

# Family History Technology: A Survey of 10 Hard Problems

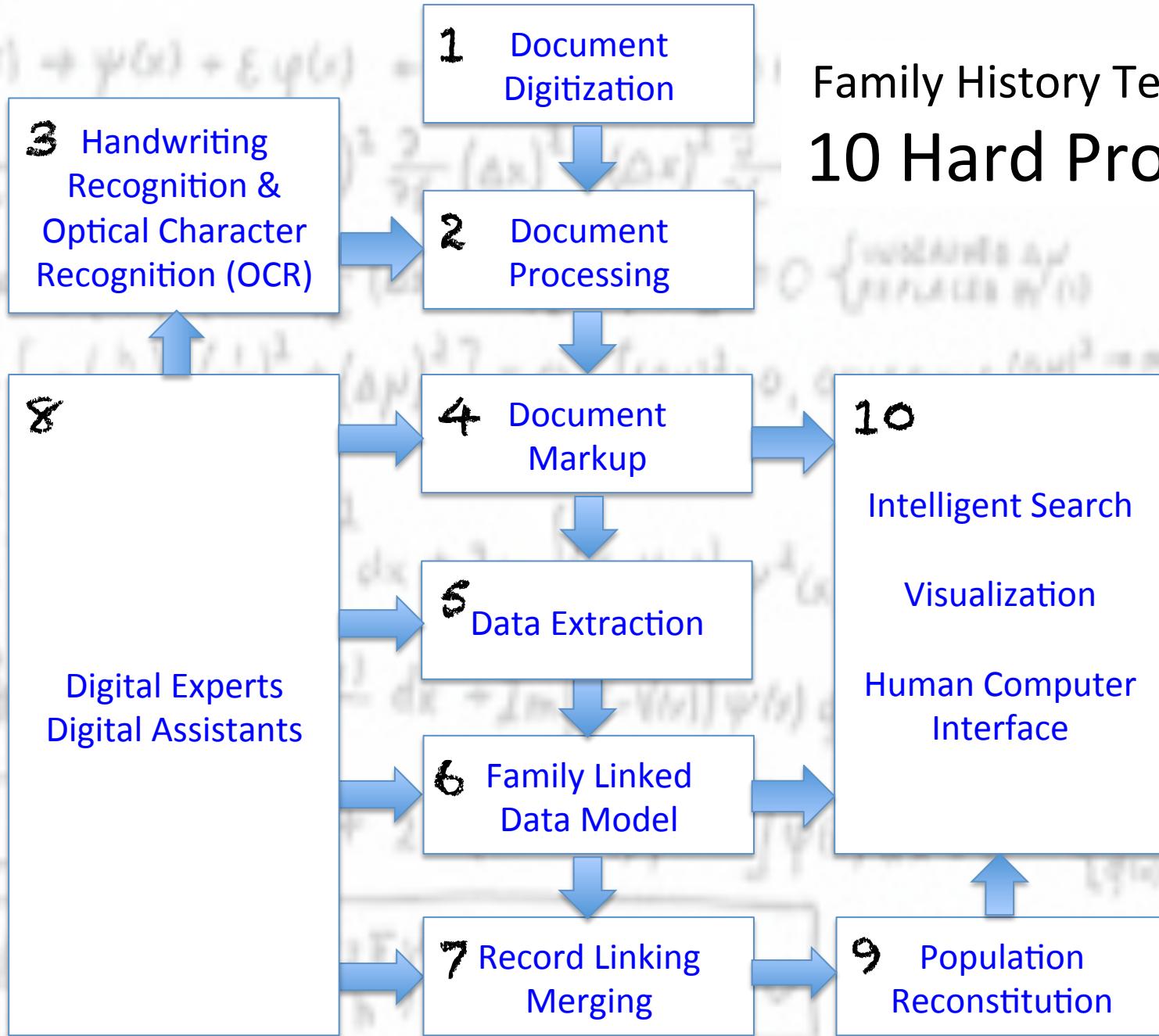
Dr. Doran Wilde  
Brigham Young University  
Dept. Electrical and Computer Engineering

