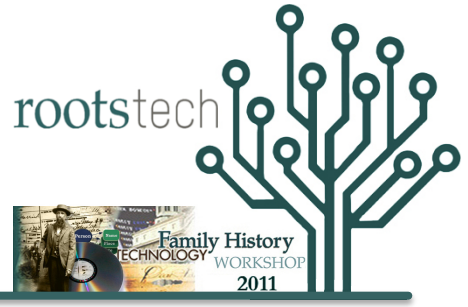
Family History Technology:

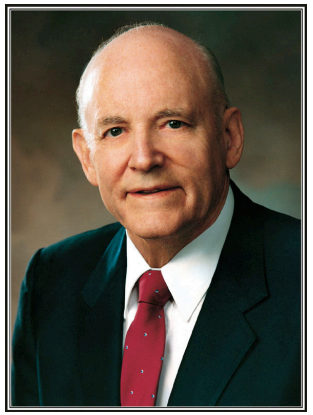


Where have we Been? ...

... and where are we GOING?

William Barrett
Department of Computer Science
Brigham Young University

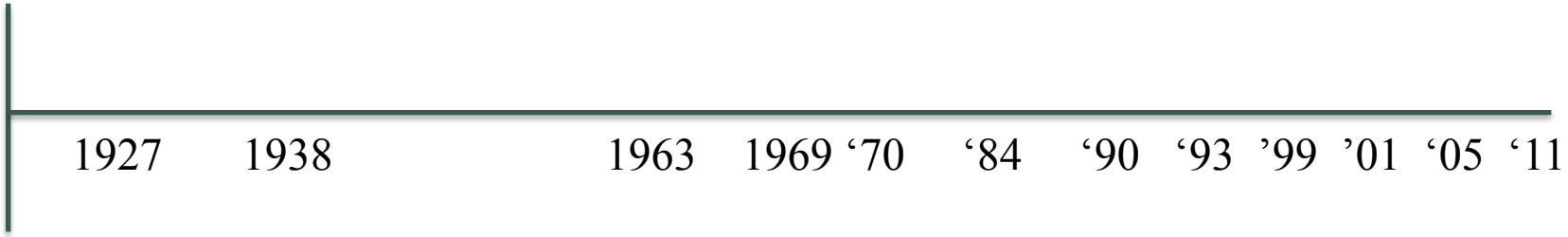
Family History Technology



"The Lord has guided the development of information technology and accelerated its role in work for the dead, and will continue to do so. However, we stand only on the threshold of what we can do with these tools. I feel that our most enthusiastic projections can capture only a tiny glimpse of how these tools can help us - and of the eternal consequences of these efforts."

-President Howard W. Hunter, 100th anniversary of the Genealogical Society of Utah

Technology Milestones in Family History



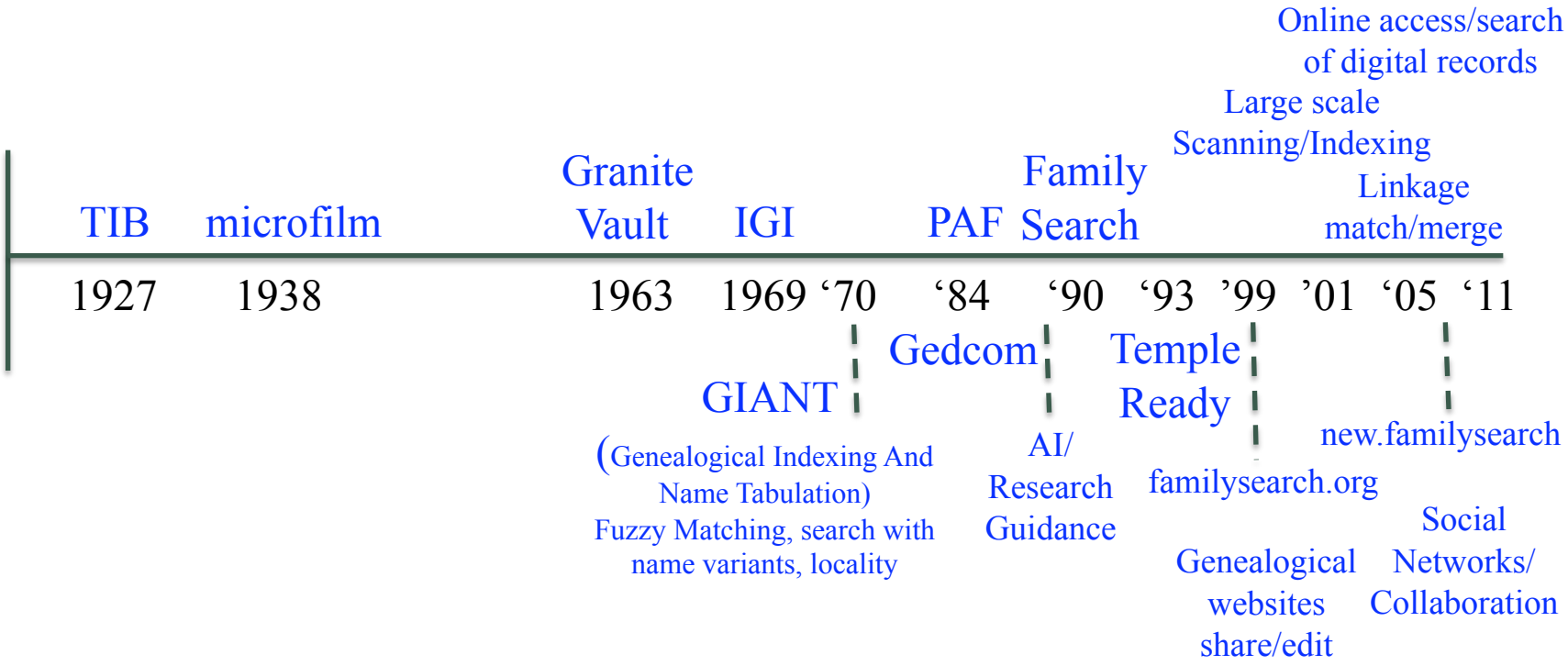
Past Century

Technology Milestones in Family History

rootstech

Family History TECHNOLOGY WORKSHOP 2011

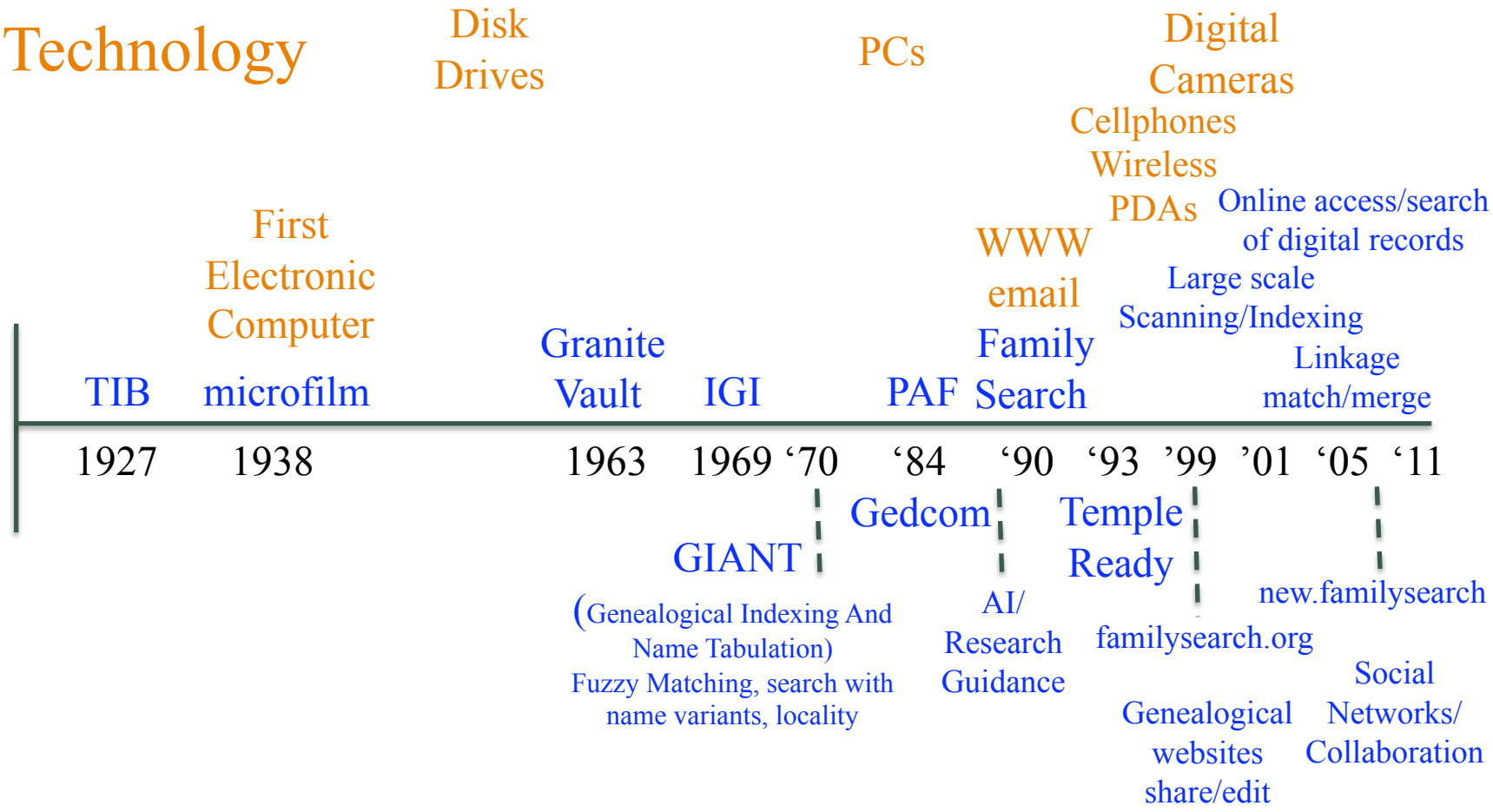
Family History



Technology Milestones in Family History

rootstech

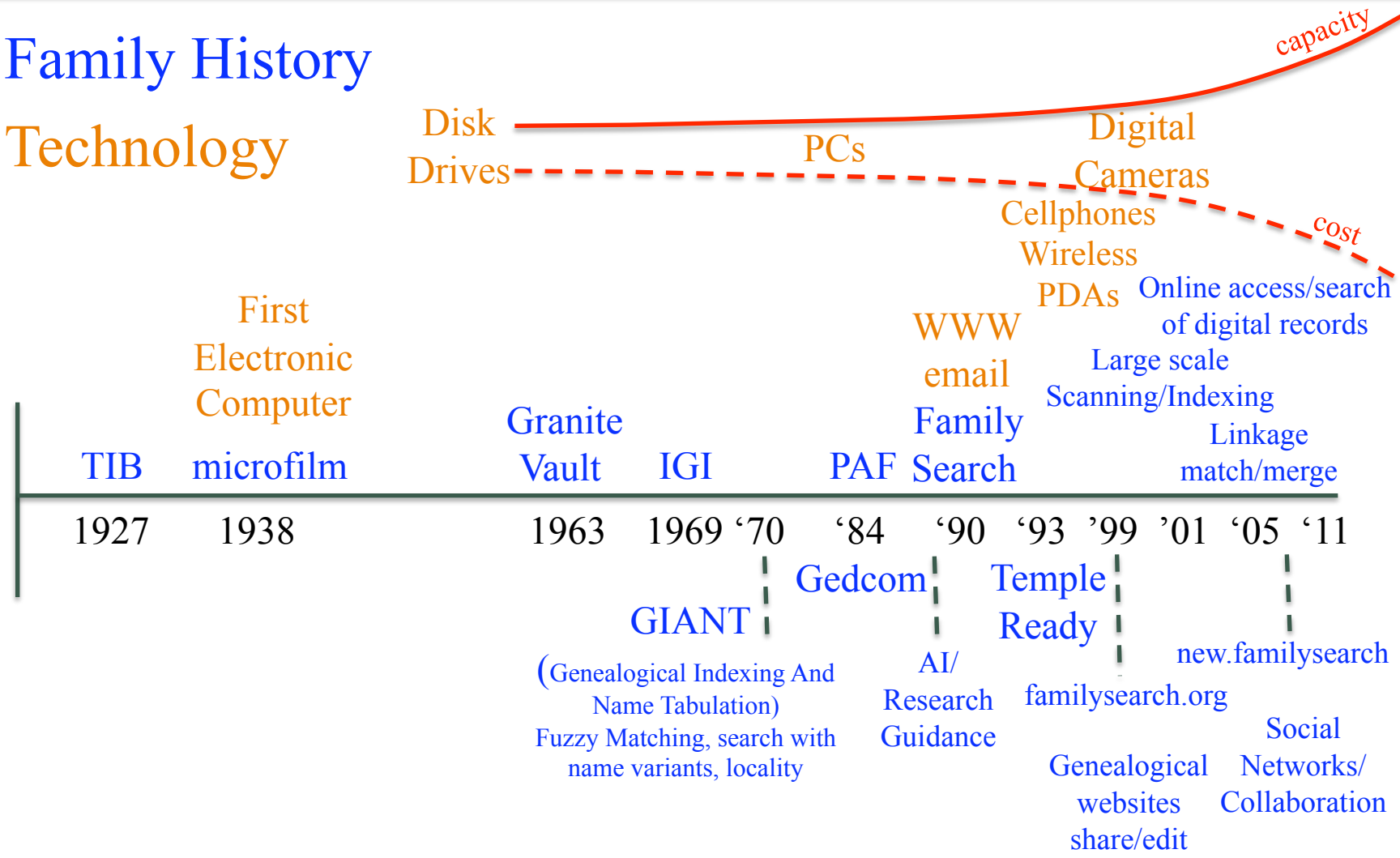
Family History TECHNOLOGY WORKSHOP 2011



Technology Milestones in Family History

rootstech

Family History TECHNOLOGY WORKSHOP 2011

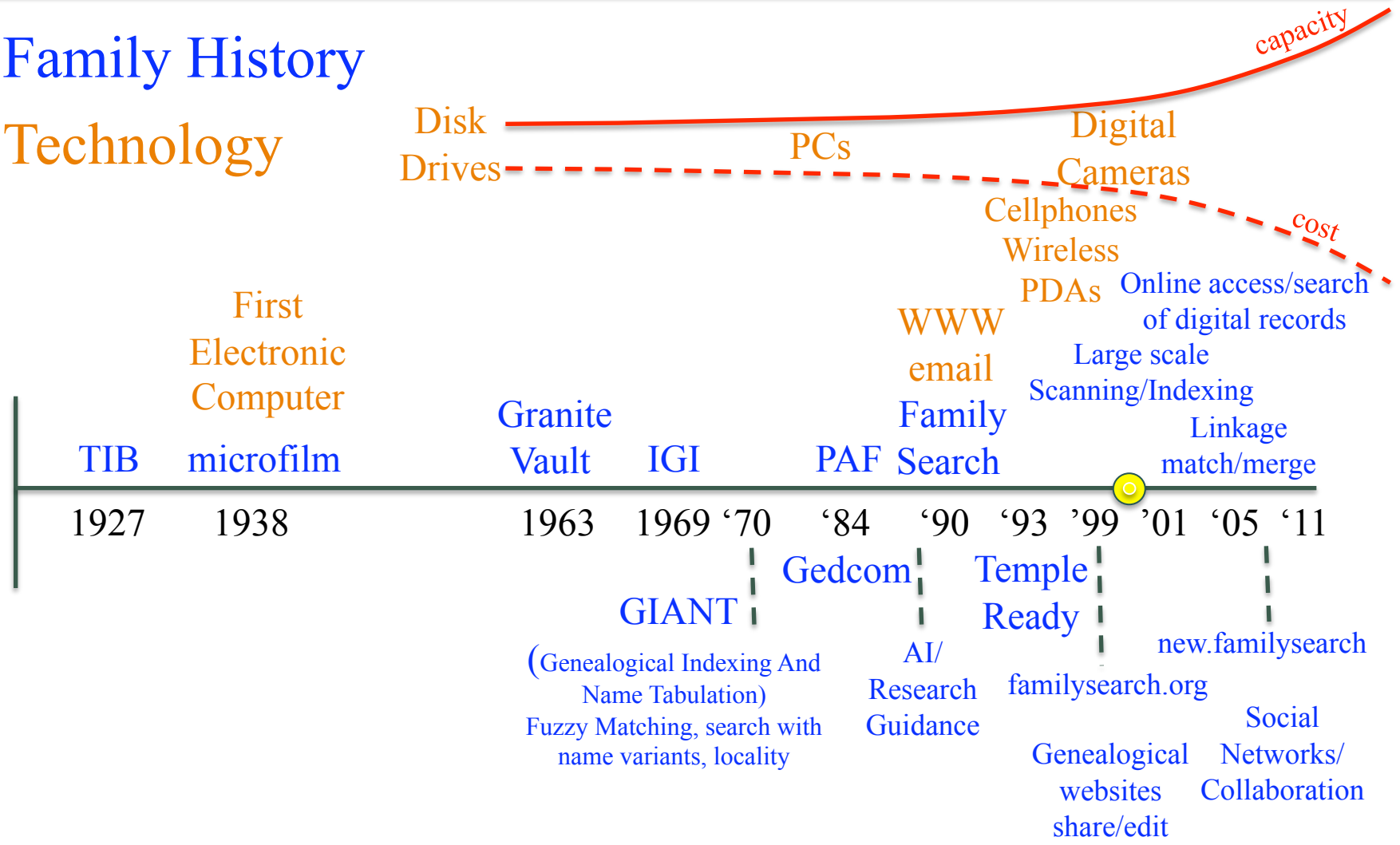


Technology Milestones in Family History

rootstech

Family History TECHNOLOGY WORKSHOP 2011

Family History Technology



Family History Technology Workshop (fht.byu.edu)

rootstech

Family History TECHNOLOGY WORKSHOP 2011

BYU | BRIGHAM YOUNG UNIVERSITY Search BYU GO

Family History Technology Workshop

- Home
- General Information
- Call for Papers
- 2011 Registration
- Local Arrangements
- Contact Us

