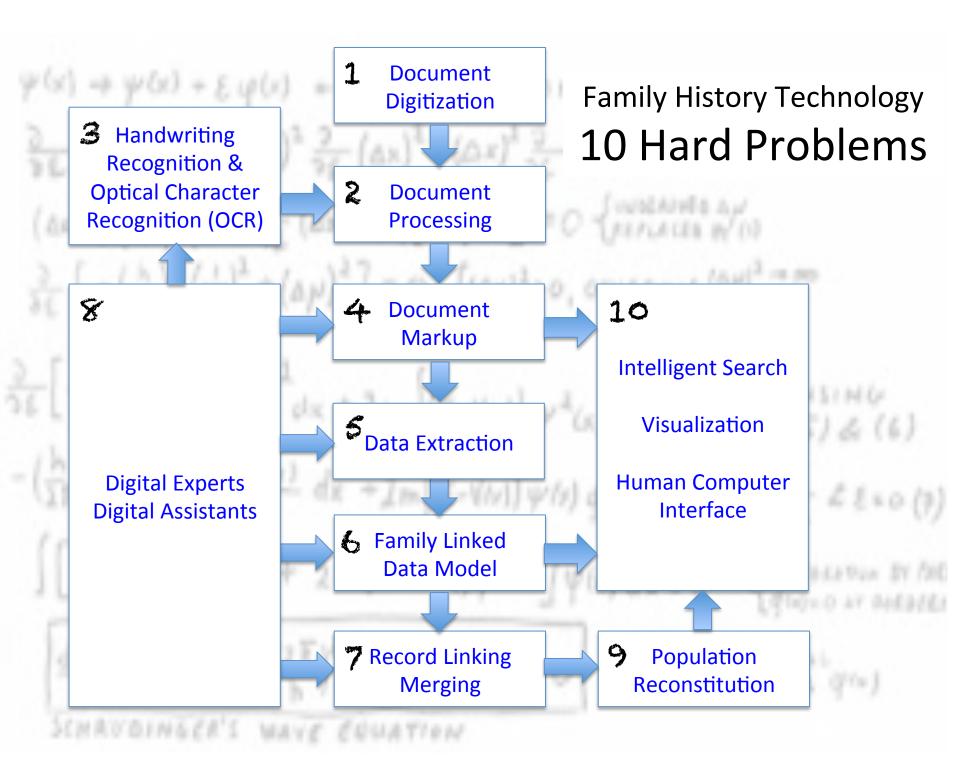
Family History Technology: A Survey of 10 Hard Problems

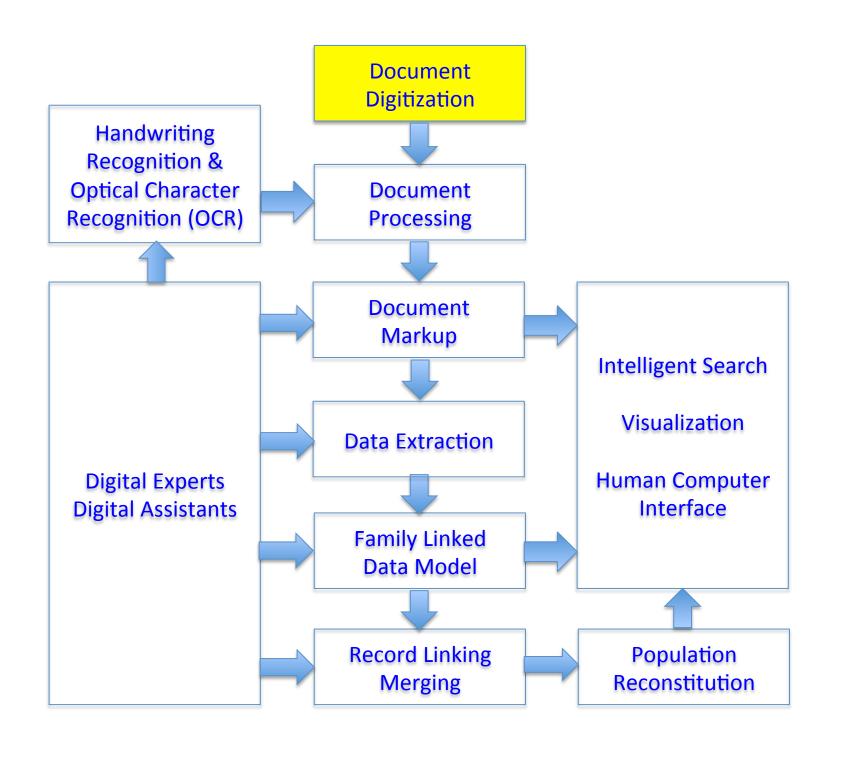
Dr. Doran Wilde
Brigham Young University
Dept. Electrical and Computer Engineering

Family History Technology Workshop 2014

Decomposing a Really Really Hard Problem

- The goal is to build a common lineage-linked family tree for the whole human race based upon available records.
- Decompose this seemingly impossible challenge into 10 hard problems.
- Allows researchers to focus on a part of the overall problem.
- Illustrated by a running example.