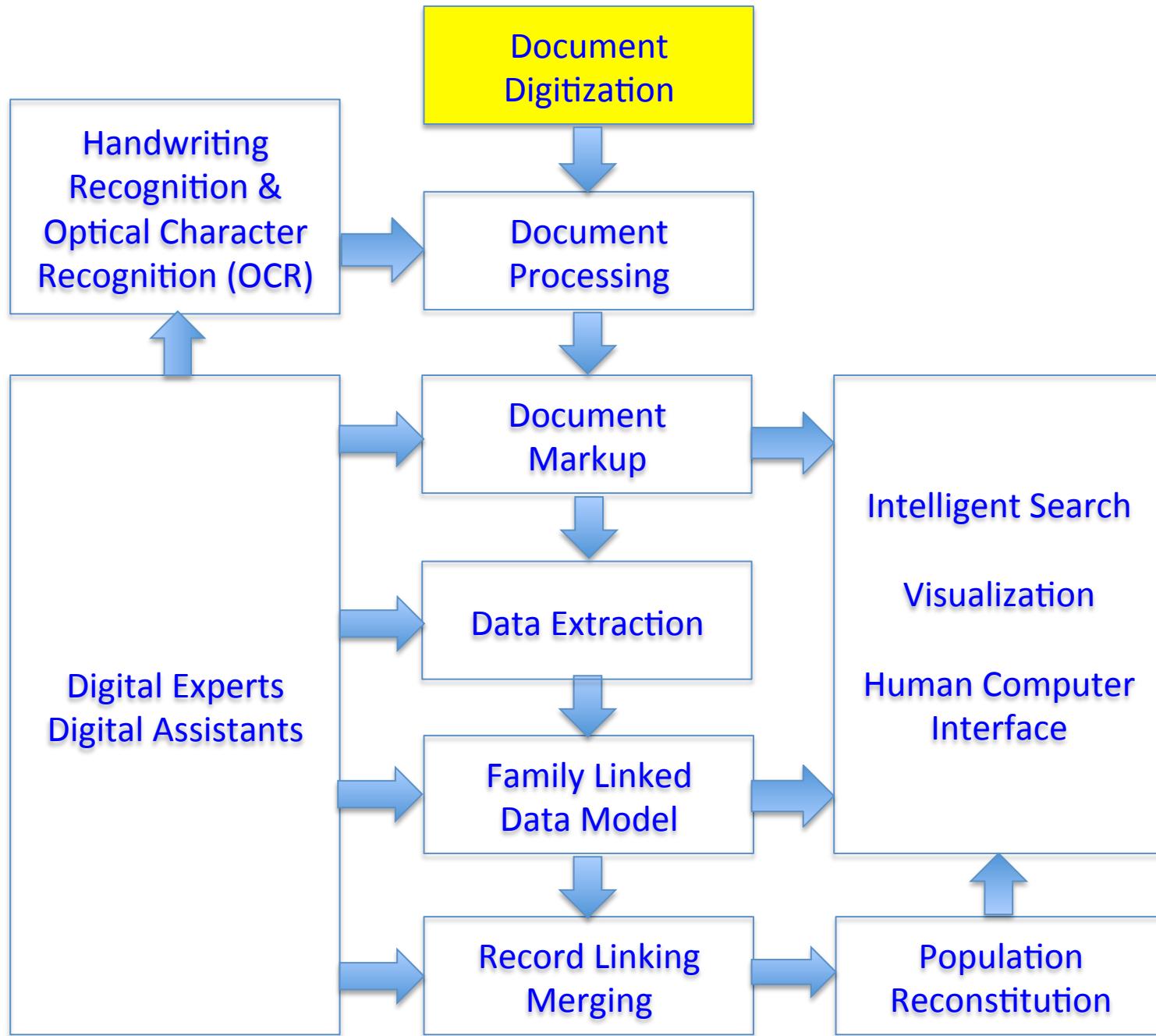
Family History Technology Workshop 2014

# Decomposing a Really Really Hard Problem

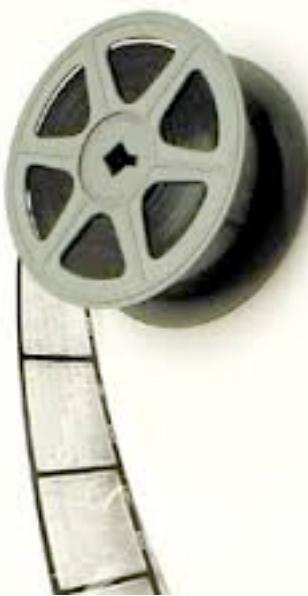
- The goal is to build a common lineage-linked family tree for the whole human race based upon available records.
- Decompose this seemingly impossible challenge into 10 hard problems.
- Allows researchers to focus on a part of the overall problem.
- Illustrated by a running example.