Related Conferences & Workshops

- Computerized Genealogy Conference
- Annual BYU Genealogy & Family History Conference

Click here for information on the 2011 Conference

10th Annual Workshop on Technology for Family History and Genealogical Research

April 28, 2010
Salt Lake Convention Center

Current Technology Innovation in the CS Department

- Digital Roots
- OnePage Genealogy
- Relationship Finder

Important Dates

January 21, 2011
Papers Due
February 11, 2011
Day of Conference

Previous Workshops

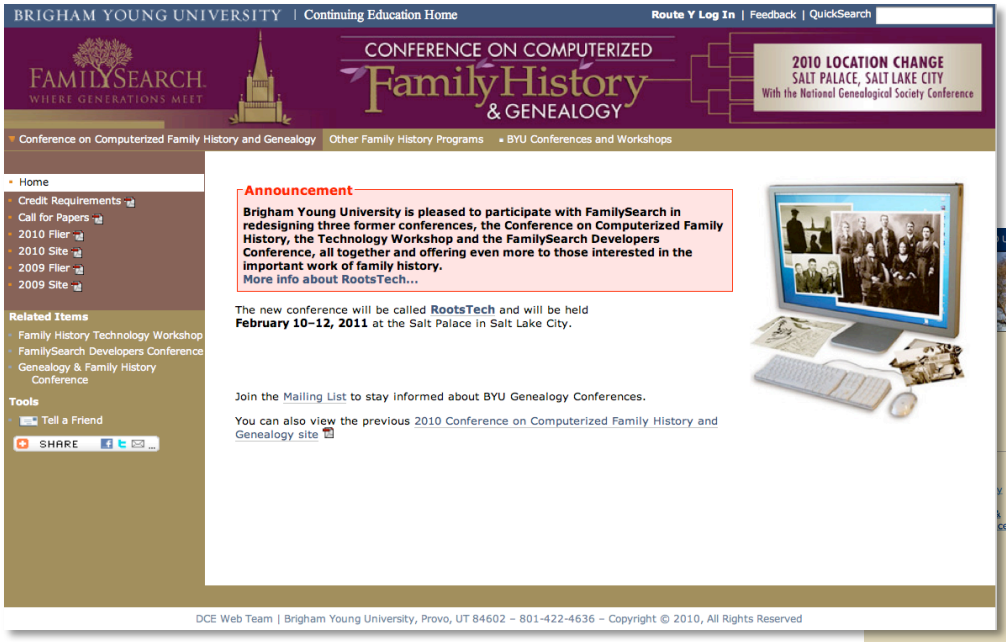
- 2010
- 2009
- 2008
- 2007
- 2006
- 2005
- 2004
- 2003
- 2002
- 2001

Sponsored by: BYU Computer Science • CS Alumni Association • BYU

Creators + Users

rootstech

Family History TECHNOLOGY WORKSHOP 2011



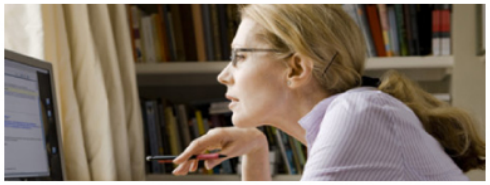
Users + Implementers + Creators = Rootstech

rootstech

Family History TECHNOLOGY WORKSHOP 2011

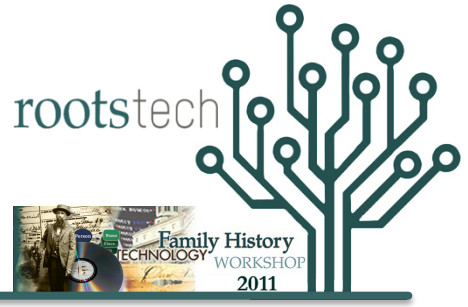
Technology Users

Technology Implementers



Technology Creators

Advances (FHTW today)



- Record Match – Patrick Schone

Evaluation/merge – precision/recall - useable

- Historical Social Networks – Doug Kennard

  History by association (Ancestral Pioneer Past – 1997)

- Record Linkage – Randy Wilson

  Machine Learning Algorithms

  – improves with data, large training sets, features

- Folk Date Pattern – Justin Selliger – user-friendly date range input
- Deep Zoom (Pivot) – Engh, Gehring

  Intuitive, interactive, visually explore, browse, cull large

  data corpus, including photos

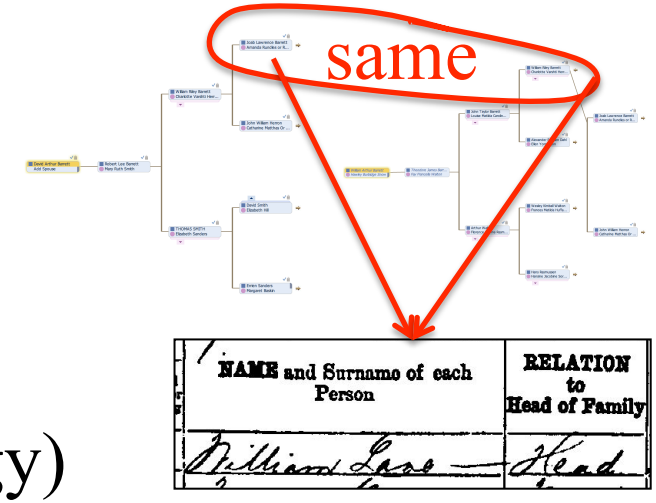
- Preserving, promoting perpetuating innovation, solutions, challenges: "Competition" – Luke Hutchison

Advances (Beau Sharbrough)

- availability of records
- innovative linking of records

2 trees link to same vital record

⇒same person



(Ancestry's "shaky leaf" technology)

- Family History site-building technologies (TNG)
- Knowledge Base of Records

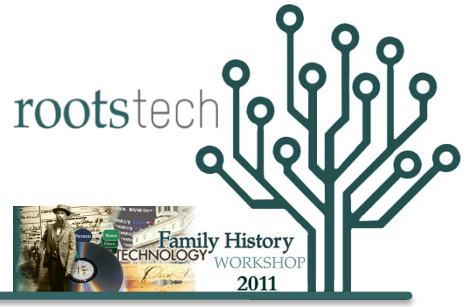
+

Individual Tree Information

→ **Research Task List**

(GenSmarts)

Advances (Dallan Quass)



- Matching people in family trees with records
- Large-scale, accessible infrastructure for collaboration and social networking

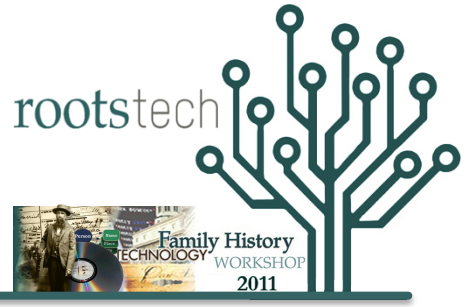
- Facebook, Twitter, Google Wave

- Increased access to genealogical information

- Billions of searchable records (Ancestry, FamilySearch)
- Increase in number of microfilms indexed (“ “)

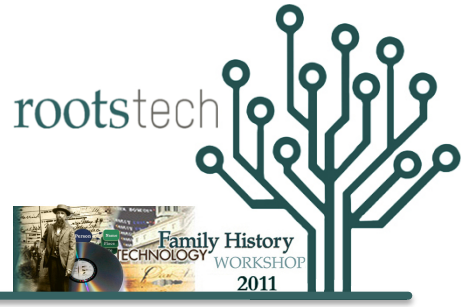
- Genealogy Wiki’s for sharing of information:
  - Wiki.familysearch.org, WeRelate

Advances (Heath Nielson)



- Interactive Genealogical Websites where content can be changed, added to and shared)
- Increased access to source records
- Increased storage capacity and bandwidth enabling the scanning, storing and serving of records across the Internet

Advances (Kirk Duffin)

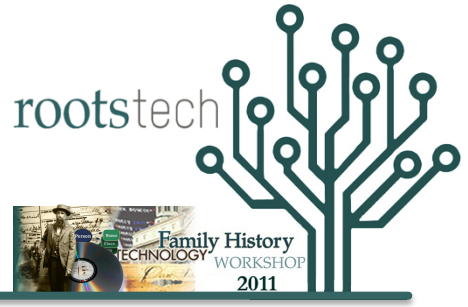


- Online searching and browsing through
  - Hierarchical digital image transmission
  - Automatic scanning and processing of microfilm
  - Associated storage technologies

- Progress in automatic matching / merging

- Human-assisted computing where full automation is not yet possible (e.g. handwriting, speech recognition)

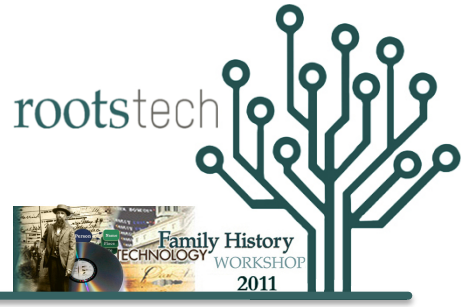
Advances (Jack Reese)



- Scanning, storage, indexing and search technologies:

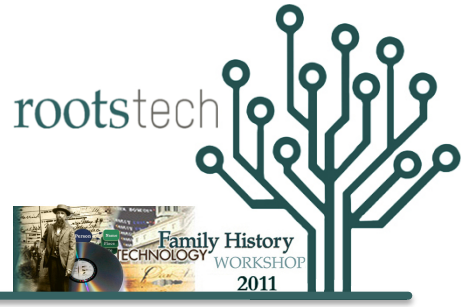
  lowered cost, time => increased productivity, success
- Digital imaging (high-speed, high-res, low-cost replacing – film-based technologies)
- Bandwidth (high-speed, low-cost, last-mile to in-home delivery – replacing trips to archives, libraries, etc.)
- Search engines (aggregate and make searchable collections previously fragmented and distributed around the world)
- Hinting (finds and suggests likely matching records automatically showing connection to existing tree)
- Social networking (collaboration, quickly learn what has been done, less duplication and reinventing the wheel)

Advances (Alan Eaton)



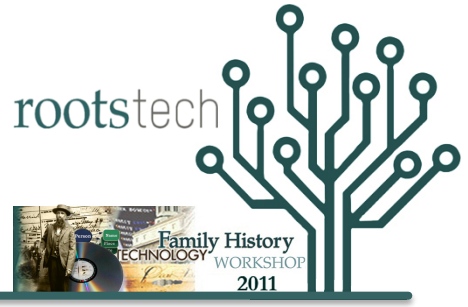
- Single, unified, online family tree wherein everybody's view is sharable and preserved (OGF)
- Images tied to indexes
- Source standardization
- Mass digitization of records
- Indexing utilizing the masses
- Velocity of indexing
- Cloud computing concepts (including GenealogyCloud.com)

Advances (Bill Harten)



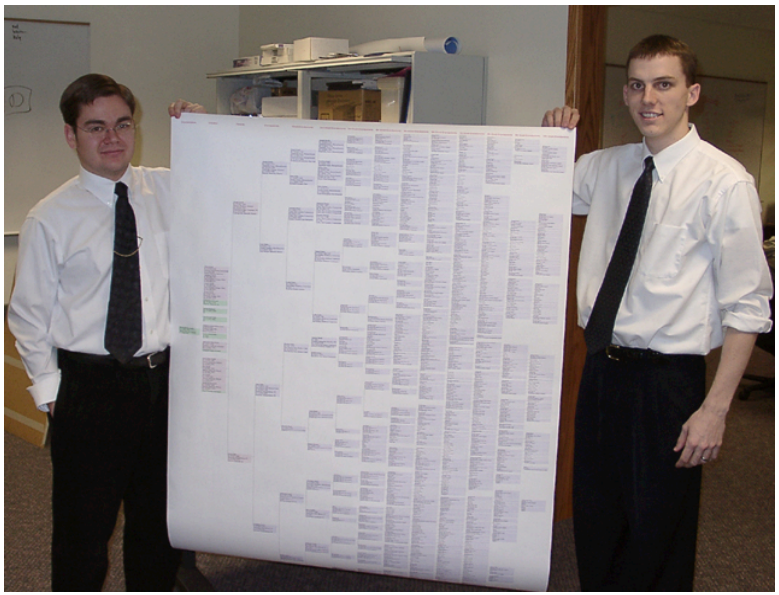
- **Key Advances (That We Have Yet To Take Advantage Of)**
- Affordable storage and delivery infrastructure for evidence images
- The union of LDS official ordinance records with extended linkage
- Research guidance tools to help researchers select the best source for their research objective

Advances (FHTW)



- Transliteration of non-roman text – Lonsdale
- Automating Exposure for Microfilm Scanning
  – Nielson, et al (Scanstone)
- Machine-learning, neural net: Record Linkage, matching
  – Giraud-Carrier, Pixton
- Social Networks for Family History – Giraud-Carrier
- Geo-Location Personal Names – Brown, Londsdale
- WeRelate.org World's largest genealogical wiki – Quass

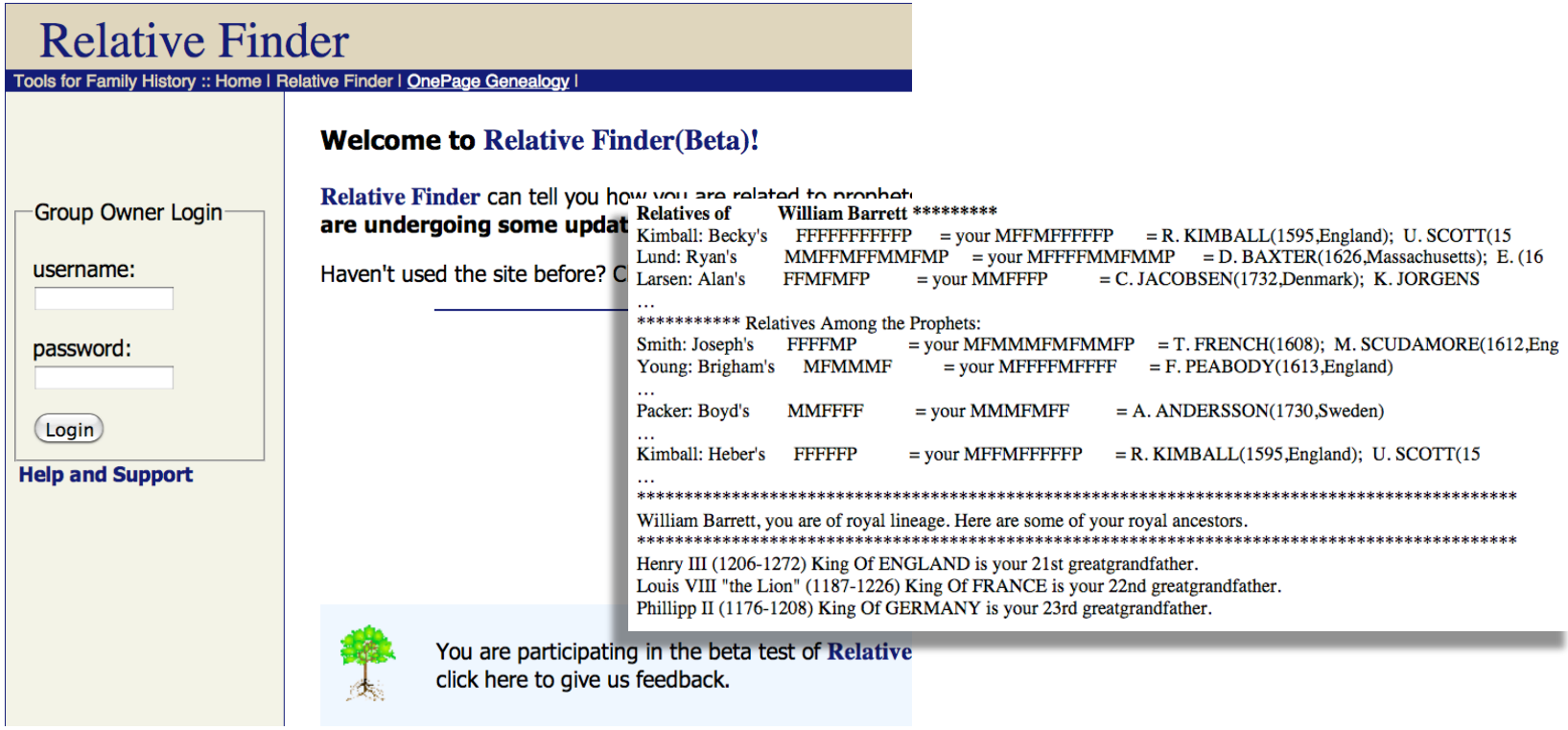
www.onepagegenealogy.com



14-generation pedigree
on 1 page

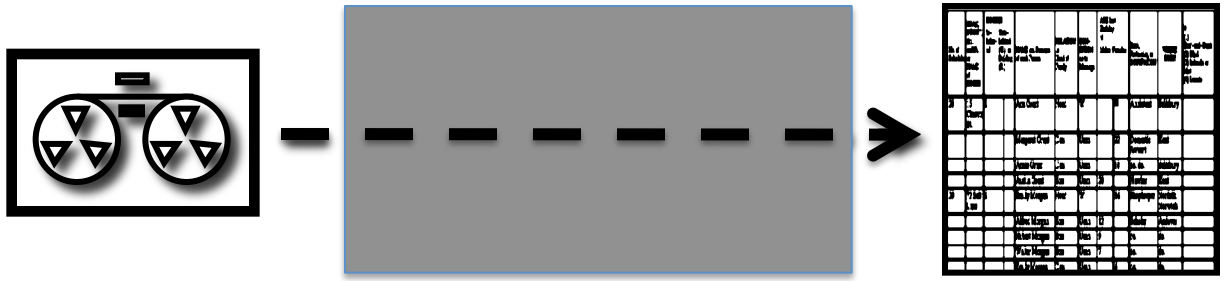
Relative Finder

(http://roots.cs.byu.edu/digroots/)



Find our how you are related to your friends and many other famous people, leaders, kings, queens, etc.