Document Digitization and Preservation



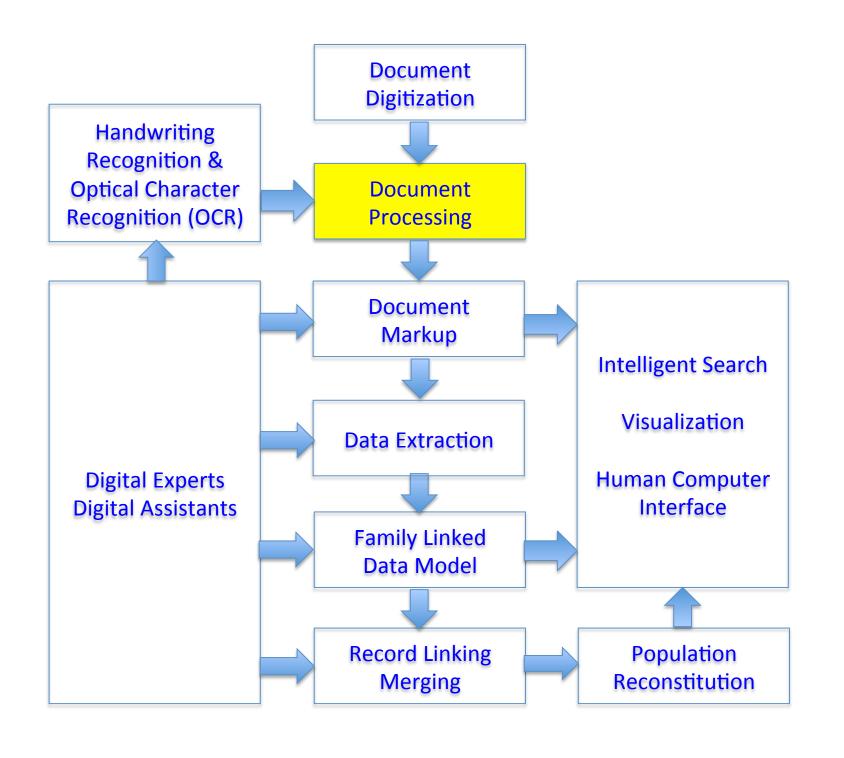
Preservation, Accessibility, and Enhancement of Records

Juni sta formibble Side Sorens Sahari Kashja Anne Dortha. Sot In The Martie Bland Sold Che Hans Sahar sugger Son Lars Han Jen Justan Mans hlansen , Tens Nielson , Christopher Lasson , Marie Per Hrandure Frogenis fight. Jane 12 Diguit blog Jaguen Gades Von Niels Torgensfort den 212thay sia; Janes Jamestrad i Rickan balsaftah. Saldan Midame Mathiefen og fomfo. Mathiefen fon thyslings, nogla Guardud og Round. tie Burantain Hang istagh Jens Hielsen fra Tions Jan 13th Martin Sprindoby Sich Strans i Nastine Palar Anna Sothe.

and har frequency of this Decorn is Later Nelson The Sor Oliver.

+ 1890. Maren Hand Persons. 1/2 prinit blastoth lens torgenens bake i haven dallas proprie Softman, Land Christensen, Peter Olsen, Kenrich Torgensen og Lats Andersens fall Side . -28m Soutago Brougas Thinge Ledersens Oou Soft Anders Saldon Johan Kroger of Holbech. Hanning Pederson, Toren Michelson, Hani 26 may Hans Sorewers Jane folk haven Jaron to Juplace Suffer Andersen, Else Ole Rasmufens fifting fra And Sind no honge Mans Sore noon Jakon Sol Ingelowalling Nich Whe The smoven, him Jenson, him hopenson hotes house Rasmusen, hand Sonson, Hans histolien, Magrethe Oie Rasmusen, Marie Per Olso Br. frogland graphic. of Luxue Sol Stans Jenons Jaha Anne . frafrance To Contract +1888 milethen, Molladow, Ole Poulson Miroden Nels Naisen 6th marks flow Bowneys & Hans Christiansens Bow fine who for a lation 10 July Gran Sold Peder Victoeno Von Ole Sta Nasbije in Millen But a drefa Niels Phristian Salvara: foothang Rosen, Niels helven foren pap. Trin: sall populations, moleran, Soren hichelan Mole Hole 28 2 Jobb Jacob Vejchels Jakan Anne Sophie Magdalene Michelsen Hans Hendrichsen Else Christens 28 May Janua Och Peder Chans Jakes Stone i Gamble Sold Son Stadame Topic family the Berrown haven has mind I 23 as other Jons Nielson, Ola Sabarren. Soulinto Trafant finder Niels Nielson, Ham hichelson e dem Taro Christmans Bon tolk Hans. Taron Medame Show tolk to the Hans. Taron Medame Show tolk to the Hans. The Staron Show the Show the Show the Short Staron Son Jakes Both to the Short Man tolk to the Short makes the Short Son Man Joseph Tongen Honor Joseph Medam tongen the Man Joseph Man tongen to the Medament Joseph Taron Son Joseph Mayreth Manus. molentan, Anne Sorenos. Edem jausen Jons Hansens John toth Else Sallx sta: Mudana Hald, Oth Hansen, tens Lersen, Ander Jonsen Theffe Nov: lod quaring Christen Nelven Both in Dator Else. Jatiethe Melle Nov. Logged Dorsen, Ole Anne Jons Olsens . -218 August Elan Grundolf Ingune Godes Darnit Ellon Topsene 180 part of the Some John Stander Hand of the Color of the Stander Rasmison, Ausers Herborichsen . 22 p Trivit: Let Soron Nichon Degn Sohn nu Tou Sens. for Son na Joseft: School Borer Lancen Landfunderson, Peder Olson, Anne Hendricks - 1885 Anne Hendricks - 1885 1712 p. Smith: Jangon Hans Sensen i No stoje nu Julius Johanne til Ausbin. Judon & softent fift: Hans Michelsen Hans Inderson, Nicht Seansen, magrathe Hans. Horas Tripik: John Arve Jouvens John & Karen Jullana Inform Sugaria, Niels Nielsen, Oke Hansen, Peter Sorensen, mag 872 po Trinib. gold Gunglundi Bijeljerge Eder Nichons Jakon, Johans Jason Garagna Superin Joseph Mincloon, Ola Te Semen Mana Sorens - 4875.

