

# Genealogy Discovery

Özgür Akgün, Tom Dalton, Alan Dearle, Graham Kirby

## Data Linkage

Identifying records that correspond to the same entity in multiple data sources. Many application areas:

- Health Sector
- National Censuses
- National Security
- Digital Libraries
- E-Commerce
- Social Science Research
- Economics

## Genealogy Linkage

Using historical records to reconstruct family histories. Motivation:

- For individuals: desire to learn about ancestors
- For historians and historical demographers: answering research questions.
  - Change in family sizes over time
  - Relation of wealth and causes of death
  - Patterns in occupation changes
  - And many more...



Administrative Data Research Centre Scotland



Page 2]		The undermentioned Houses are situate within the Boundaries of the									
Civil Parish of		Quoad Sacra Parish of		School Board District of		Parliamentary Burgh of		Parliamentary Division of		Royal Burgh of	
Bracadale		Bracadale		Bracadale							
Municipal Burgh of		Police Burgh of		Burgh Ward of		Town of		Village or Hamlet of		Island of	
No. of	HOUSES	NAME and Surname of each Person.	RELATION to Head of Family.	CONDITION as to Marriage.	AGE (last Birthday)	PROFESSION or OCCUPATION.	Employer.	Employed.	Mother Employer nor Head of Family but working for her account.	Gaelic or G. & E.	Visitors 1. Work and Death 2. Blind, Infirm, Imbecile, or Idiot.
4	1	Henry Nicholson	Head	Widow	54	Cotter			Inhabitants of Braehead	Scots	2
5	1	John McLean	Wife	Mar.	47	Minister of Braehead Parish	X		North East	Scots	18
		Johnnie & McLean	Wife	Widow	53	Monitor of P.P.			Braehead Kirkby	Scots	
		Johnnie & McLean	Daughter	Widow	9	Scholar			Braehead Kirkby	Scots	
		Johnnie McLean	Wife	Widow	47				Braehead	Scots	
		Johnnie Scott	Daughter	Widow	14				Braehead	Scots	
		Mary Mc Donald	Serv.	Widow	21	Governess	X		Fifeshire Palmerins	Scots	
		John Mc Lean	Wife	Widow	50	House Maid Domestic			Monktonshire Palmerins	Scots	
		Alexander Mc Lean	Daughter	Widow	25	Cook			Monktonshire Palmerins	Scots	
		Janet Mc Lean	Head	Widow	40	Mistress House	X		Kilmuir	Scots	
		James Mc Lean	Wife	Widow	30	Mistress House			Braehead	Scots	
		Robert Mc Lean	Wife	Widow	32	Writer			Braehead	Scots	
		Robert Mc Lean	Wife	Widow	32	Letter & Correspondency Office	X		Braehead	Scots	
		John Mc Lean	Wife	Widow	32	Writer			Braehead	Scots	
		Barbara Mc Seal	Daughter	Widow	8	Writer			Braehead	Scots	
		Jeanne Mc Seal	Wife	Widow	6				Braehead	Scots	
		John Mc Seal	Wife	Widow	70				Braehead	Scots	
		Catherine Mc Seal	Mother	Widow	70	Writer			Braehead	Scots	
		Janet Mc Seal	Sister	Widow	34	General Servant Domestic	X		Braehead	Scots	
		Catherine Mc Seal	Wife	Widow	50	General Servant Domestic			Braehead	Scots	
5	5	Total of Houses...		Total of Males and Females..	15				Total of Windows.		
Scot.—Sheet A.]											

## Automatic Linkage

- Doing linkage by hand is time consuming, tedious, and error-prone.

## Computer Science can help!

- Automatic linkage algorithms
- Start by preparing the data: digitising, transcribing, cleaning
- Robust comparison methods
- Using statistical tools
- Efficient algorithms that are able to receive input from domain experts
- Can handle very large amounts of data
- We are happy to talk more about the underlying techniques!

## Contact us

- [ozgur.akgun@st-andrews.ac.uk](mailto:ozgur.akgun@st-andrews.ac.uk)
- School of Computer Science
- University of St Andrews



The family tree of Ludwig Herzog von Württemberg (ruled 1568-1593)  
Painting from 1585 - Currently in Landesmuseum Stuttgart