The Digital Microfilm Pipeline*



Film
in

Searchable
Text Out

* Ty Davies, Jake Gehring

The Digital Microfilm Pipeline

rootstech

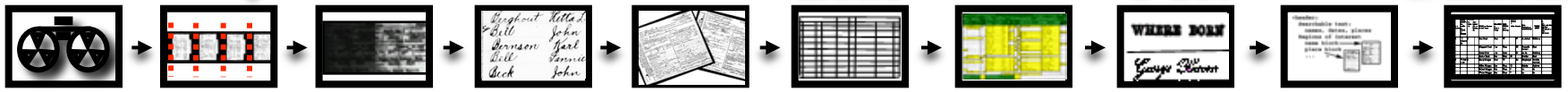
Family History TECHNOLOGY WORKSHOP 2011

What's inside
the box?

What processing steps have to happen
in between?

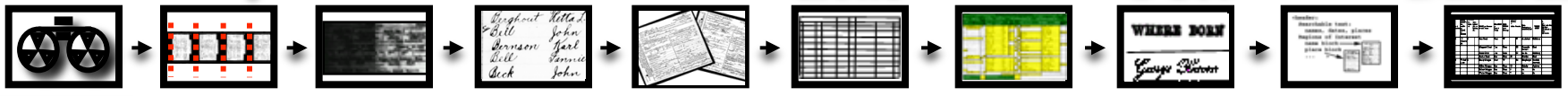
The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Waypointing

Intelligent Indexing

Index/ Search by
- Word
- ROI
- Pattern

Full Text Index/ Search

Intelligent Browsing

Visual Index/Search by Zone, Label, ROI

Scanning the Granite Vault

2.5 Million rolls of microfilm

~ 1300 images per roll

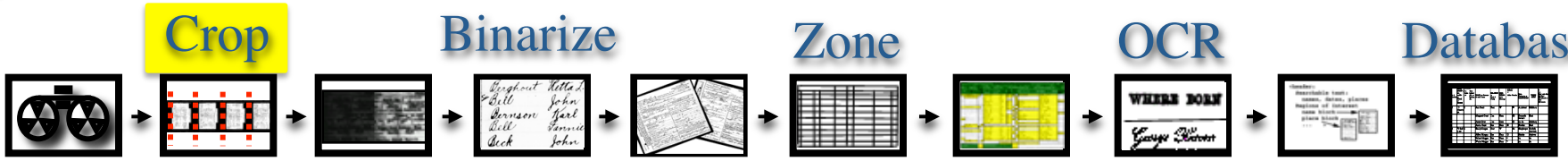
Up to 25 Megabytes per image

1300 x 25,000,000 x 2,500,000

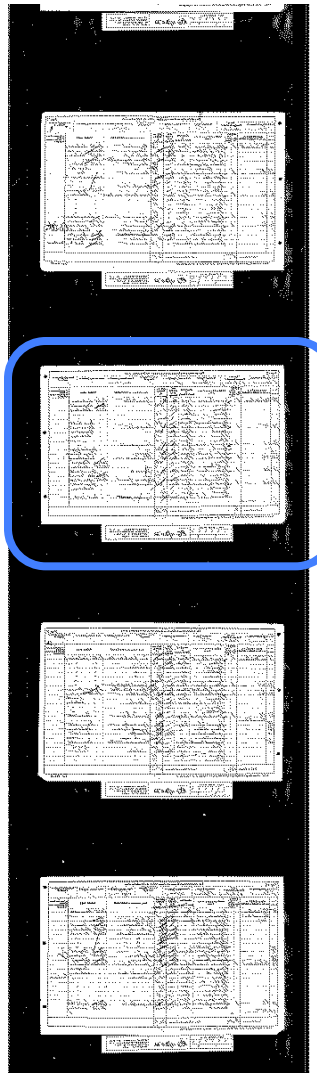
= 81.25 Petabytes

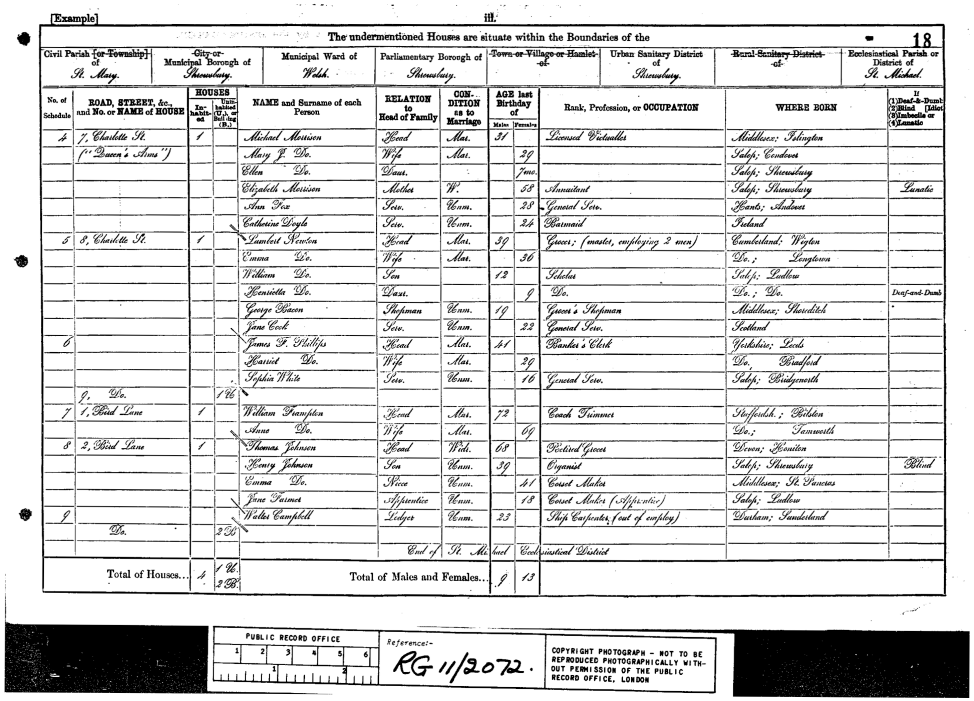
The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Cropping





[Example] iii.

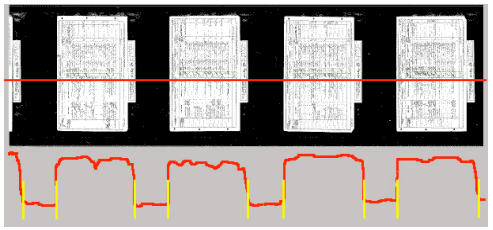
The undermentioned Houses are situate within the Boundaries of the — 18

Civil Parish [or Township] of St. Mary. | City or Municipal Borough of Shrewsbury. | Municipal Ward of Wyle. | Parliamentary Borough of Shrewsbury. | Town or Village or Hamlet of — | Urban Sanitary District of Shrewsbury. | Rural Sanitary District of — | Ecclesiastical Parish or District of St. Michael.

No. of Schedule	ROAD, STREET, &c., and No. or NAME of HOUSE	HOUSES Inhabited	HOUSES Uninhabited (U.), or Building (B.)	NAME and Surname of each Person	RELATION to Head of Family	CONDITION as to Marriage	AGE last Birthday of Males	AGE last Birthday of Females	Rank, Profession, or OCCUPATION	WHERE BORN	If (1) Deaf-&-Dumb (2) Blind (3) Lunatic (4) Imbecile or (5) Idiot
4	7, Charlotte St. ("The Queen's Arms")	1		Michael Morrison	Head	Mar.	51		Licensed Victualler	Middlesex; Islington	
				Mary J. Do.	Wife	Mar.		29		Salop; Condover	
				Ellen Do.	Daur.			7 mo.		Salop; Shrewsbury	
				Elizabeth Morrison	Mother	W.		53	Annuitant	Salop; Shrewsbury	Lunatic
				Ann Fox	Serv.	Unm.		23	General Serv.	Wants; Andover	
				Catherine Doyle	Serv.	Unm.		24	Barmaid	Ireland	
5	8, Charlotte St.	1		Lambert Newton	Head	Mar.	39		Grocer; (masters, employing 2 men)	Cumberland; Wigton	
				Emma Do.	Wife	Mar.		36		Do.; Longtown	
				William Do.	Son		12		Scholar	Salop; Ludlow	
				Henrietta Do.	Daur.			9	Do.	Do.; Do.	Deaf-and-Dumb
				George Bacon	Shopman	Unm.	19		Grocer's Shopman	Middlesex; Shoreditch	
				Jane Cook	Serv.	Unm.		22	General Serv.	Scotland	
6				James T. Phillips	Head	Mar.	41		Banker's Clerk	Yorkshire; Leeds	
				Harriet Do.	Wife	Mar.		29		Do. Bradford	
				Sophia White	Serv.	Unm.		16	General Serv.	Salop; Bridgenorth	
	9, Do.		1 H.								
7	1, Bird Lane	1		William Hampton	Head	Mar.	72		Coach Trimmer	Staffordsh.; Bilston	
				Anne Do.	Wife	Mar.		69		Do.; Tamworth	
8	2, Bird Lane	1		Thomas Johnson	Head	Wid.	68		Retired Grocer	Devon; Honiton	
				Fanny Johnson	Son	Unm.	39		Organist	Salop; Shrewsbury	Blind
				Emma Do.	Niece	Unm.		41	Corset Maker	Middlesex; St. Pancras	
				Jane Farmer	Apprentice	Unm.		18	Corset Maker (Apprentice)	Salop; Ludlow	
9				Walter Campbell	Lodger	Unm.	23		Ship Carpenter, (out of employ)	Durham; Sunderland	
	Do.		2 B.								
						End of St. Michael Ecclesiastical District					
	Total of Houses...	4	1 U. 2 B.	Total of Males and Females...			9	13			

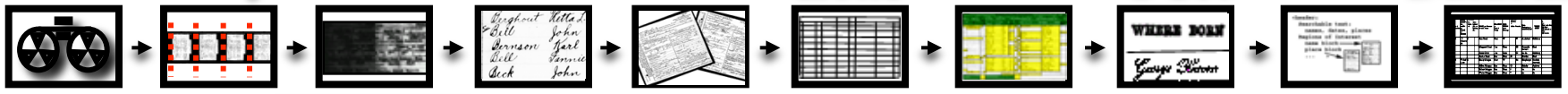
PUBLIC RECORD OFFICE Reference: RG 11/2072. COPYRIGHT PHOTOGRAPH – NOT TO BE REPRODUCED PHOTOGRAPHICALLY WITHOUT PERMISSION OF THE PUBLIC RECORD OFFICE, LONDON

- Scan continuously in Ribbons →
- Crop in batch, exploit consensus →

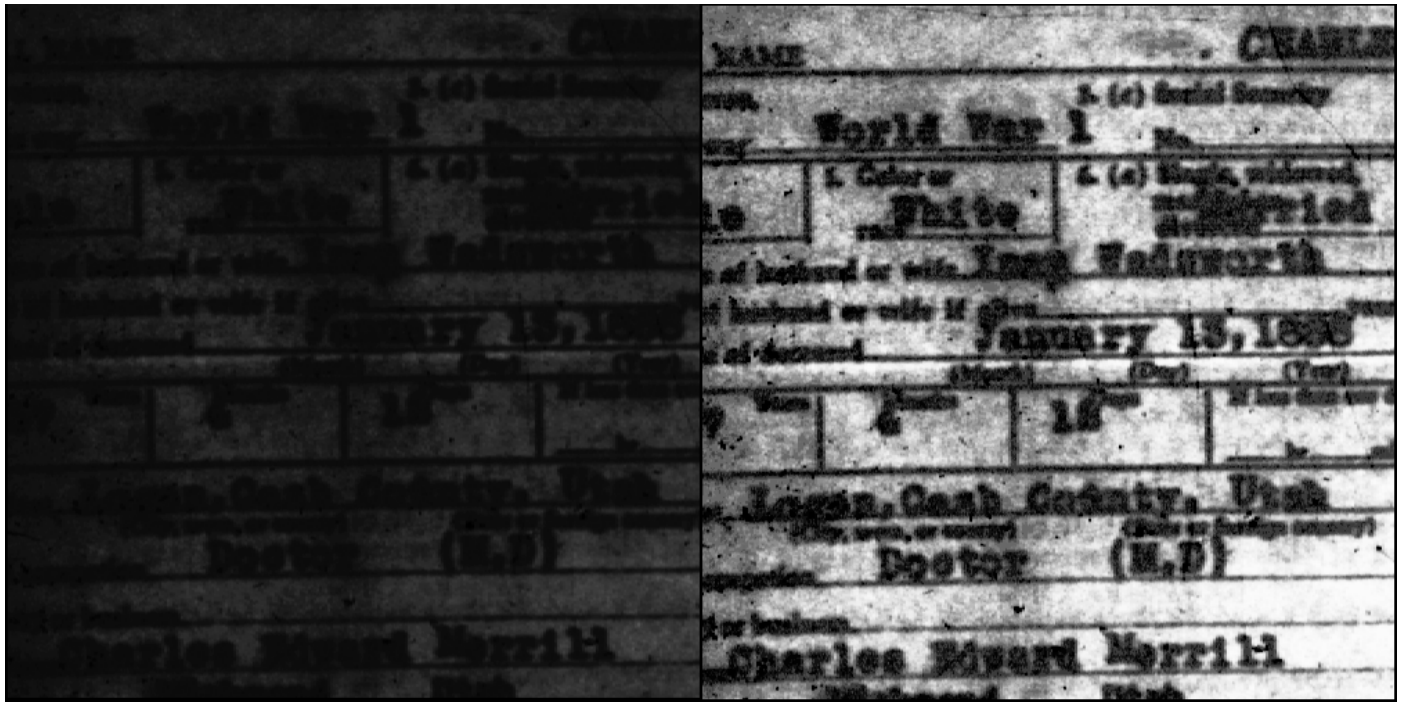


The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011

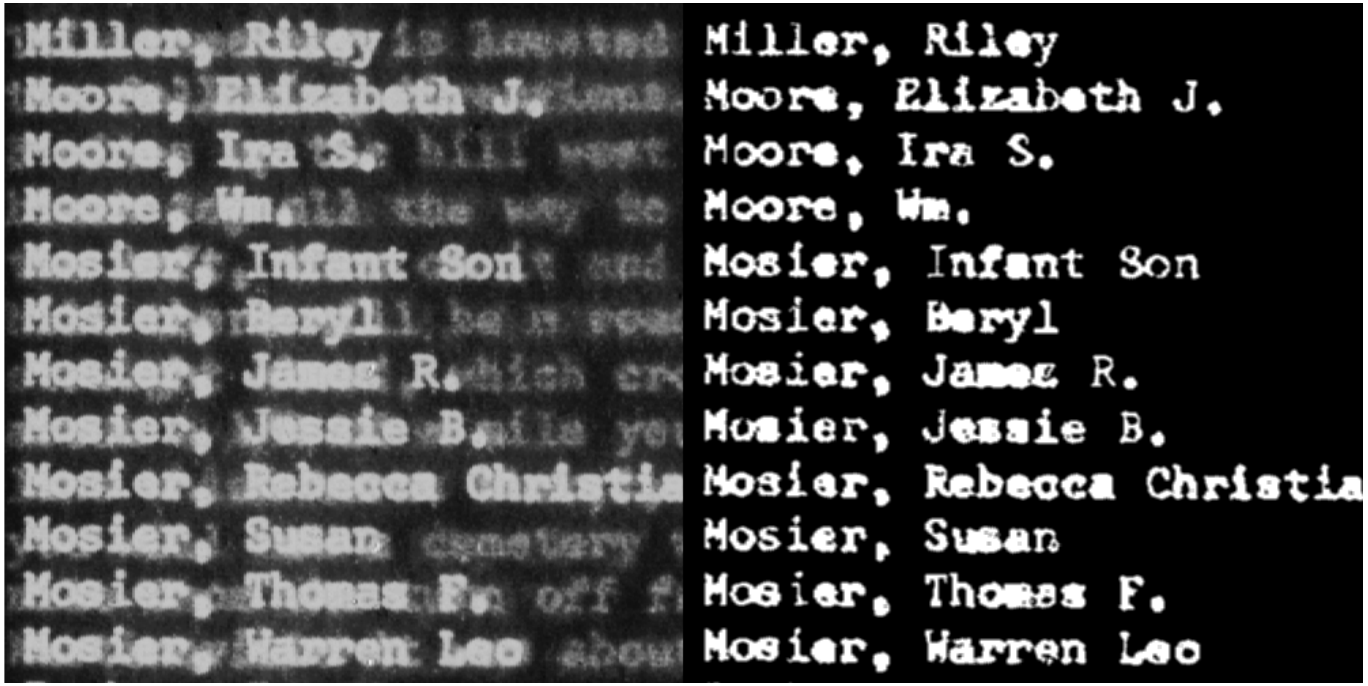


Histogram Enhancement



(Left) Original scanned record (Right) After Enhancement

Threshold Enhancement



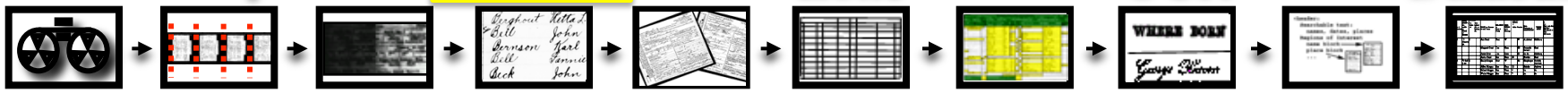
(Left) Original scanned record (Right) After thresholding for Bleed-through removal, with Enhancement, and Antialiasing