42 Clov: View Mayer I Mana Son, Hand, Januaroth - dash the Selout a program Julian Hans messelver, Molenmen, Thomas Habreng Jen, Kirsten Sovens, Karen Knudsts

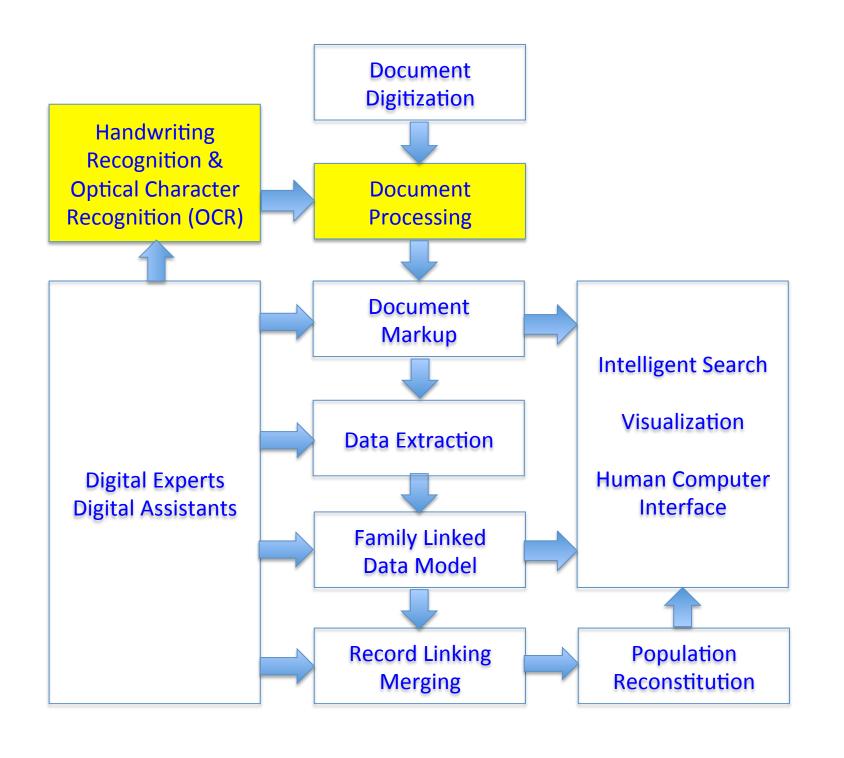


Document Processing

Spatially parse the page into blocks to give context and meaning to what is written.

Zoning.

