlett/	96
	1809	Dorothy /Ollyver/	1810	Dorothy /Oliver/	96
	1798	Edmund /Ollyver/	1799	Edmund /Oliver/	96
	2040	Eleanor /Adams/	2041	Eleanor /UNKNOWN/	94
	459	William /Titcomb/	1054	William /Titcomb/	94
	2037	Mary /UNKNOWN/	2039	Mary /Cross ?/	92
	267	Mary /Cruttenden/	932	Mary /Cruttenden/	92
	2110	John /Ferreby/	2111	John /Ferreby/ Esquire	90
	2391	Sara /Elson/	2392	Sarah /Elson/	86
	2011	Samuel /Banks/	2012	Samuel /Baker/	84
	2076	Alice /Hay/	2077	Alice /Hurlman/	84
	2075	Alice /Huroman/	2076	Alice /Hay/	84
	2075	Alice /Huroman/	2077	Alice /Hurlman/	84
	2320	Bethia /Newhall/	2323	Elizabeth /Newhall/	84
	1791	UNKNOWN /Winslow/	1792	John /Winslow/	84
	306	Margaret /UNKNOWN/	309	Margaret /UNKNOWN/	80
	1840	Thomas /Winslow/ esq.	2093	Thomas Esq. /Winslow/	80

PAF Insight - Merge

– Differences highlighted

The screenshot shows a 'Details' window with a blue title bar and a close button. It contains two tabs: 'Individual' and 'Spouses and Children'. The 'Individual' tab is active, showing a comparison between two individuals. The left pane is titled 'Primary Individual - 2040' and the right pane is titled 'Duplicate Individual - 2041'. Both panes show the name 'Eleanor /Adams/' and 'Eleanor /UNKNOWN/' respectively, with the names highlighted in pink. Below the names, there are fields for 'Sex: Female', 'B:', 'C:', 'D:', and 'Baptism'. The 'Baptism' field has a checkbox. The 'Endowment' field is also present. At the bottom of each pane, it says 'Last Modified: 05 Jul 2002 07:41 PM'.

Field	Primary Individual - 2040	Duplicate Individual - 2041
Name	Eleanor /Adams/	Eleanor /UNKNOWN/
Sex	Female	Female
B:		
C:		
D:		
Baptism		
Endowment		
Last Modified	05 Jul 2002 07:41 PM	05 Jul 2002 07:41 PM

PAF Insight - Merge

– Differences highlighted

The screenshot shows a 'Details' window with a blue title bar and a close button. It contains five tabs: 'Individual', 'Spouses and Children', 'Parents and Siblings', 'Notes', and 'Sources'. The 'Individual' tab is selected. The window is split into two panes. The left pane is titled 'Primary Individual - 2040' and shows details for Eleanor /Adams/. The right pane is titled 'Duplicate Individual - 2041' and shows details for Eleanor /UNKNOWN/. Differences between the two records are highlighted in pink. In the left pane, the marriage date 'Jun 1634' is followed by a pink box containing 'Marshfield, Plymouth ...'. The sealing date '17 Feb 1984' is followed by a pink box containing 'JRIVE'. In the right pane, the marriage date 'Jun 1634' is preceded by an unchecked checkbox and followed by a pink box. The sealing date is preceded by an unchecked checkbox and followed by a pink box. Both panes have a 'Children' section with an empty list box and scroll buttons.