Stroke Enhancement

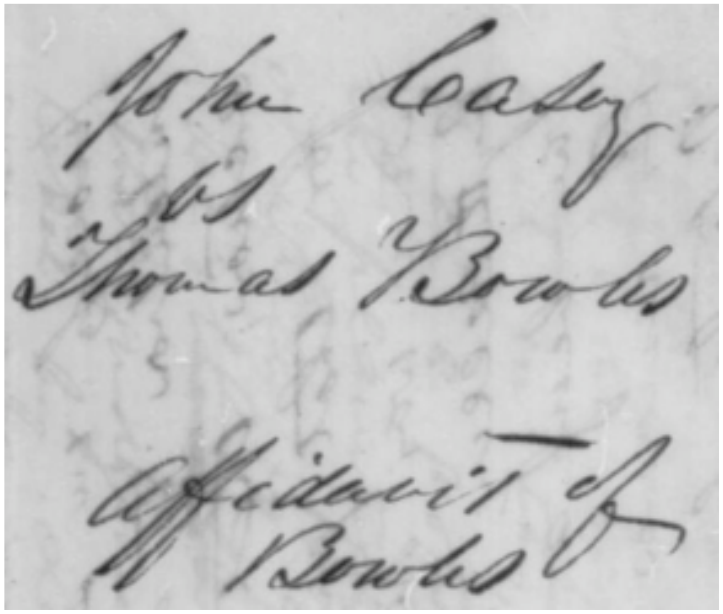
(Left) Original scanned record (Right) After Matched Filter Stroke Enhancement

The Digital Microfilm Pipeline

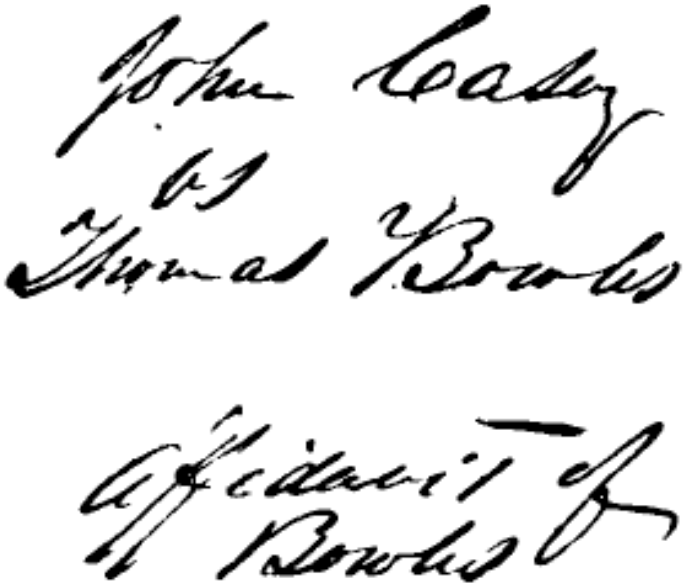
rootstech
Family History TECHNOLOGY WORKSHOP 2011



Binarization



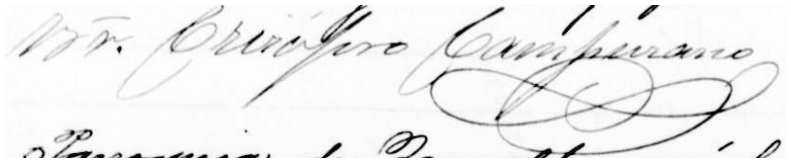
(a) Original Image



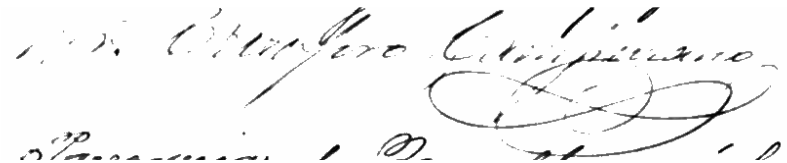
(b) Binarized Result

Median Filter Subtraction for background elimination

Binarization



(a) Original with background removed



(b) First set of text pixels (S_1)

(c) Second set of text pixels (S_2)

(d) Third set of text pixels (S_3)

.

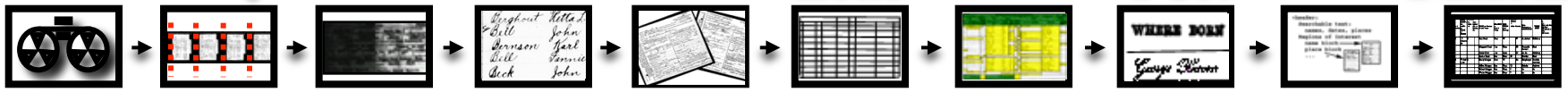
.

(f) Result ($S_1 + S_2 + S_3 + S_4$)

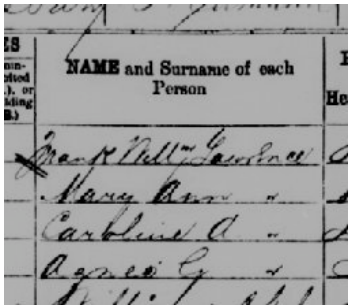
Recursive Otsu Intensity Parsing and Accumulation

The Digital Microfilm Pipeline

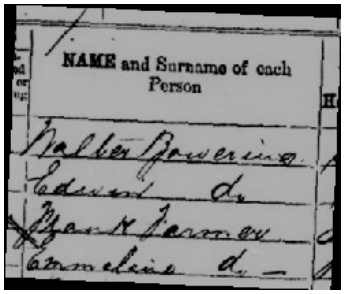
rootstech
Family History TECHNOLOGY WORKSHOP 2011



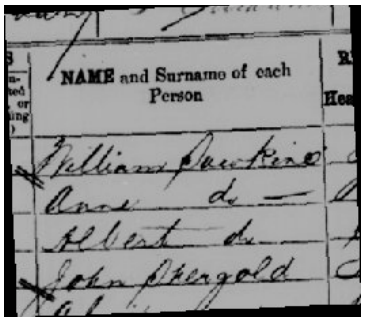
Registration



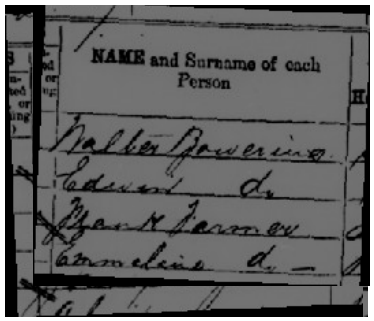
+



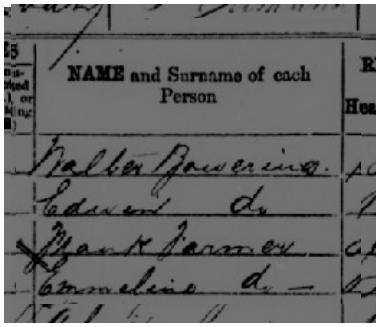
+



=



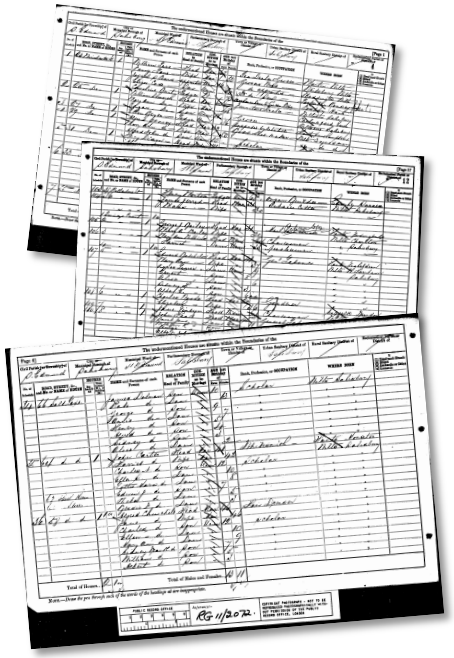
Not registered



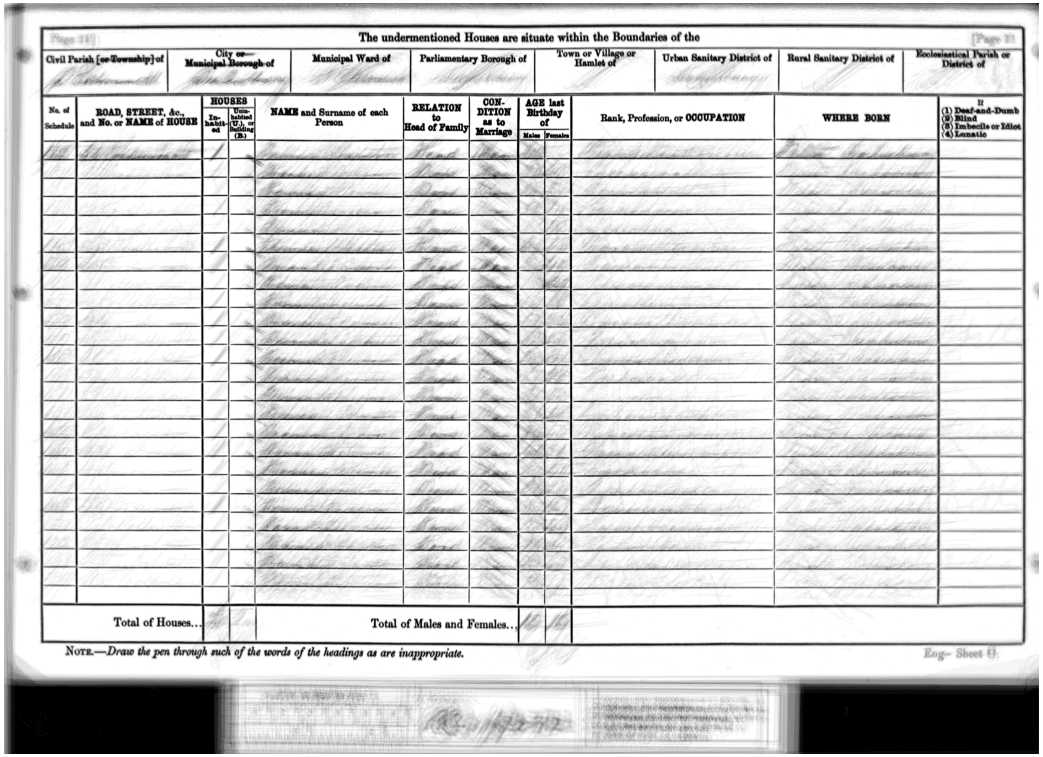
Registered

Documents must be aligned, rotated and scaled to register them

Registration



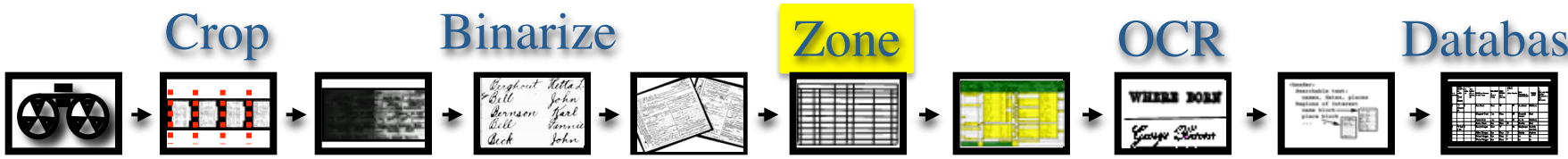
Register



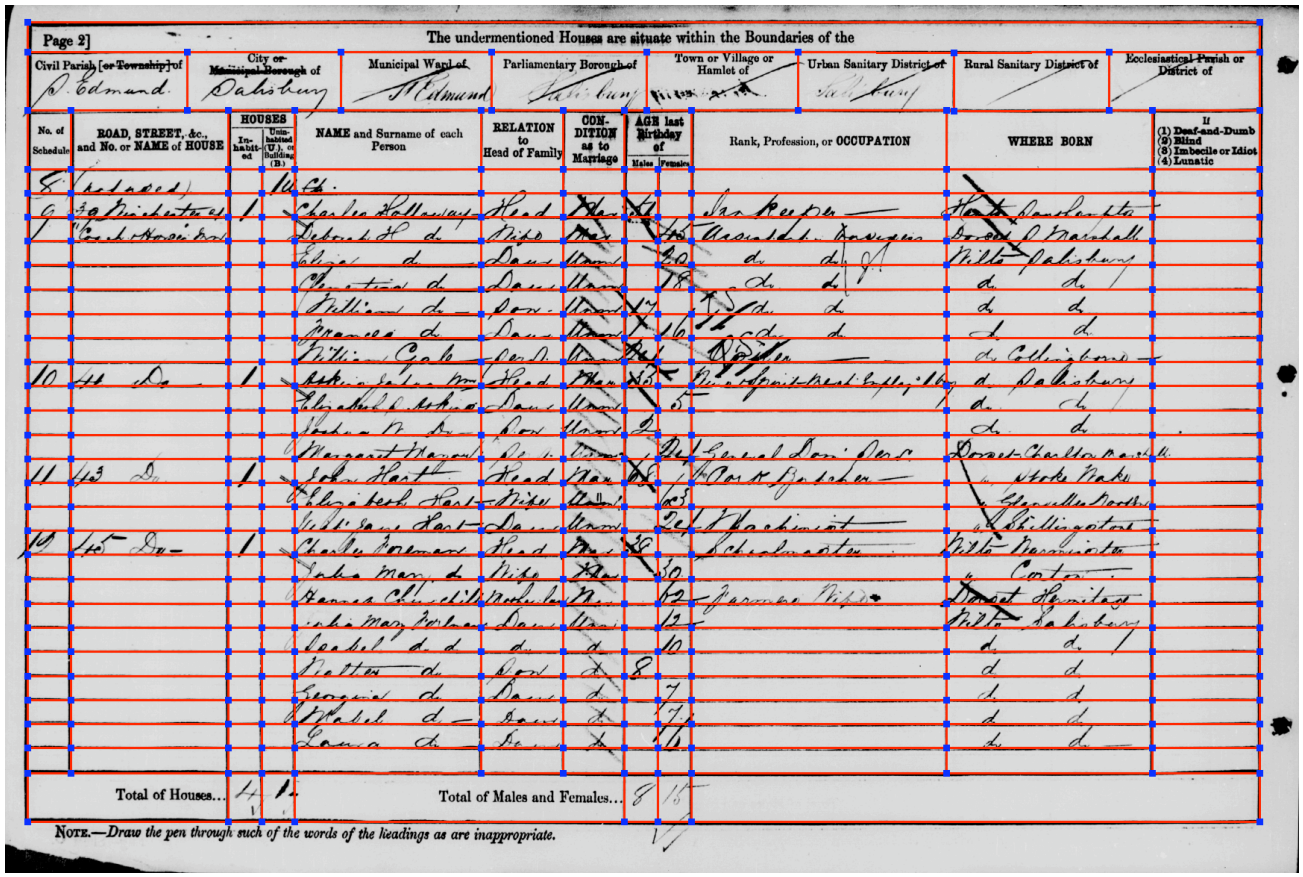
Alignment, rotation and scaling using the Fourier-Mellin Transform allows registration to sub-degree and sub-pixel accuracy.