Tunii sta Gamadobh Miels Screw Parbari Nasbje Anne Dorthe. Ist Daw To Martie blan Sold Che Mans Sakno & washe Sou Lars Han Jen Jastan Mans Mansen, Jens Sichen, Christopher Farsan, Marie Per Al radhors Fraffaus Gift. Ince Po Diguil blog Jaguer Gades Voic Niels Torgen of How It thory in gover framedonat a Richard balosoftat. Saldan Madame Sathise for on fourte til Bowanfain Henn ürlagh Tem Nielsenska Typinge Thathirten for Hyllinga, nogh granded og Round. -Jan 13th Martin Symundo by Sich Woons i Nastine Jahar Anna Sothe. . and box free from gaper ships Descrip in Solar Nelson Ander Oliver. , + 1890. Maren Hand Persons . 172 p Trinit blastoth tens torgenam lake i haven dather Prophy Softwin, Lars Christensen, Peder Cloen, Henrich Jorgensen og Lars andersens fel file . -Inu 17th April gineen och theke Marie Sorond auch Down Johan hirizer & Ristante pe Budaflas fragen. Apopula Roma, for at lastre, ficher Office, Minde Shiperfue, the This found the Base Share for the Maje Hay high Prosinages Ingul Johan Robiger of Moldeth. 1810-Thige Sedervens Con Soll Anders Saldon 20m Sveetag po. 3 Rougas. planning Perowen, Loven Michelsen, Hani 26 th May Hans Sorensens and folk haven Surson to Tropens Suffer Anderson, Else Ole Rasmufens fufri, fra Whe Trasmoven, firm Jenson, Siais Rosmont hotal Rousic And Suplivar far Int. -4 Down p. 3 Rouge & Mans Sore noew father Soll Ingelow athen Nich Rasmusen, hair Jonson, Hans thickelien, Magrethe of Jusic Soll Stand Jensons gather Anne . profrange to Carrie Die Rasmusen, Marie Per Hotolor frofond guffsrin +1888 pri Entfra, Molinton, Ble Poulson Miroden Niels Niels Niels Niels bling Bounger Ham Christiansen Bin gjan while glack Niels Christian Sallar : Joshan Roma Niels Nielsch Joben 1m July Spring Jobh Peder Victoeno Von Ole Sya Nasbije in higher pap. Trin sall frepuntion, moleran, Soren Michelson Mila Stola 28 2 Jobs Jacob Vejchels Jakan Anne Laphie Magdalone. michelsen, Hans Hendrichsen, Else Christens 28- May Janua Soft Paro Closes Jaket Anne a Gamble . Lood in Madame This Sumper Birrown haven has mad Paro Other Sons Nichon , Ole La barren . Joelinto Profund Saide Niels Nielson, Ham hilhelsen Cosen fars Christeniens Son Hold Mans, Lasonen Medame Shoel Soft 1891 Rechot Kendrich trocensen, Christon Viction, Ob Learnen School Son Cheggift 54 Million Vill Millions Don Mistr West brught chargement of white on Section, the Bright State on thems Toronson Jorgen Hender Johnson, guinable that Jude new Juffer Mayrethe Mann. molentan, Anne Sorenot. Codem jain and Sons Hansons John & Joth Else Sould An: Mudame Heals, Othe Hansen, Jens Person, Anders Jensen polyte for lod grande Christen Victor Wohne Jutor Cle. Jahren Welver John Jufg I Dorsen, Ol Anne Icns Olsens. 21 Thingieth Dan figure dobt Ingune Gales Jarry Ellon Jensene Basmison, Anders Herrichson. 19th ps. Trinis: Footh Howagen Hans Olive tiansons Jakas Billa Marie Sadarie France Support Sothe Jened Niels Niels Marie Son, Anders Honsich sen. +1882 220 p Trivit les Soron Viction Dogn Sohn per Con Sons. for Sonn Joseph: Athoric Sever Larson, Terreffens lons on, Jades Olice, Anne Hendrights 1885 17th p Simil Jangen Hans Jousen i No styr ne Julia Tohanne fil Burben. 21 rgs Tornite Joth Arve Jonvens John & Karen Jallana Inform La Donn Frofent Sift: Acans Micholsen Stans Inderson, Saythun, Niels Nielsen, Oke Hansen, Perco forensen, man Wiels Jeanson, magrothe Hand. Res po Trinih: gold Gunstmant i Bijelje rav Eder Nichons Jakon, Johann Jasson: I hafnut highin, Josep Mincloon, Ob Jedemen, Mars 42 Clar: View Nature i Morres Son, Hand Grand James Solands 42 Jour p Criptan: Song fin and Niels Sovensons Son 29 Then v. finan 451882 Libbs Sou Jens, Graghi Ristan Sallar Joren Sickon lan Sorens on Mans Henrichsen, Anne Mans Jorens & population States of Barrens Son Zans with some his the South of th Below in program Suprim Hons merolica, Molenton, Thomas Habren, Magate bird. Levenin Justine hour, Der Bleen, Law Christensen Hanston Jen, Kirsten Sovens, Karen Knudsts



Document Processing and Handwriting Recognition

Lines
passed from
image to
handwriting
recognizer,
and collates
the results
back into an
annotated
transcript of
the record.