Details

Individual | Spouses and Children | Parents and Siblings | Notes | Sources

Primary Individual - 2040

Eleanor /Adams/

Spouse

Kenelm /Winslow/ (1599-1672) - RIN 1734

Marriage

Jun 1634 **Marshfield, Plymouth ...**

Sealing

17 Feb 1984 **JRIVE**

Children

Duplicate Individual - 2041

Eleanor /UNKNOWN/

Spouse

Kenelm /Winslow/ (1599-1672) - RIN 1734

Marriage

☐ Jun 1634 ☐

Sealing

☐ ☐

Children

PAF Insight – Compare & Sync

Compare1.paf

List: Changed or New Records

Previous Next Update Add

Status	Prim. RIN	Primary Name	Sec. RIN	Secondary Name	% Match
	637	Jeremiah /WOODBURY/	1	Jeremiah /Woodbury/	98
	3058	Elizabeth /West/	3	Elizabeth /West/	96
	3503	Mary /West/	4	Mary /West/	98
	3504	Hannah /West/	5	Hannah /West/	98
	3506	John /West/	6	John /West/	69
	3509	Ruth /West/	7	Ruth /West/	76
	3510	Samuel /West/	8	Samuell /West/	98
	3514	Thomas /West/	9	Thomas /West/	92
	3074	Sarah /WOODBURY/	11	Sarah /Woodbury/	80
	3100	Abigail /WOODBURY/	12	Abigall /Woodbury/	60
	3882	Benjamin /WOODBURY/	13	Benjamin /Woodbury/	72
	639	Elizabeth /BARTLETT/	15	Elizabeth /Bartlett/	95
	2282	John /WOODBURY/	16	John /Woodbury/	90
	2295	Electa /WOODBURY/	18	Electa /Woodbury/	98
	2288	Jerusha /WOODBURY/	19	Jerusha /Woodbury/	90
	2296	John /WOODBURY/	20	John /Woodbury/	88
	2290	Mary Elizabeth /WOODBU...	21	Mary Elizabeth /Woodbury/	90
	2292	Susannah /WOODBURY/	22	Susannah /Woodbury/	88
	2299	Lucy /WOODBURY/	23	Lucy /Woodbury/	98
	2300	Isaac /WOODBURY/	24	Isaac /Woodbury/	98
	2303	Sybil /WOODBURY/	25	Sybil /Woodbury/	96
			26	Mary /Woodbury/	
	2304	Ward /WOODBURY/	27	Ward /Woodbury/	98
			28	Mary /UNKNOWN/	
			29	Elvira /Stevens/	
	3844	Jeremiah /WOODBURY/	30	Jeremiah /Woodbury/	93
	2284	Sybella Jerusha /TUCKER/	31	Sibella Jerusha /Tucker/	95
	2996	Sarah C. /WOODBURY/	32	Sarah /Woodbury/	90
	2653	Mehitable /WOODBURY/	33	Mehitabel /Woodbury/	68
	2998	Jerusha C. /WOODBURY/	34	Jerusha /Woodbury/	90
	3278	Elizabeth /WOODBURY/	35	Elisabeth /Woodbury/	93
	3279	Diana /WOODBURY/	36	Diana /Woodbury/	77
	3280	Theodore /WOODBURY/	37	Theodore /Woodbury/	93

PAF Insight – IGI Search

IGI search Demo.paf

List: Needs Ordinances Region: Norway Searching: ☒

Previous Next Update Search Year Range Exact Stop

Status	Prim. RIN	Primary Name	IGI Name	% Match
Done: 11 Matches	1	Sarah Ellen /PAPE/		
	1	Sarah Ellen /PAPE/	Sarah Ellen /PAPE/	92
	1	Sarah Ellen /PAPE/	Sarah Ellen /Pape/	88
	1	Sarah Ellen /PAPE/	Sarah Ellen /Pape/	84
	1	Sarah Ellen /PAPE/	Sarah Ellen /Pape/	82
	1	Sarah Ellen /PAPE/	Sarah Ellen /Pape/	82
	1	Sarah Ellen /PAPE/	Sarah Ellen /PAPE/	72
	1	Sarah Ellen /PAPE/	SARAH ELLEN /PAPE/	63
	1	Sarah Ellen /PAPE/	Sarah /Pape/	53
	1	Sarah Ellen /PAPE/	Sarah /Pape/	47
	1	Sarah Ellen /PAPE/	Sarah /Pape/	47
	1	Sarah Ellen /PAPE/	Sarah Ellen /Pape/	26
Done: 6 Matches	2	Henrici /PAPE/		
	2	Henrici /PAPE/	Henrici /PAPE/	88
	2	Henrici /PAPE/	Henrici /Pape/	74
	2	Henrici /PAPE/	Henrici /Pape/	74
	2	Henrici /PAPE/	Henrici /Pape/	74
	2	Henrici /PAPE/	Heinrich /Pape/	60
	2	Henrici /PAPE/	Heinrich /Pape/	60
65 Hits, 17 Matches	3	Johannes /Hansen/		
	3	Johannes /Hansen/	Johannes /HANSEN/	80
	3	Johannes /Hansen/	Johannes /Hansen/	76
	3	Johannes /Hansen/	Johannes /Hansen/	74
	3	Johannes /Hansen/	Johannes /Hansen/	74
	3	Johannes /Hansen/	JOHANES /HANSEN/	71
	3	Johannes /Hansen/	Johannes /Hansen/	68
	3	Johannes /Hansen/	Johannes /Hansen/	66
	3	Johannes /Hansen/	JOHANES /HANSEN/	64
	3	Johannes /Hansen/	JON /HANSEN/	34
	3	Johannes /Hansen/	Johanes /HANSEN/	32
	3	Johannes /Hansen/	John /Hansen/ Vinter	32
	3	Johannes /Hansen/	JOHANES /HANSEN/	31
	3	Johannes /Hansen/	JOHANES /HANSEN/	29

PAF Insight

– Matching methods

- Fewer false matches
 - Weighted scoring
 - Date extrapolation
 - Situational rules

PAF Insight

– Matching methods

- More valid matches
 - Weighted scoring
 - Date extrapolation
 - Phonetic analysis
 - Nickname list
 - Situational rules

PAF Insight

– Matching methods

- Weighted scoring factors
 - Given names, Surnames
 - Birth/Christening date
 - Death/Burial date
 - Spouse
 - Marriage date
 - Parents

PAF Insight

– Matching methods

- Date extrapolation and interpolation
 - Birth/Christening date
 - Death/Burial date
 - Marriage date
 - Recursive

PAF Insight

– Matching methods

- Phonetic analysis
 - Multi-level weighting
 - Scandinavian patronymics
 - Abbreviations
 - Other, nationality-specific

PAF Insight

– Matching methods

- Situational rules
 - Child with same name as older deceased sibling
 - Weighted scoring adjustments for certain combinations

PAF Insight tuning

– Human Perception Is Key

- Adjustments to weighted scoring
- Situational rules
- Lower limit to match percentages

Summary

- Improvements were made in user interface and matching method
- Designed for genealogists, not techies
- Tuned to reflect human perception