The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011

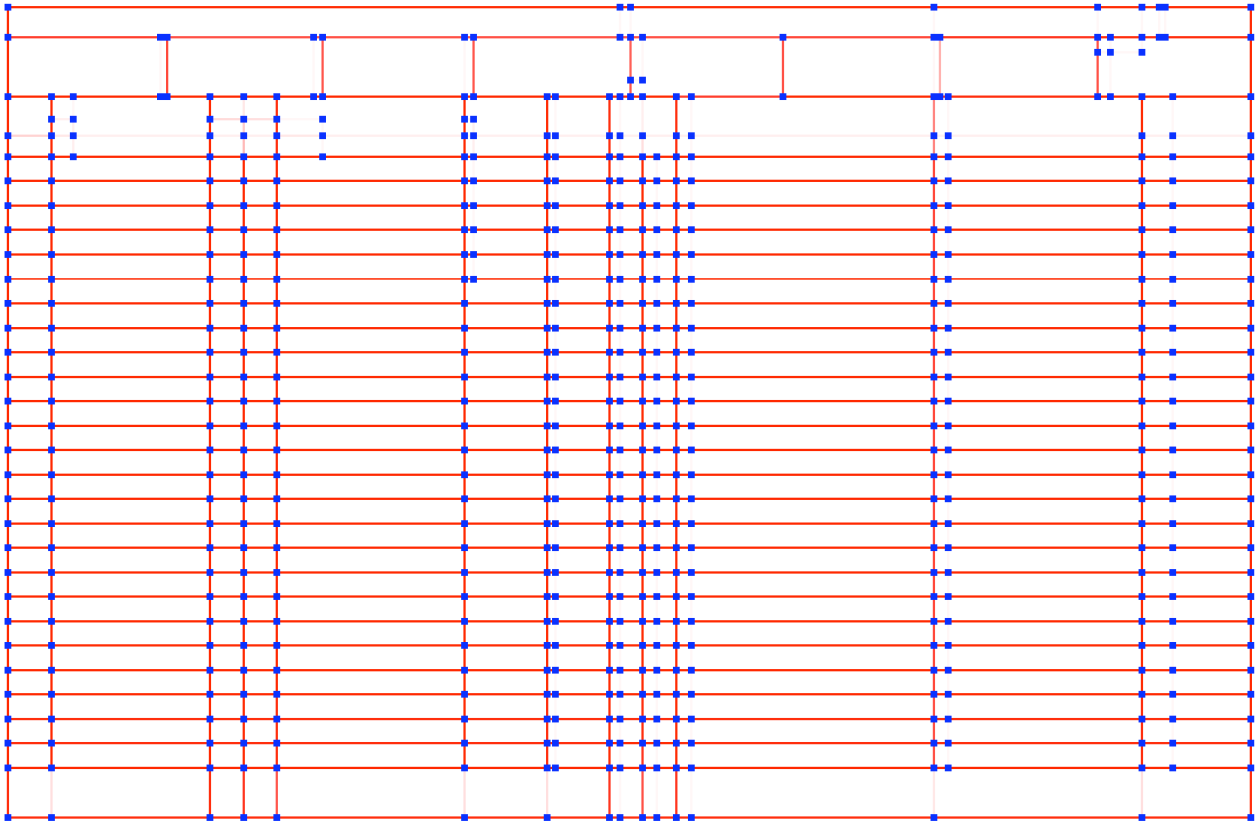


Zoning



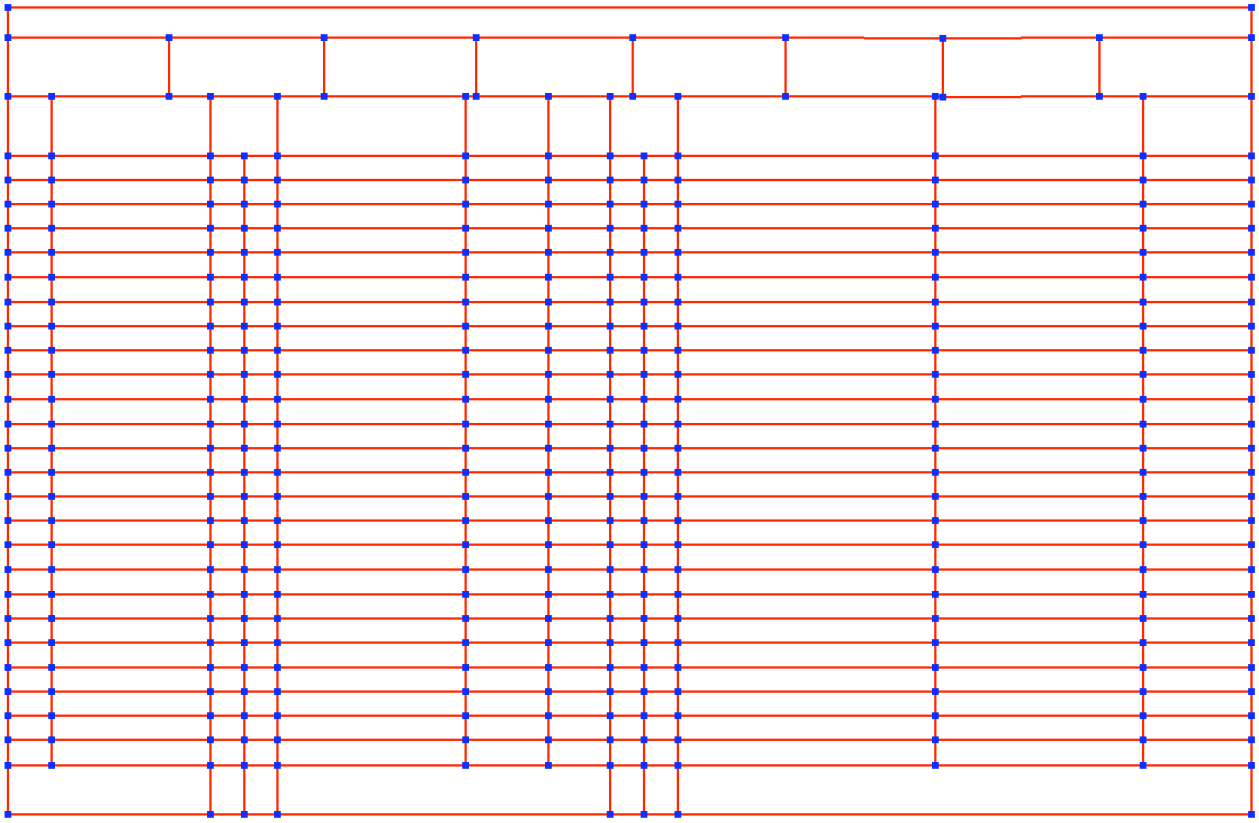
Line detection and local refinement for document zoning

Zoning by Consensus



Consensus from 28 frames

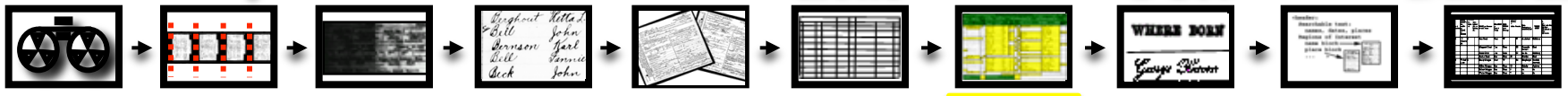
Zoning by Consensus



Final Zone Template

The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Waypointing

Intelligent Indexing

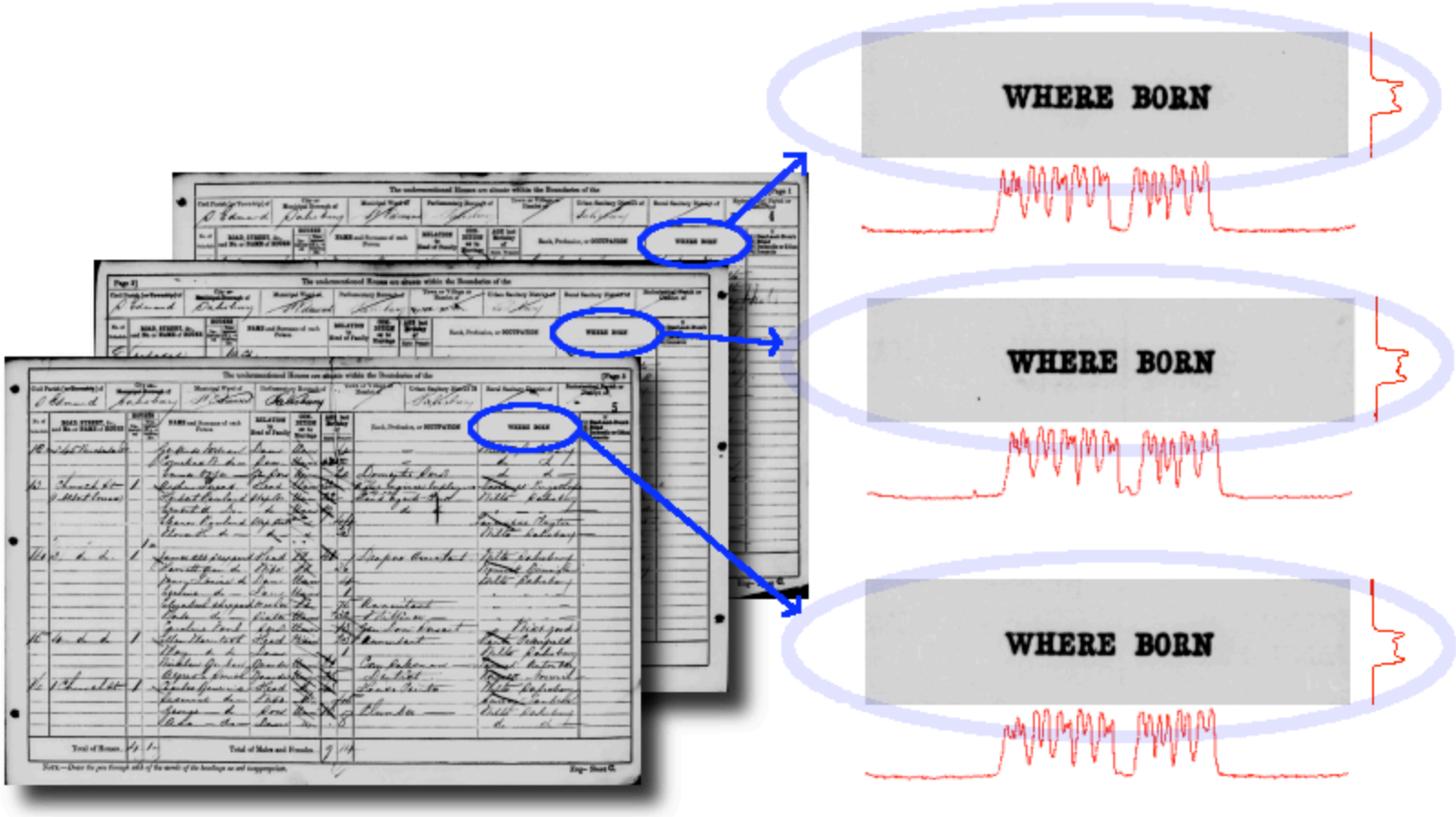
Index/ Search by
- Word
- ROI
- Pattern

Full Text Index/ Search

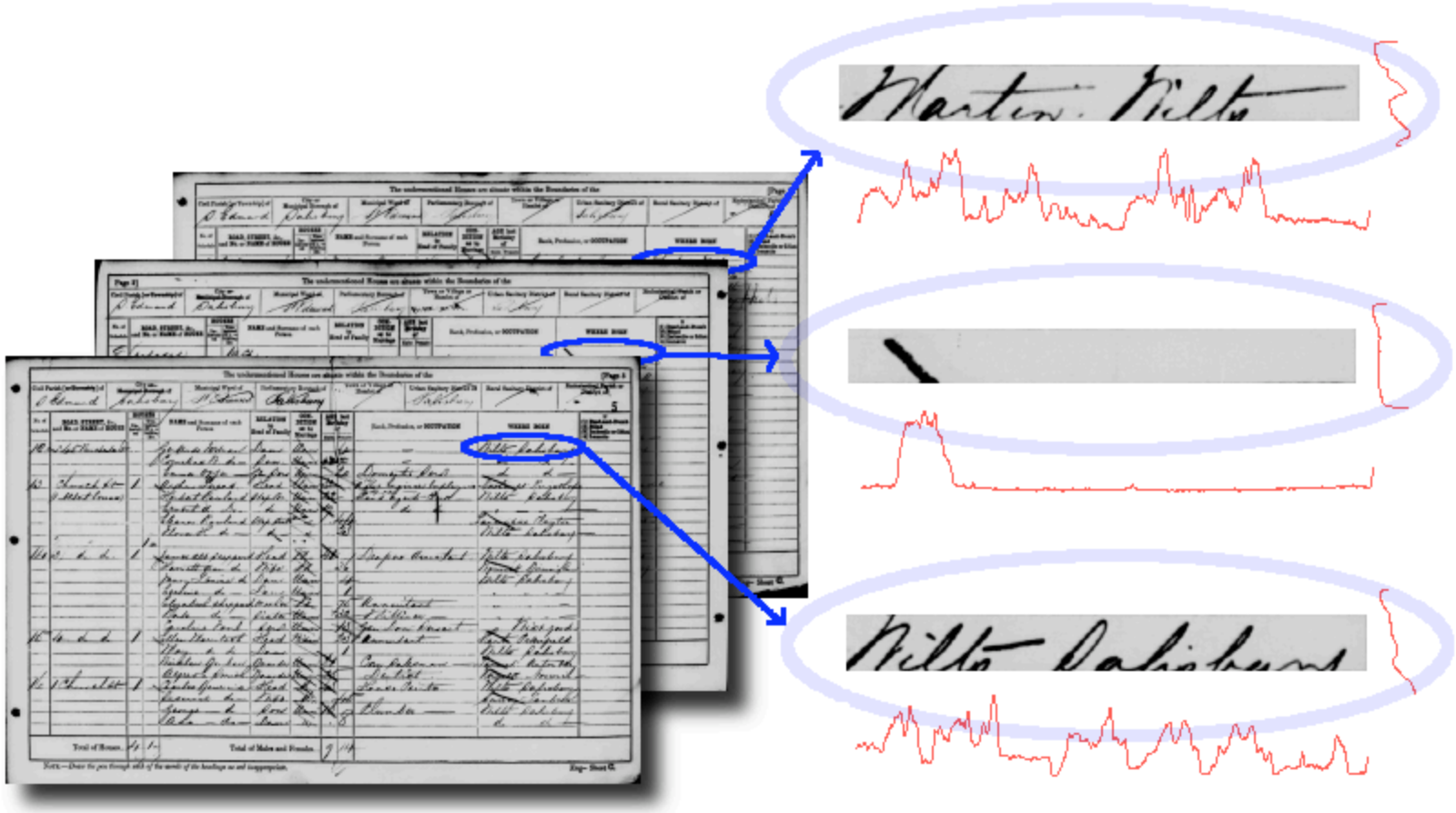
Intelligent Browsing

Visual Index/Search by Zone, Label, ROI

Label Machine Print



Label Handwriting



Labeled Image

rootstech Family History TECHNOLOGY WORKSHOP 2011

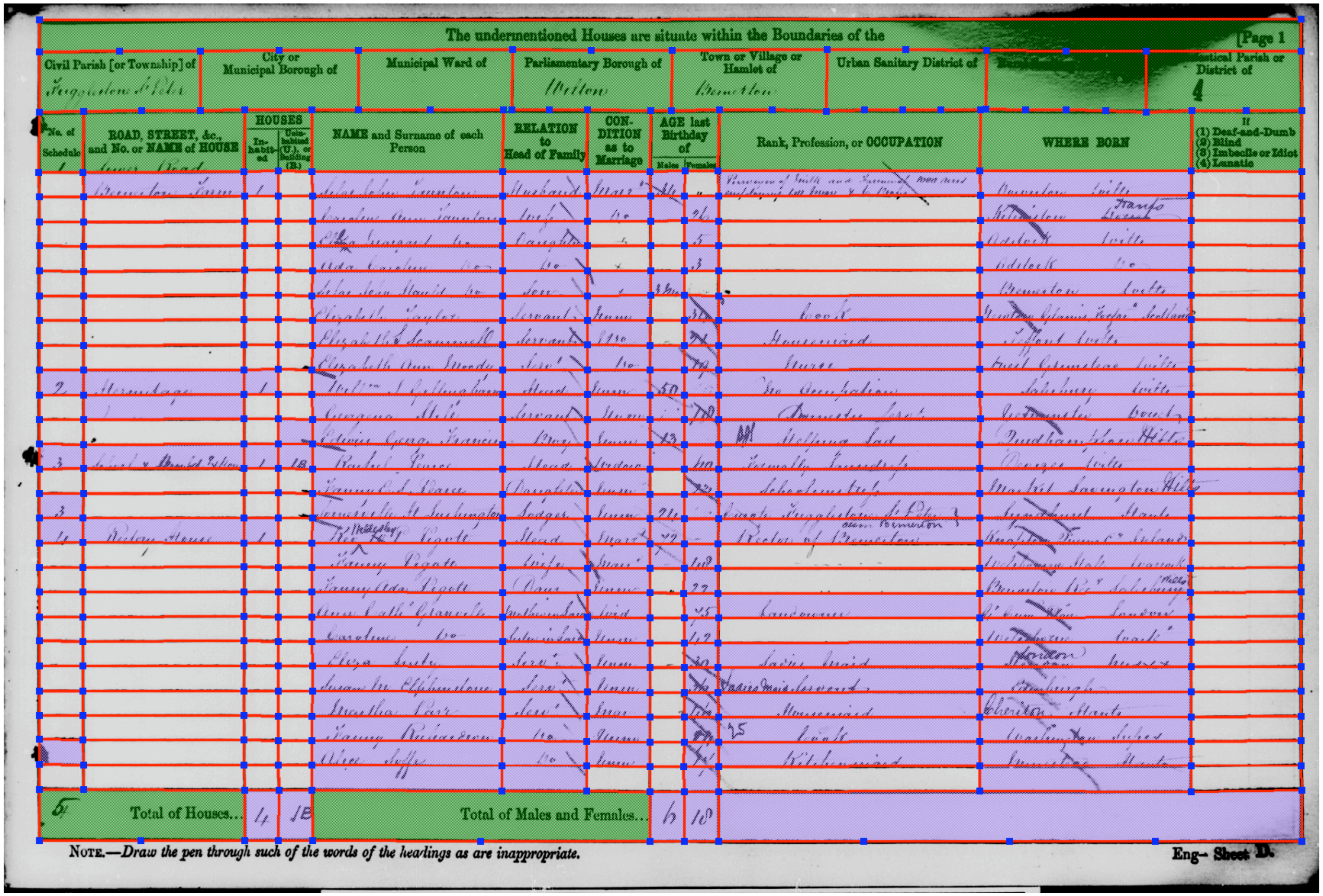
[Example] iii.