Juni sta Gamabil, Siels Sorens Rubari Nashje Anne Dorthe. I'm Jan 78 Marti Alan Sobt Clar Mans Sakar Sugar Son Ears Ran Jen. Jassan Mans Hansen, Jens Sielson, Phristopher Bow 10 Diguil blog daguen Gades Vow Niels Torgen for Son 212 their sia. Larson, Marie Per Peratinto Trafant Gula. gour grandward a Rickon baloughan Salara Madame Sather for og foungs tie Borranfain Henn istagh tens Nielsendra Lyonge Mathicken for Hydlinger, nogla gaarbad og Round. -Jan 13th Martin Symmatoby Sich Solomon Nastine Fallar Anna Sotthe. and box for offere gaper ships Descript Sect Nelson The Sorr Olden, + 1890. Maren Hand Persons. 179 - prinit blassoth Sens Jorgenans Pater & haven dallown Proper Softman, Land Christenson, Peder Cloen, Henrich Jorgenson og Jun 19th April gineendobs Stocke Marie Sorand work Som Johan Project of Richard po Boulafor fragen Apopula hour faller Sarpe John Son John Son Son Project of Richard Stranger Star Star Sarpe Mary Will David Consular Strand Lats Andersens fall Silfe . . 1810-20m Vondag po 3 Rougas Thinge Sedervens Cou Solh Anders. Laktorn Johan Kroger of Holbech. Flanning Ledowen , Loven Michelson , Hans 26 thay Hans Sorewoons parter folh haven das on the tropper fulls And ersen, Else Ole Rasmifens fin pring for Ble Trasmusen, Prois Jenson, Siais Rosmon Motos Rouse 4 Dour p. 3 Kongan Hland Jote noew father roch Ingent Rasmasen, hand Sonson, Hans thickelven, magrethe of Jusic Sol Stand Sensons father Anne . propriety to Carrie Die Rasmusen, Marie Par Ofrailer frofont juftrick +1888 Miletfon, Molenton, Ole Poulson Mission Niels Niels Niels Niels Niels Niels fly Blowage & Hans Christiansens Son Grandshog leron 10 Julis Sprin Solt Peder Vielsens Son Ole Sya Nas bije in Mither New Christian Jadhan Foothung Rosen Mile Michen Joren 12/20 Trin: Jall fighaut from Mole xxu, Joren Michelsen Mile Stole 28th Said Vajenels Sahar Anne Sophie Magdalene michelien Hans Hendrichsen Else Christens 282 May Journal Poor Clans Jahn Stone a Gambly Substant Hadane Birosen haven Rasmain. Peder Other, Jens Nielson, Ob To Derson. Jadbrow Krahnut Kant. Niels Nielson, Ham hithelsen thro Christmans Bon Hold Hans, Asimon Madama Show, Sidje +1891 Reason to the Aright of mounts, Christma Millon, Ba Barrers, Sidje with the Madama San Hope letter the the things of the state of the Madama San Hope letter to the Madama San Hope letter to the San San San San Madama Madama Honor. Todorsen, which the the Table and Sight Mayorthe Manus. molentan, Anne Sovenst. Edem jausud Jons Hansens John & Joth Else Sallinn: Mudame Head, Oth Hansen, Jens Person, Ander Jensen Takefte Nov tod granden Christen Nelsendebann Jakos Else. Ladius. Anne Jens Olsens. 18 High Find four four North In gane gades fair of Ellon Jensene Wasana Protte for Angebor Molinto, Thing I Joseph, Ol Prasmisen, Auser Herrichsen 199 ps. Trinik. Tobb Howagen Hand Olivistiansons Jaken Bille Marie 199 ps. Trinik. Tobb Howagen Hand Olivistiansons Jaken Bille Marie Sallani Haphan Juphan Dothe Jend. Niels Nielson, Hand Hall Jen, Huder Honrich son. +1882 272 p Trist for Jorga Viction Dogn John nu Don John for John Joseph: Schon , Peder Lansen, Tensfinstonsen, Jades Olies. Anne Meneriche 11855 (1 Julian Le June) 172 p. Smith ganger Heans Fouseni No styre ne July & Johanne til Justin. Jagan Grofint Sift: heard Michelsen Hears Inderson, Worge Trigit: John Arve Jonsens John & haven Julan Frage Sugaria, Niels Nielsen, Oke Hansen, Perce forensen, man Wiels heansen, magrathe Hand. 200 po Trinik John Gunstmanti Bichjerge Eder Nielsons Jakot, Johans dalinan Jonghan Suprim, foren hindson, Ole Je demen Main 13 Abril go Epiphan: Stong fine and Niels Sovensons In 29 Jan v. Spine Topsen Anne Hans Sorens - +1875 441882 John Bou Jons, braghi Rinkow Jadhar Joren Helson van Joren Stelson van Jorenson, Mans Mennithen, Amelhan Joren, Jaghand Jorens 42 ald : Vais Sulsen i Morens Son, Hand, Spring och - dallaga. Oded Gasefina Sin undel Ole Berens Son Lars. Brook a like the Dokorful toit a Googland suprin Hans merolen, Molenne, Thomas Nothern a Spoke Voice. Leven profines thoun, Per Bloom, Laro Christonson Hand to Jahr 1872 Pathorne Ole Persons -Jen, Kiroten Sorons, Raran Knudsis

Document Processing and Handwriting Recognition

28m Nordagp. 3 Rougas. Thige Sedervens Over toth Hoders. darburn planning Sederson, Toren Michelson, Hanis Anderson, Else Ole Rasmufens suffra find find from far Suff. [Annotations are add

[Annotated transcript]

[Christening Record]

[Source: Church records, Oro, Frederiksborg, Denmark]

[Image 112]

[Heading] 1810

[Margin Note] 2^{den} Sondag p. 3 Konger.

[Body]

[Line 1] Thyge Pedersens Son dobt Anders. Faddern

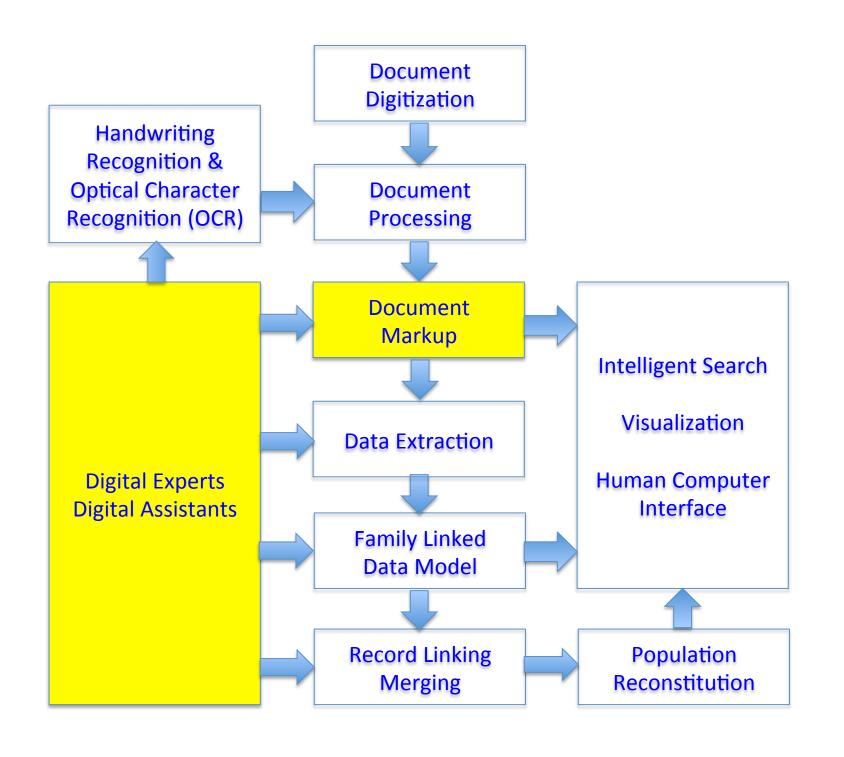
[Line 2] Henning Pedersen, Soren Michelsen, Hans

[Line 3] Andersen, Else Ole Rasmusens hustru, Præ

[Line 4] stens hustrue bar det.—

[Annotations are added or inferred by the document processor with human input.]

A transcript is a faithful copy that preserves the characteristics of the original as much as possible, making a useful copy for further analysis and research.

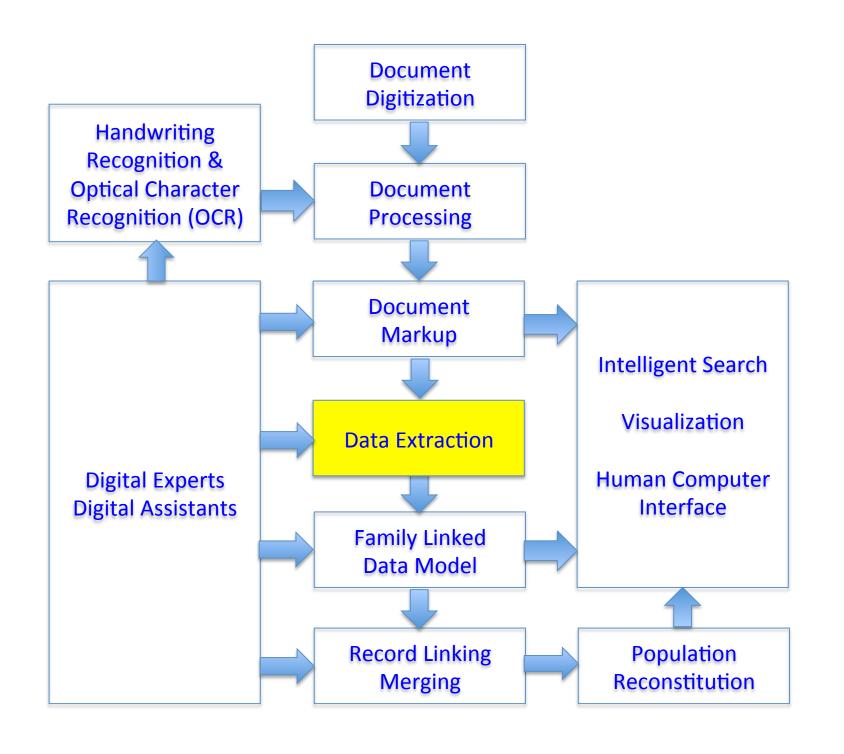


Document Markup

[Source: Church records, Oro, Frederiksborg, Denmark] (Danish) [Image 112] Categorize text into semantic [Heading] **1810** elements: names, dates, places, etc. [Margin Note] 2^{den} Sondag p. 3 Konger. Adds markup to the document to [Body] convey this information. [Line 1] Thyge Pedersens Son dobt Anders. Faddern Uses Date Authority [Line 2] Henning Pedersen, Soren Michelsen, Hans Uses Name Authority [Line 3] Andersen, Else Ole Rasmusens hustru, Præ **Uses Place Authority** [Line 4] stens hustrue bar det.— [Source: Church records, Oro, Frederiksborg, Denmark] (English [Image 112] Translation) [Heading] **1810** [Margin Note] 2nd Sunday past 3 Kings. [Body] [Line 1] Thyges Pedersen's Son christened Anders. Witnesses [Line 2] Henning Pedersen, Soren Michelsen, Hans

[Line 3] Andersen, Else Ole Rasmusen's wife, the Priest's

[Line 4] wife carried the child. —



Facts

- Also called assertions (facts with associated confidence ratings).
- Answer the questions:
 - what (fact type)
 - who/role (might be multiple persons)
 - when (date)
 - where (place)