The undermentioned Houses are situate within the Boundaries of the — 18

Civil Parish [or Township] of	City or Municipal Borough of	Municipal Ward of	Parliamentary Borough of	Town or Village or Hamlet of	Urban Sanitary District of	Rural Sanitary District of	Ecclesiastical Parish or District of
St. Mary.	Shrewsbury.	Welsh.	Shrewsbury.		Shrewsbury.		St. Michael.

No. of Schedule	ROAD, STREET, &c., and No. or NAME of HOUSE	HOUSES Inhabited	HOUSES Uninhabited (U.), or Building (B.)	NAME and Surname of each Person	RELATION to Head of Family	CONDITION as to Marriage	AGE last Birthday of Males	AGE last Birthday of Females	Rank, Profession, or OCCUPATION	WHERE BORN	If (1) Deaf-&-Dumb (2) Blind (3) Imbecile or (4) Lunatic
4	7, Charlotte St.	1		Michael Morrison	Head	Mar.	31		Licensed Victualler	Middlesex; Islington	
	("Queen's Arms")			Mary F. Do.	Wife	Mar.		29		Salop; Condover	
				Ellen Do.	Daur.			7 mo		Salop; Shrewsbury	
				Elizabeth Morrison	Mother	W.		58	Annuitant	Salop; Shrewsbury	Lunatic
				Ann Fox	Serv.	Unm.		28	General Serv.	Hants; Andover	
				Catherine Doyle	Serv.	Unm.		24	Barmaid	Ireland	
5	8, Charlotte St.	1		Lambert Newton	Head	Mar.	39		Grocer; (master, employing 2 men)	Cumberland; Wigton	
				Emma Do.	Wife	Mar.		36		Do.; Longtown	
				William Do.	Son		12		Scholar	Salop; Ludlow	
				Henrietta Do.	Daur.			9	Do.	Do.; Do.	Deaf-and-Dumb
				George Bacon	Shopman	Unm.	19		Grocer's Shopman	Middlesex; Shoreditch	
				Jane Cook	Serv.	Unm.		22	General Serv.	Scotland	
6				James F. Phillips	Head	Mar.	41		Banker's Clerk	Yorkshire; Leeds	
				Harriet Do.	Wife	Mar.		29		Do. Bradford	
				Sophia White	Serv.	Unm.		16	General Serv.	Salop; Bridgenorth	
	9, Do.		1 B								
7	1, Bird Lane	1		William Frampton	Head	Mar.	72		Coach Trimmer	Staffordsh.; Bilston	
				Anne Do.	Wife	Mar.		69		Do.; Tamworth	
8	2, Bird Lane	1		Thomas Johnson	Head	Wid.	68		Retired Grocer	Devon; Honiton	
				Henry Johnson	Son	Unm.	39		Organist	Salop; Shrewsbury	Blind
				Emma Do.	Niece	Unm.		41	Corset Maker	Middlesex; St. Pancras	
				Jane Farmer	Apprentice	Unm.		18	Corset Maker (Apprentice)	Salop; Ludlow	
9				Walter Campbell	Lodger	Unm.	23		Ship Carpenter (out of employ)	Durham; Sunderland	
	Do.		2 B								
					End of	St. Michael	Ecclesiastical District				
Total of Houses...		4	1 U 2 B	Total of Males and Females...			9	13			

Zoned and Labeled Image


The undermentioned Houses are situate within the Boundaries of the [Page 1

Civil Parish [or Township] of Fugglestone St Peter | City or Municipal Borough of | Municipal Ward of | Parliamentary Borough of Wilton | Town or Village or Hamlet of Bemerton | Urban Sanitary District of | ...stical Parish or District of

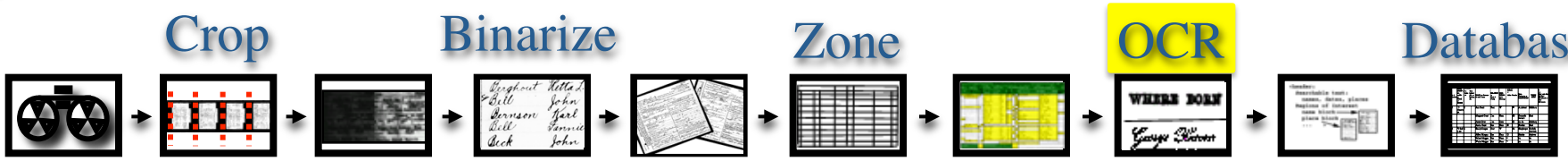
No. of Schedule	ROAD, STREET, &c., and No. or NAME of HOUSE	HOUSES In-habit-ed	HOUSES Unin-habited (U.), or Building (B.)	NAME and Surname of each Person	RELATION to Head of Family	CON-DITION as to Marriage	AGE last Birthday of Males	AGE last Birthday of Females	Rank, Profession, or OCCUPATION	WHERE BORN	If (1) Deaf-and-Dumb (2) Blind (3) Imbecile or Idiot (4) Lunatic
	Total of Houses...	4	1B	Total of Males and Females...			6	18			

NOTE.—*Draw the pen through such of the words of the headings as are inappropriate.*

Eng- Sheet D.

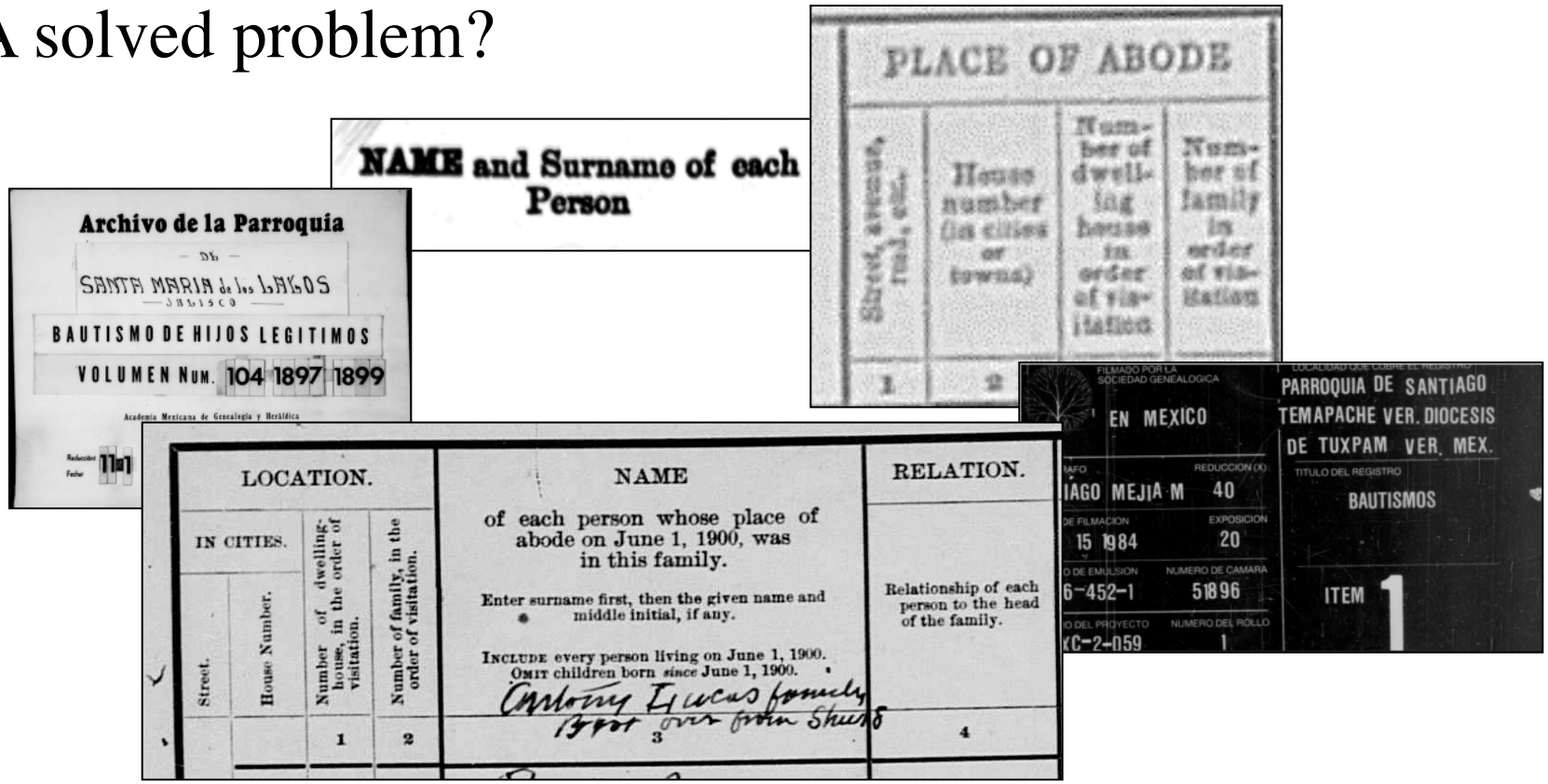
The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Machine-Printed OCR

A solved problem?

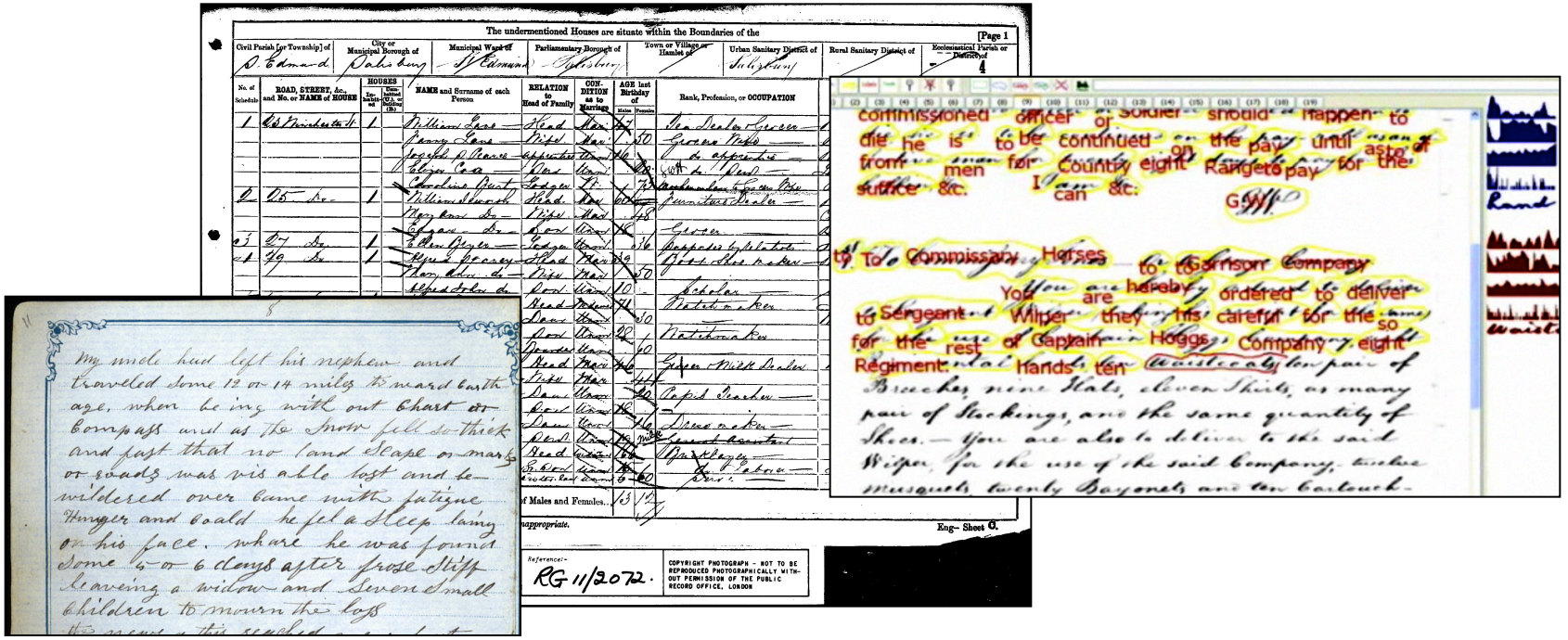


Sort of … depending on resolution, clarity, noise, etc.

Handwriting Recognition

rootstech Family History TECHNOLOGY WORKSHOP 2011

Definitely ***NOT*** a solved problem

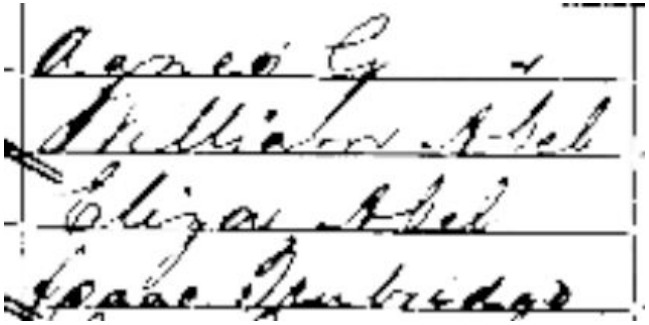


Cursive, off-line handwriting recognition is an area of active research

Handwriting Recognition Challenges/Opportunities

rootstech Family History TECHNOLOGY WORKSHOP 2011

- Image quality is not always good
  - Fading of documents/acetate-based microfilm
  - Original document written with quill-pen



- But documents were usually written meticulously
  - Older handwriting more regular, so simpler to match than modern handwriting
  - Different approach required (e.g. stroke ordering problem)

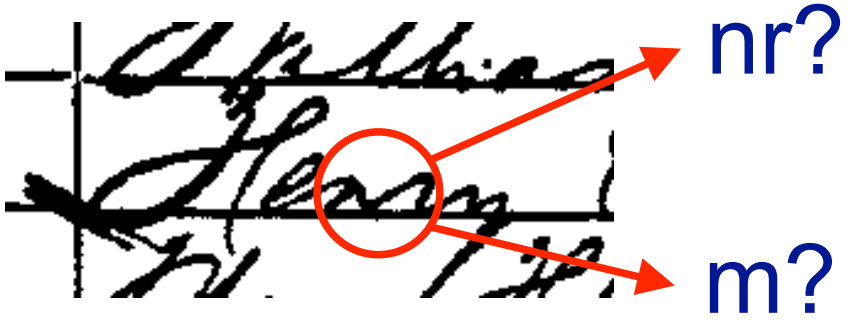
Handwriting Recognition
Past Approaches

rootstech

Family History TECHNOLOGY WORKSHOP 2011

- An example of some steps in the recognition process:
  - *Handwriting style clustering*

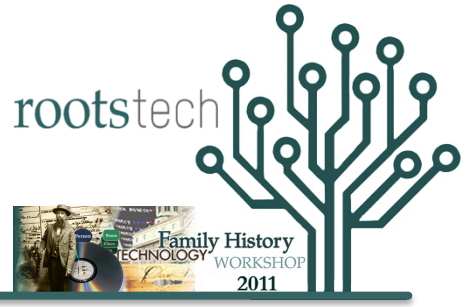
  - *Letter recognition*



  - *Approximate string matching*

Smith
Smythe

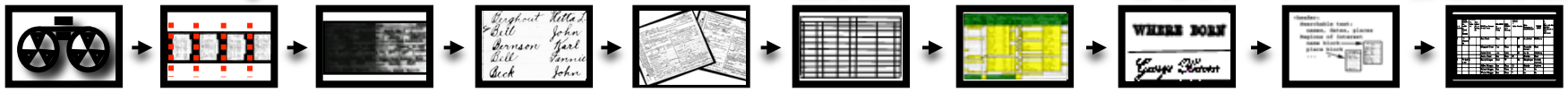
Handwriting Recognition Future Approaches



- Exploit manually-indexed, ground-truthed data to train algorithms
- Training may be required for each new handwriting style / time period
- (Semi-) Automated extraction system: Intelligent assistant to indexer, or use as "Indexer B."
- Use image morphing/pattern recognition: don't have to recognize explicitly, only implicitly
- Use contextual information to increase confidence

The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Waypointing

Intelligent Indexing

Index/ Search by
- Word
- ROI
- Pattern

Full Text Index/ Search

Intelligent Browsing

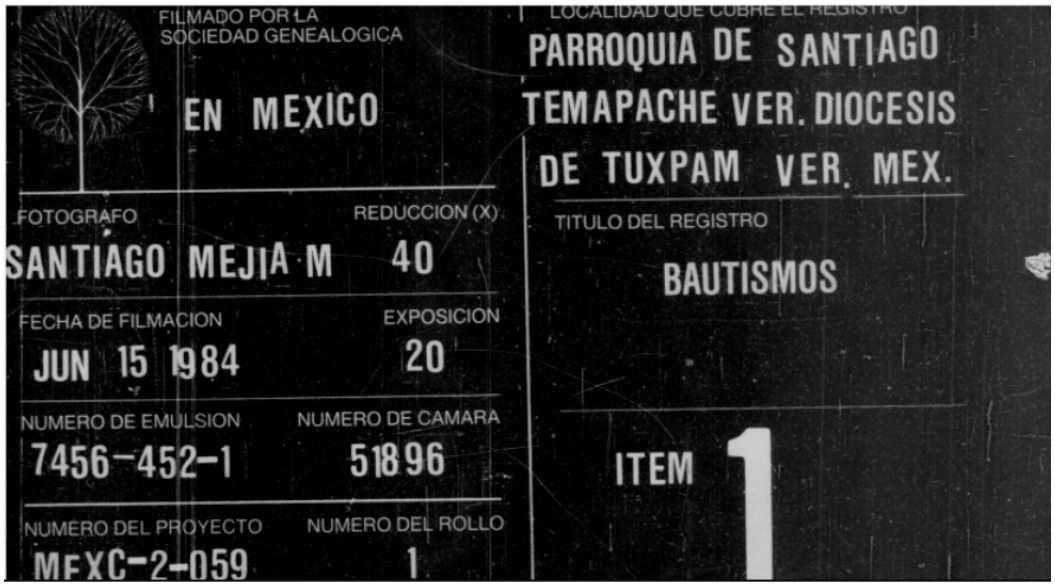
Visual Index/Search by Zone, Label, ROI

Automated Waypointing:
Extract content from titleboards*

rootstech Family History TECHNOLOGY WORKSHOP 2011

Focus on key fields:

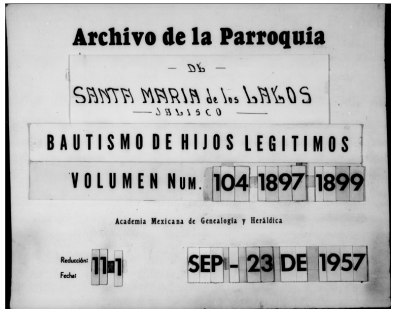
- Record type
- Place
- Date range
- Repository
- Film number



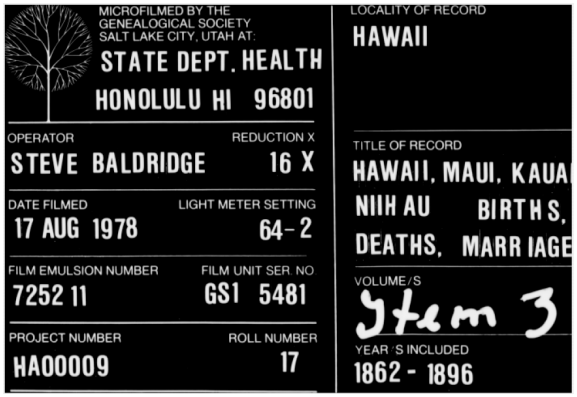
OCR/Index the text

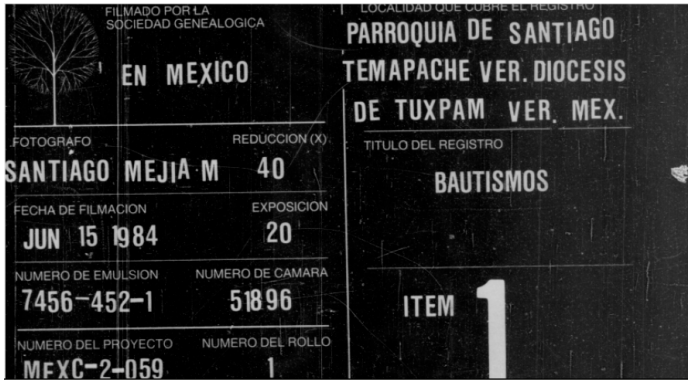
* David Ouimette, Jake Gehring

Automated Waypointing:
Group titleboards uniformly



. . .





. . .

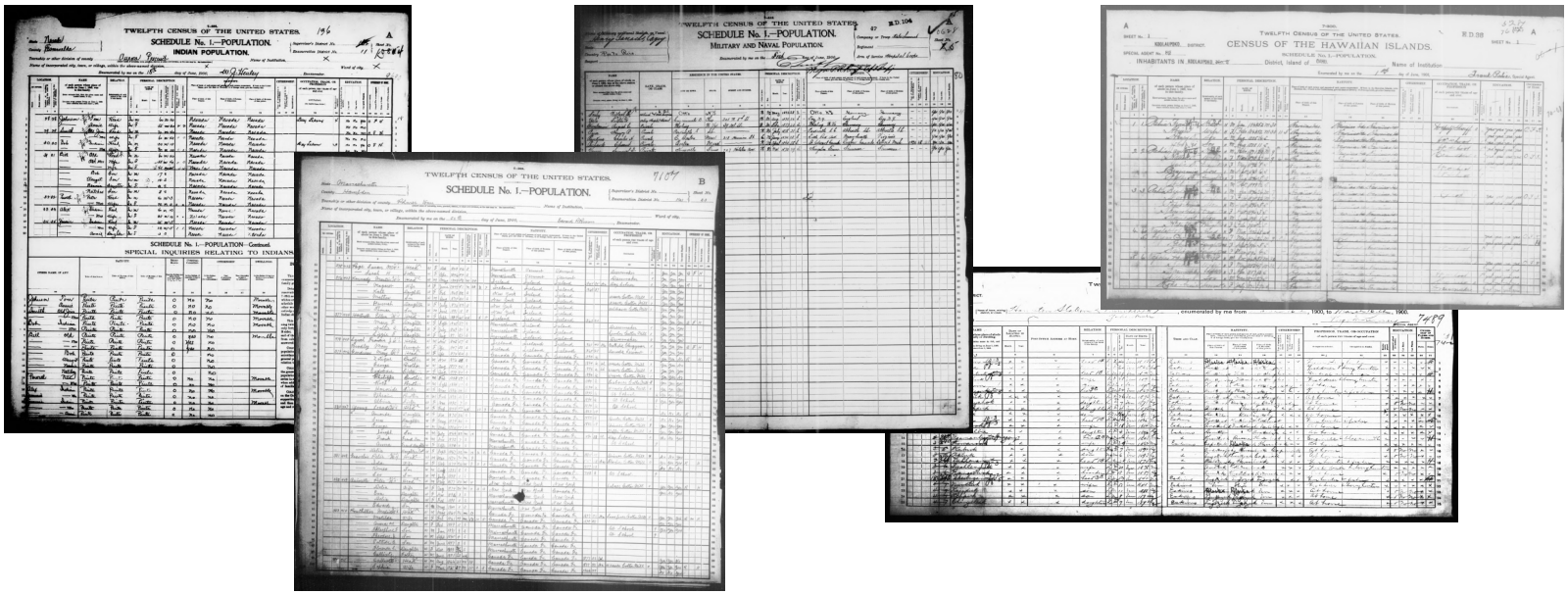
To facilitate collection analysis and indexing prep

Automated Waypointing:

Classify images that look alike

rootstech Family History TECHNOLOGY WORKSHOP 2011

Example 1: There are five distinct forms used in the 1900 census. They need to be treated differently when indexing

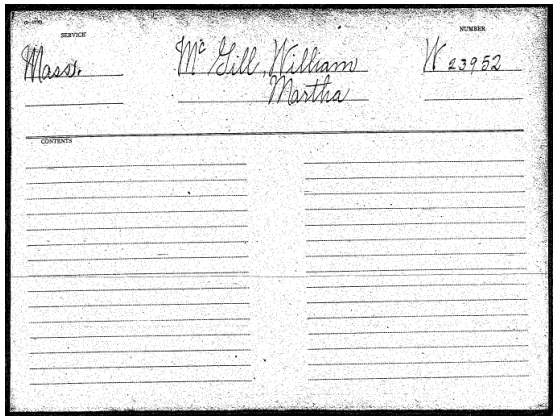


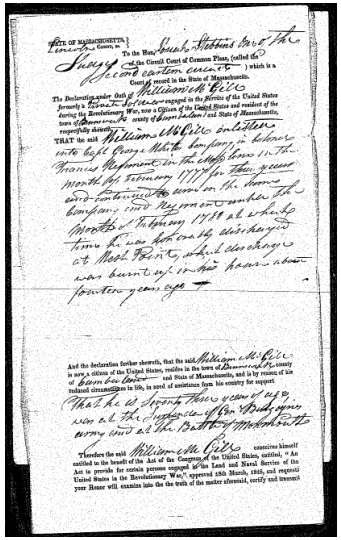
Tool for index batching and automated waypointing

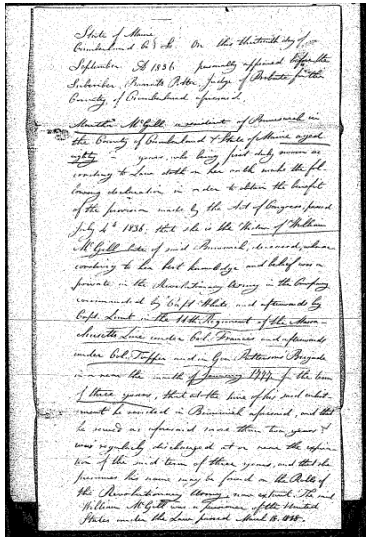
Automated Waypointing:
Classify images that look alike

rootstech Family History TECHNOLOGY WORKSHOP 2011

- Tool for waypointing and index batching
- Example 2: Create brief indexes (e.g., Rev War pensions)
  - Isolate images to send to indexers (e.g., a single card image which precedes dozens of pages in a pension packet)
  - Waypoint entire set of images, including indexed images

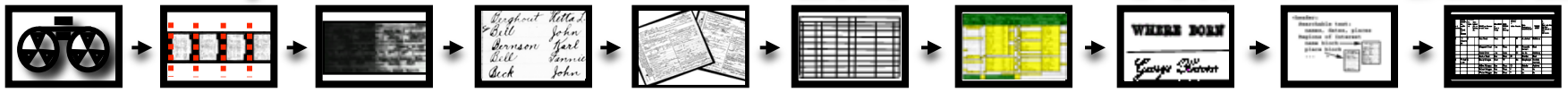






The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Crop
Binarize
Zone
OCR
Database
Scan
Enhance
Register
Label
Text/
Bitmaps
Waypointing
Intelligent
Indexing
Index/
Search
by
• Word
• ROI
• Pattern
Full
Text
Index/
Search
Intelligent Browsing
Visual Index/Search by Zone, Label, ROI

Intelligent Indexing

rootstech

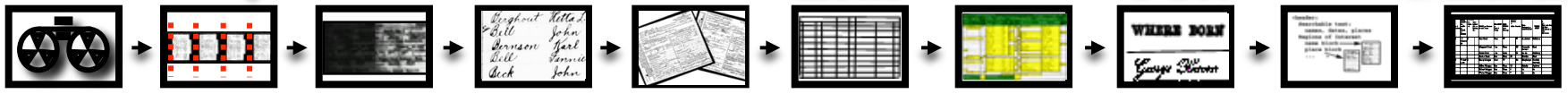
Family History TECHNOLOGY WORKSHOP 2011

Agent "looking over the shoulder" of the indexer:

- Name glyphs + ground truth = Training Set for recognition algorithms
- Agreement/disagreement with other indexer weights prior/conditional probabilities
- "Remembers" image patterns and characters from previously entered names to prompt indexer and refine probabilities and training set
- Exploits contextual constraints from surrounding data

The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Waypointing

Intelligent Indexing

Index/ Search by
- Word
- ROI
- Pattern

Full Text Index/ Search

Intelligent Browsing

Visual Index/Search by Zone, Label, ROI

Intelligent Browsing

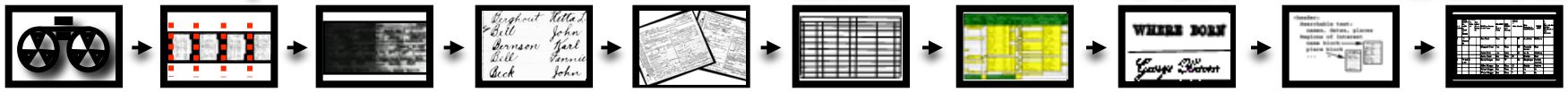
rootstech

Family History TECHNOLOGY WORKSHOP 2011

- Use automated waypointing for light to fine-grained, hierarchical frame to field level snap-click browsing
- Capture and "remember" the frames and fields the user is browsing. Ask: "5 others users have interest in this name – would you like to know who they are?"
- "Amazon Browsing: "Users who looked at this also looked at the following frames/fields/names …"
- Hyperlinking of fields to source data or related information in this or other collections

The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Waypointing

Intelligent Indexing

Index/ Search by
- Word
- ROI
- Pattern

Full Text Index/ Search

Intelligent Browsing

Visual Index/Search by Zone, Label, ROI

Search by Words/Patterns

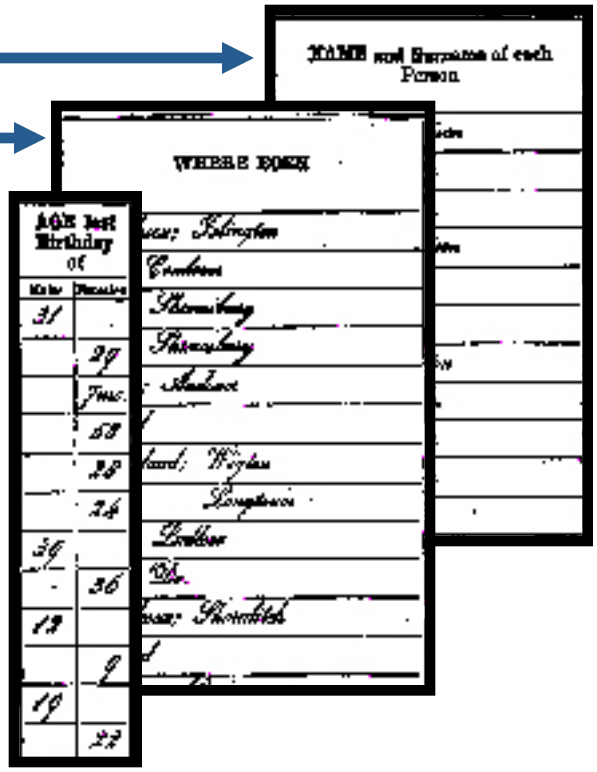
<header:

Name fields

Place fields

Date fields

Regions of Interest:



Partial Search:

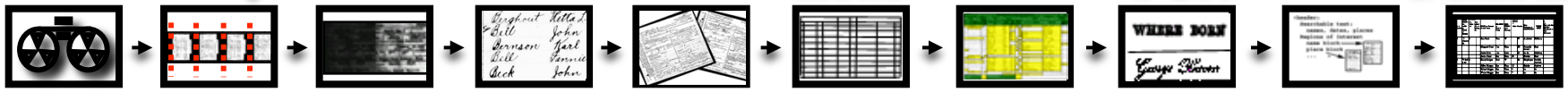
Words, Bitmaps,

Regions of Interests

Logical Groupings

The Digital Microfilm Pipeline

rootstech
Family History TECHNOLOGY WORKSHOP 2011



Waypointing

Intelligent Indexing

Index/ Search by
- Word
- ROI
- Pattern

Full Text Index/ Search

Intelligent Browsing

Visual Index/Search by Zone, Label, ROI

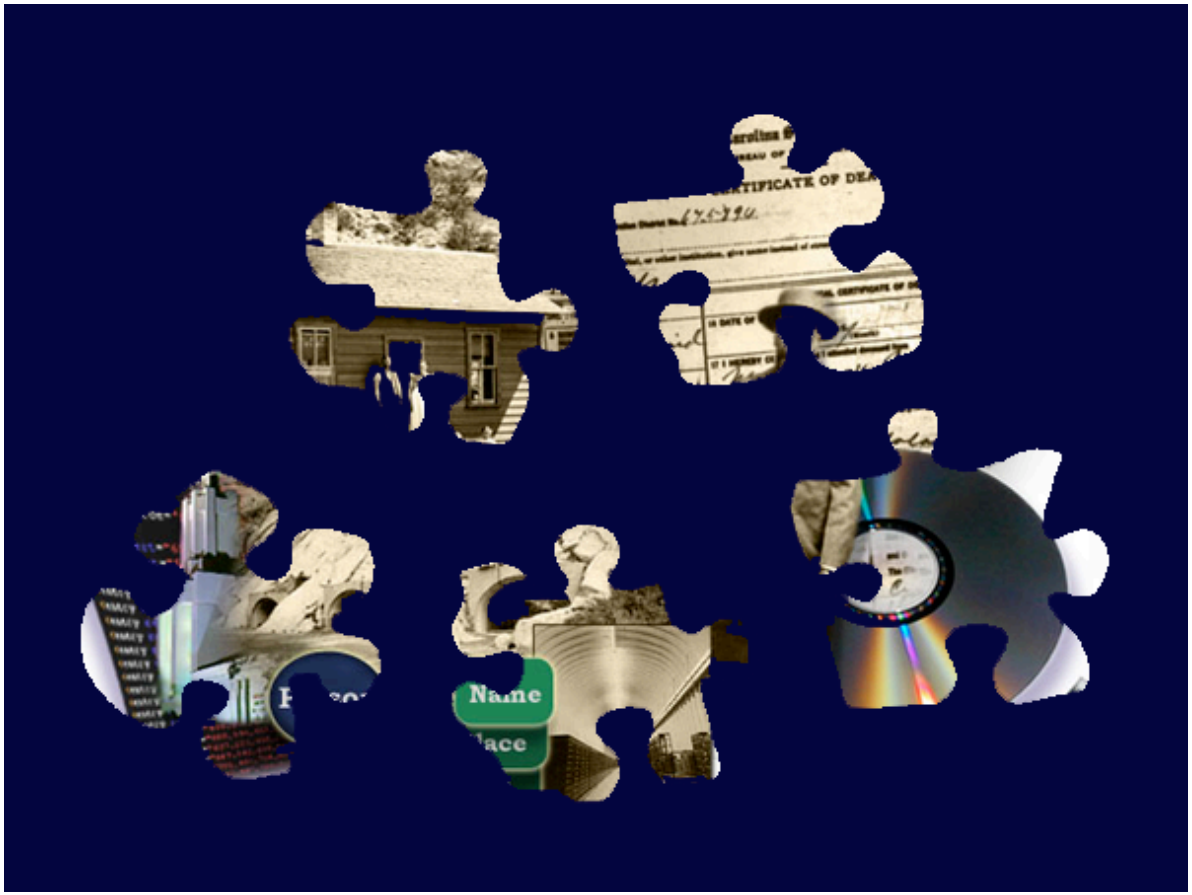
Full Text Database Search

rootstech
Family History TECHNOLOGY WORKSHOP 2011

No. of Schedule	ROAD, STREET, &c., and No. or NAME of HOUSE	HOUSES Inhabited	HOUSES Uninhabited (U.) or Building (B.)	NAME and Surname of each Person	RELATION to Head of Family	CONDITION as to Marriage	AGE last Birthday of Males	AGE last Birthday of Females	Rank, Profession, or OCCUPATION	WHERE BORN	If (1) Deaf-and-Dumb (2) Blind (3) Imbecile or Idiot (4) Lunatic
28	15 Church St.	1		Ann Grant	Head	W		50	Anniutant	Salisbury	
				Margaret Grant	Dau	Unm		22	Domestic Servant	Kent	
				Annie Grant	Dau	Unm		18	do. do.	Salisbury	
				Austin Grant	Son	Unm	20		Hawker	Kent	
29	72 Salt Lane	1		Emily Morgan	Head	W		34	Shopkeeper	Norfolk, Norwich	
				Alfred Morgan	Son	Umn	12		Scholar	Andover	
				Hubert Morgan	Son	Umn	9		do.	do.	
				Walter Morgan	Son	Umn	7		do.	do.	
				Emily Morgan	Dau	Umn		4	do.	do.	