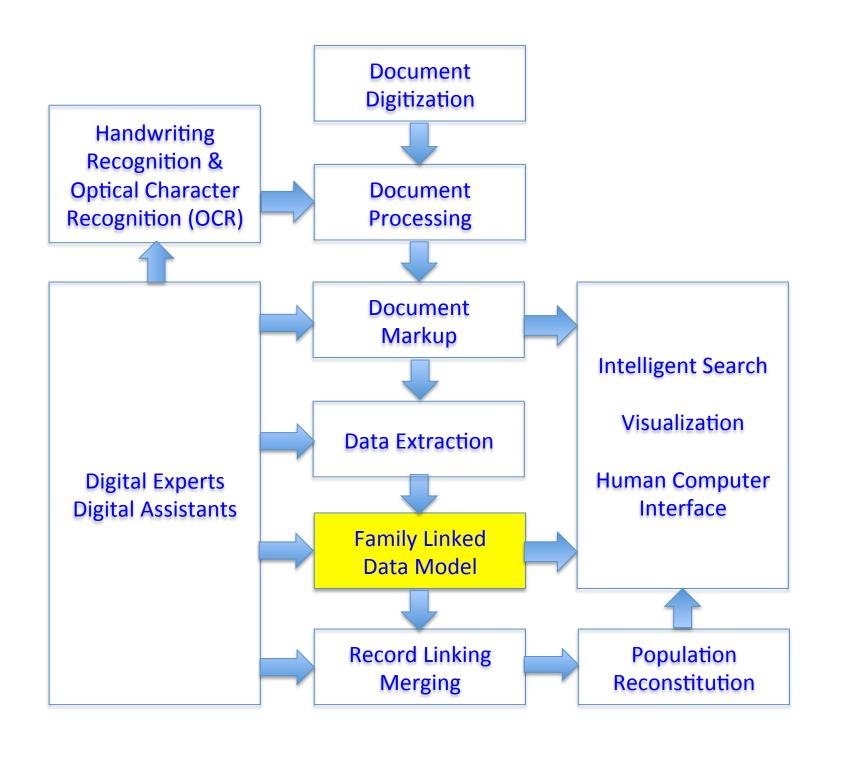


Data Extraction

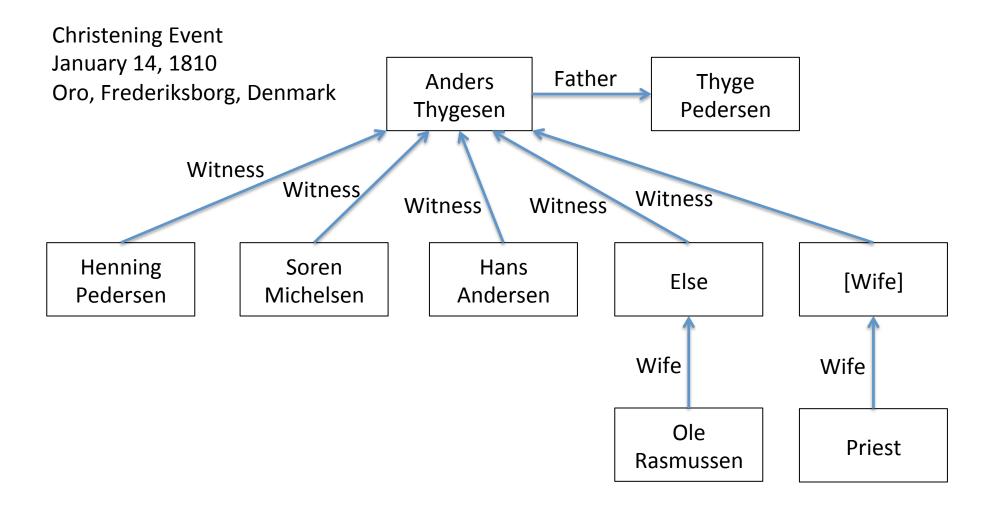
Inferred Facts:

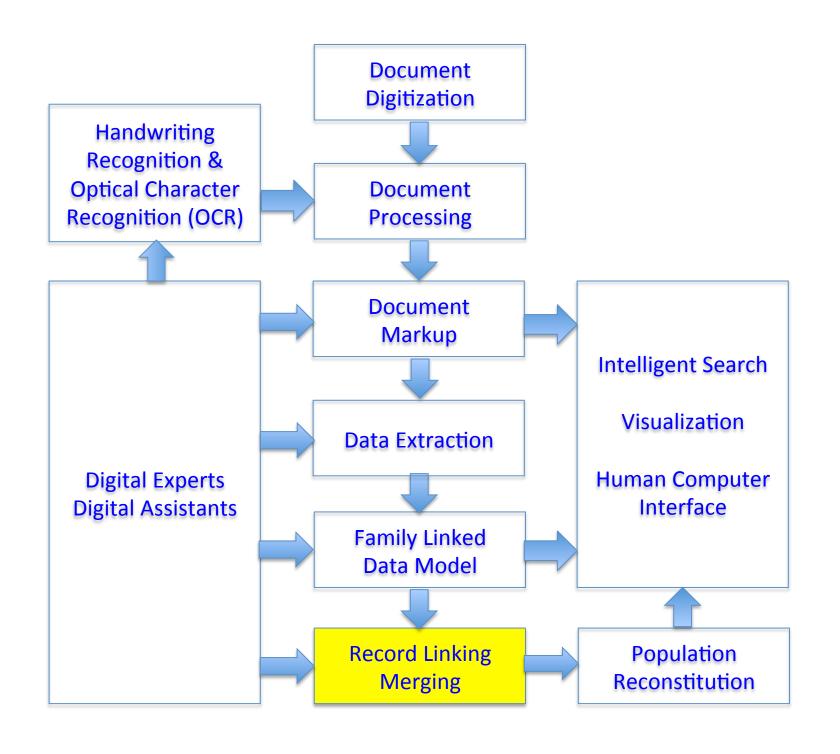
- Source Citation: Oro, Fredericksborg, Denmark Church Records, Christenings, Image 112 [Document Processor]
- Location is Oro, Fredericksborg, Denmark [Place Authority]
- 3 Kings Day is January 6. [Date Authority]
- 2nd Sunday past 3 Kings Day in 1810 is <u>January 14, 1810</u>. [Date Authority]
- Anders Thygesen was christened
- <u>Thyges Pedersen</u> was the father of <u>Anders Thygesen</u>
- Henning Pedersen was a witness to the christening of <u>Anders Thygesen</u>
- Soren Michelsen was a witness to the christening of <u>Anders Thygesen</u>
- Hans Andersen was a witness to the christening of <u>Anders Thygesen</u>
- Else was a witness to the christening of <u>Anders Thygesen</u>
- <u>Else</u> was the wife of <u>Ole Rasmusen</u>
- The Priest's wife was a witness to the christening of Anders Thygesen

(12 Facts from this simple Christening record.)

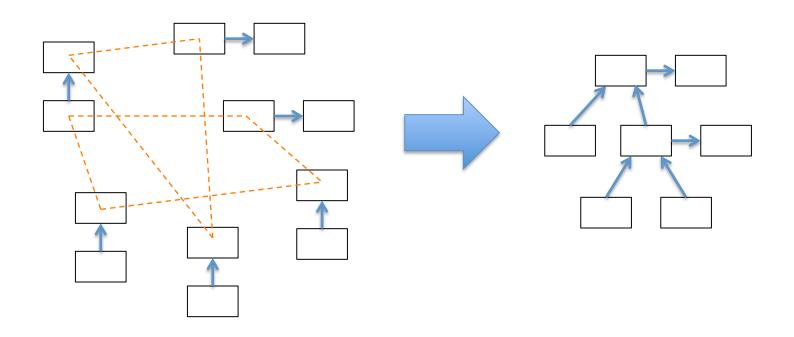


Family Linked Data Model



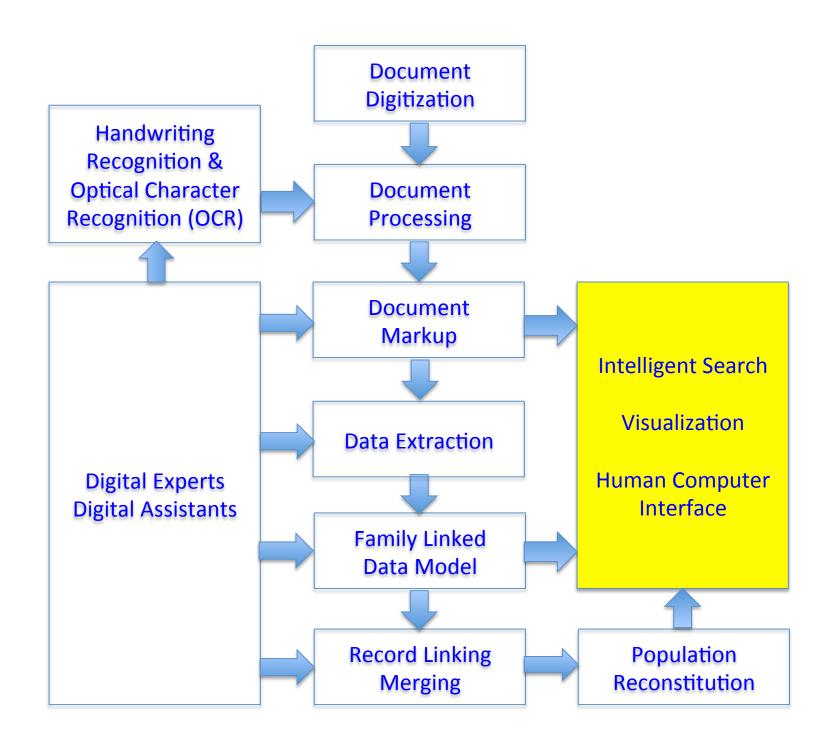


Record Linkage/Merging

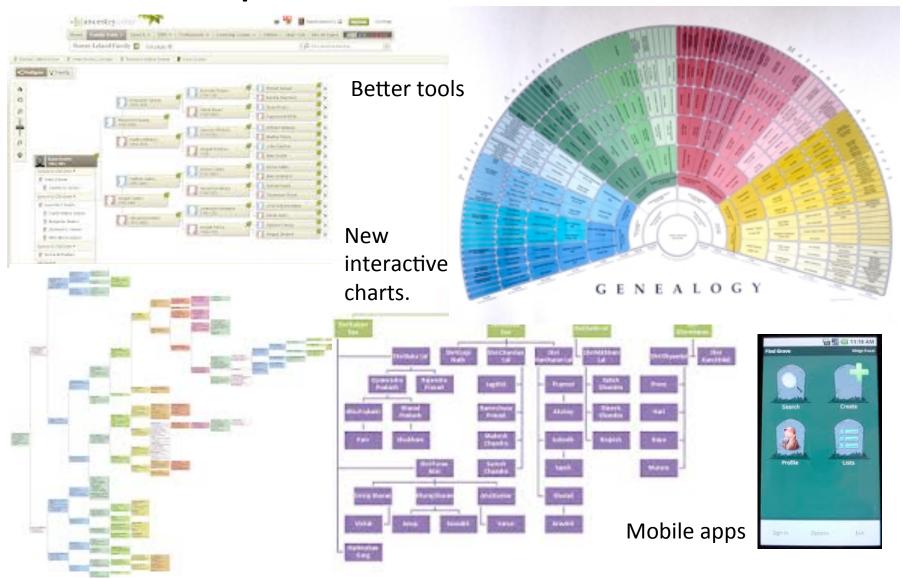


Record linking determines when two or more different records refer to the same individual.

Duplicates are merged together to build lineage-linked families.



Search, Visualization & Computer-Human Interface



Summary

- I have surveyed 10 problems that are on the critical path to a revolutionary break-through in the way we do genealogical research.
- There are undoubtedly other important problems that impede family history that haven't been mentioned or haven't yet been identified.
- The Family History Technology Workshop is held to discuss these kinds of problems.
- The hope is that technology can put the ultimate goal of assembling a family tree for all people within reach.