What are the other Pieces of the Puzzle?

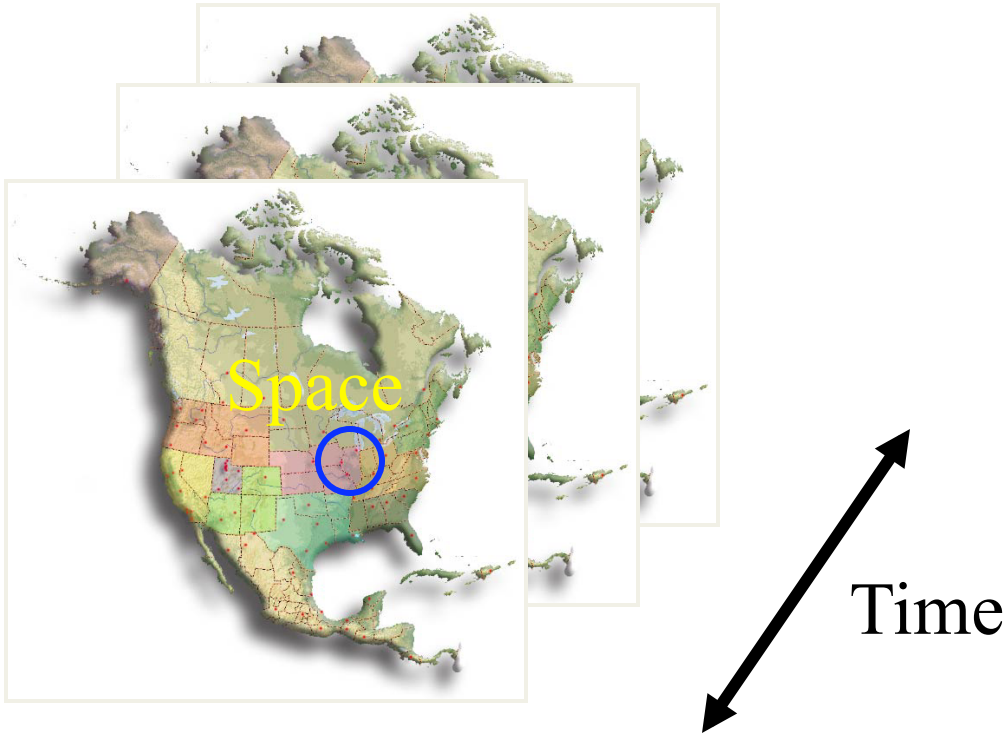
rootstech
Family History TECHNOLOGY WORKSHOP 2011



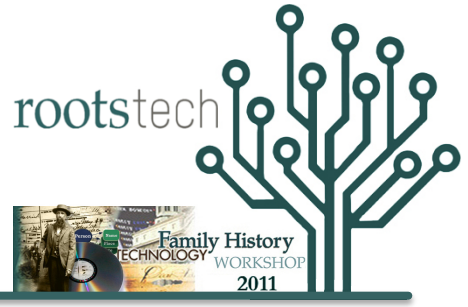
Future Directions

rootstech
Family History TECHNOLOGY WORKSHOP 2011

- Longitudinal Search and Visualization
  <proximity- space - # names - time>



Future Development: What is Needed (Dallan Quass)

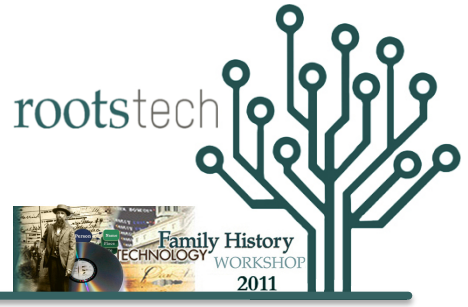


- A comprehensive catalog of genealogical source information
  - when people "get stuck" - consult the catalog to find sources

- Collaboration models that allow people to maintain their private version of the data independently – like the google wave protocol or distributed version control systems

- "Push" systems for doing genealogy: The system prompts me by posting items on my "wall" telling me interesting things I could be doing to learn more about my genealogy

Future Development: What is Needed (Heath Nielson)



- A Common Genealogical Data Model: so that developers, providers and users can structure and link data unambiguously – "stitching" trees together. (Similar problem with Medical Data Records)

- Bi-directional linking of Conclusion Data ⟷ Source Data: Access to, and sharing of conclusional and source data suggests a need to bridge the gap between the two sources of data – invites more careful interpretation Increasing access ***and*** accessibility to source records. Need more compelling and efficient interfaces and algorithms to navigate, search and find data of interest.

- 12-bit (HDR) microfilm scanning to reduce sensitivity of film scanning to changes in film density and to maximize the dynamic range of the (8-bit) output

Future Development: What is Needed (Kirk Duffin)

rootstech

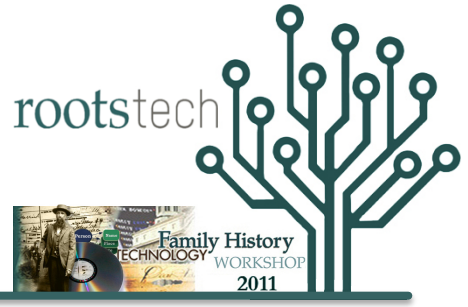
- Closer coupling and improvement in human-assisted computing.
  Why? Because data acquisition will far outstrip available human resources,
  Human resources will outstrip neede expert genealogical knowledge

- Intelligent research assistants (computers)
  - highly specialized to answer specific questions or expedite menial tasks.

- Handwriting recognition system: human transcriber transparently trains computer:
  - makes connections between strokes and output text
  - exploits name authority and locality
  - becomes expert on this recorder's handwriting and the specialized domain
  - eventually system makes suggestions; human becomes arbiter

- Geographic Assistants with specific knowledge about a geographic area over time: Could explain to a relatively new researcher how the place in question no longer exists by that name and is now known by a different name -- and by the way, it also exists in a different county now, because of boundary changes that took place in 18xx.

Future Development: What is Needed (Kirk Duffin)

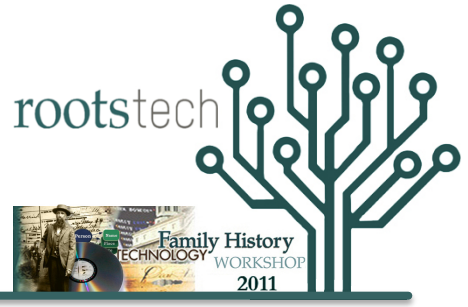


•Geographic Assistants with specific knowledge about a geographic area over time: Could explain to a relatively new researcher how the place in question no longer exists by that name and is now known by a different name -- and by the way, it also exists in a different county now, because of boundary changes that took place in 18xx.

•Defining standards and protocols so that knowledge can be shared between artificial assistants.

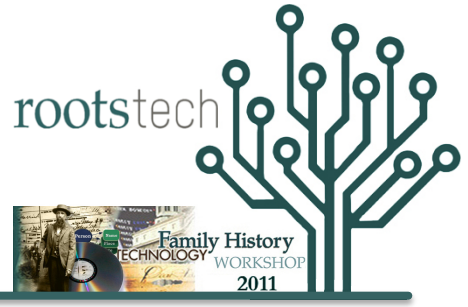
•Improvements multi-lingual knowledge and transliteration so that researchers in one country are not required to become experts in the language and culture of their ancestors before doing effective research

Future Development: What is Needed (Kirk Duffin)



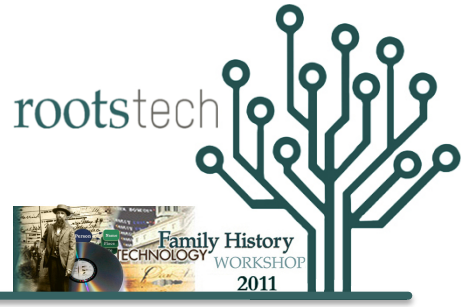
- Dynamic user interfaces and displays for multi-resolution, multi-layered views of a person's digital assets (photos, journals, documents, videos, sound recordings, etc.)
- Expand focus from individuals to families: Family history is typically an individual endeavor. But families are tracked through history, often as a set of individuals.
- Intelligent Data Mining: Analyze entire cross-sectional collections of data and collect them longitudinally. How does the 1830 census of an entire town connect with the 1840 census of the same town? And with subsequent censuses? Can family connections be automatically generated across these data sets? Then start connecting census info with that of neighboring locations. Can we automatically identify who moved? Who got married? Discover where they came from? Can this facilitate automatic creation of verifiable reference info for individuals, families? Can similar information also be gleaned from a comprehensive analysis of parish records?
- This will be of increasing importance for Church family history work because there is going to be a need to identify and do the work for those who have no descendants. These can only be found by expanding the research work from individuals to groups.

Future Development: What is Needed



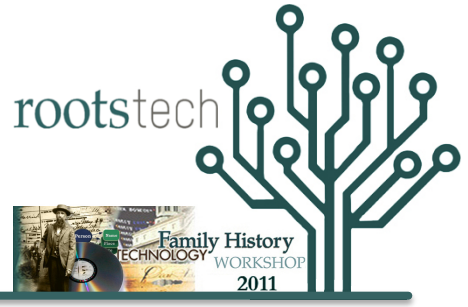
- Interchange format standards – a worthy GEDCOM successor
- Cloud/synch – same content available everywhere from my PC, to TV, tablet, smart phone, game console, car, etc…
- Better entity extraction – to create relationships between data mined from records
- Better search – interactive "hinting" search (i.e., Google instant), name-place-date-event entity extraction, etc.
- Links\connections\hints from tree framework (vitals, census, etc.) to rich media (stories, newspapers, yearbooks, photos, movies, videos, local histories, etc. adding meaning and detail beyond dates and names)

Future Development: What is Needed (Alan Eaton)



- Synchronization of the cloud(s)
- Sourcing
- Ease of consumption of genealogy services and data
- Everything fully indexed and/or structured

Future Development: What is Needed (Bill Harten)



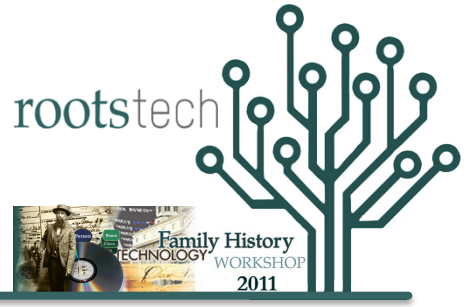
- **Challenge: Evidence-based Genealogical Computing**.

- **Evidence-centric data model and computerized evidence-linking**: Need to be able to verify, change, extend conclusions. Cannot build on other's conclusions alone – must follow source references. Need to see an image of the evidence.

- GEDCOM's syntax and extensibility was conceived for this evidence-oriented purpose, but its first and only application has been to represent only our linked conclusions, because the community demanded computerized pedigrees first. Work on linked conclusions without linked evidence will need to be done over.

- This has been understood by many since the early 90s, yet nothing substantive has happened to address it.

Back to the *FUTURE*



If hindsight is 20-20, what about foresight?

Back to the *FUTURE*

rootstech

Family History TECHNOLOGY WORKSHOP 2011

If hindsight is 20-20, what about foresight?

Back to the *FUTURE*

rootstech Family History TECHNOLOGY WORKSHOP 2011

If hindsight is 20-20, what about foresight?

What kind of Family History Technologies might be with us in the year **2020**?

Back to the *FUTURE*

rootstech

Family History TECHNOLOGY WORKSHOP 2011

Smartphones/Cameras/Recorders with

- "flex-cell" HHD display, contact scan, OCR
- HHD photo, full motion video, sound
- *IntelliSpeech* indexing, search and retrieval
- Virtual presence with multiple callers, collaborators
- Virtual networking with devices and friends
- Smart authentication and dynamic portal connect
- 1-10 terabytes of memory

Back to the *FUTURE*

rootstech Family History TECHNOLOGY WORKSHOP 2011

Personal "Wall" for digital assets

explore, view, query, organize photos, movies, music, recordings, digital documents, files, etc.

- Speech and gesture interaction and control
- Shared views, virtual presence with social network
- Virtual travel and presence to any destination
- Smart authentication, device and network connect
- 1-2 petabytes of memory

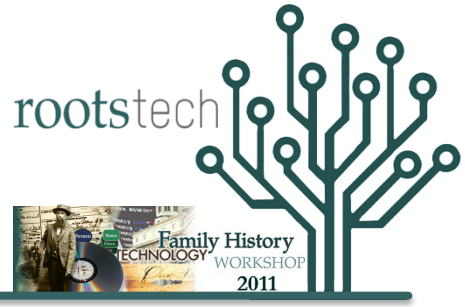
Back to the *FUTURE*

rootstech
Family History TECHNOLOGY WORKSHOP 2011

Intelligent Genealogical Research Assistant

- Organizes and keeps track of photos, movies, digital documents, files, sources, what you were working on, names, dates, etc.
- Fully speech aware and interactive
- Domain, geographical, linguistic expertise
- Suggests what to work on next, asks good questions
- Accepts assignments, searches and mines data while you sleep; presents findings in the morning

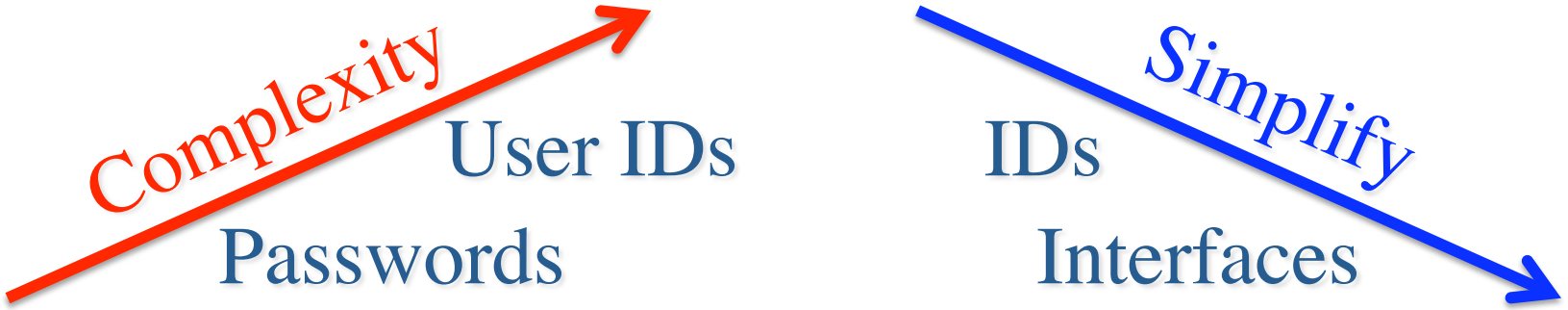
How to move *FORWARD?*



How to move *FORWARD?*

Competition!

The Need to Simplify

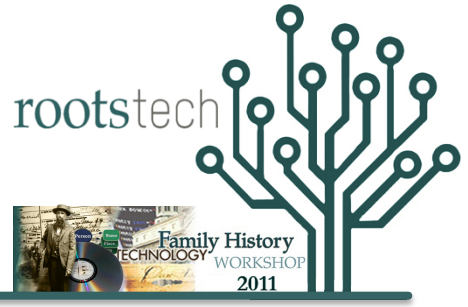


User IDs
Passwords
Security, Software,
Data Complexity

IDs
Interfaces
Data Models, Sharing,
Collaboration

There is a great need to simplify and unify technologies, algorithms and data

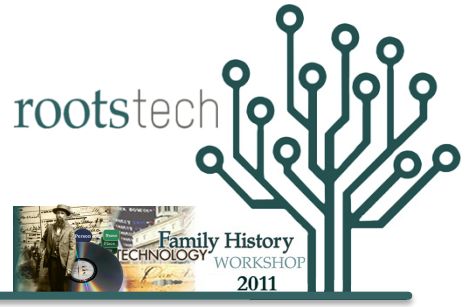
Fun + Experience + Reward



- The busy mother of 3
- A turning-the-heart experience
- Rewarding: don't go away empty-handed
- Pac-Man genealogist
- The "20-minute genealogist"

Curt Witcher genealogy – get a life

The Need to Unify



- Common Data Model
- Record Linkage and Merging
- Standardized collaboration model
- Closer Human + Computer coupling
- Genealogy and Source Data

Of all of these, perhaps the most unifying concept of all is that this is the Lord's work to link and unite families in His temples.