# Family History Technology 10 Hard Problems



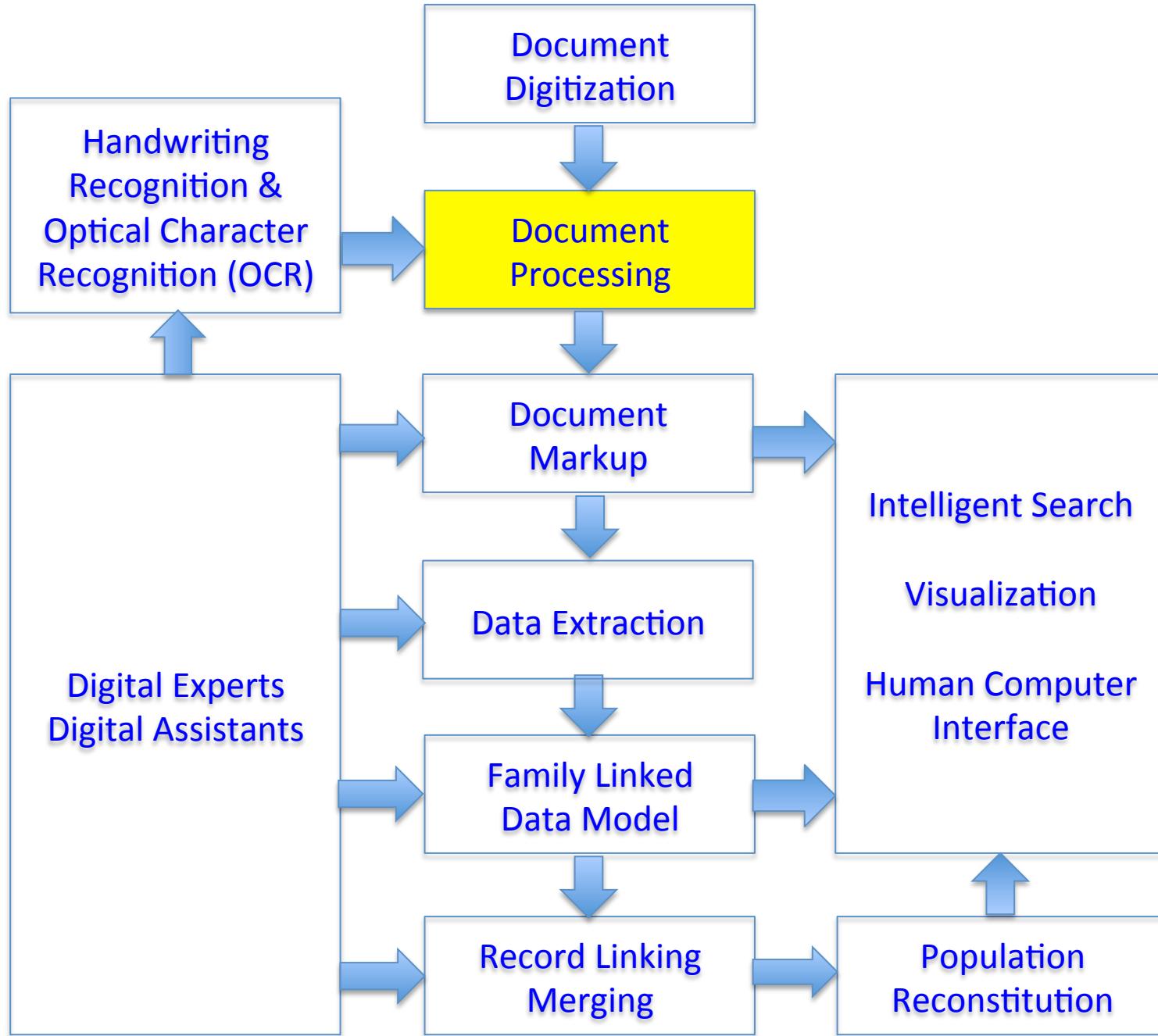


# Document Digitization and Preservation



Preservation,  
Accessibility, and  
Enhancement of  
Records

<p>Juni 1<sup>o</sup> Grunneblich Niels Sorensen døde 29<sup>o</sup> juli 18<sup>o</sup> + 1891. Døb. Ole Jørgen Hansen, Christen Hansen, Peter Hansen, Niels Hansen Hansen Christensen. Anne Hansen, Christen Hansen, Christen Hansen, Ole Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Sønner: Jørgen Hansen, Peter Olsen, Lars Christensen, Christen Hansen, Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p>	<p>Juni 1<sup>o</sup> Niels Hansen døde 1<sup>o</sup> juli. Niels Hansen, Christen Hansen Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen Larsen, Marie Per Hansen og Peder Hansen.</p> <p>12<sup>o</sup> Døb. 1<sup>o</sup> juli 18<sup>o</sup> i Nørbyen. Niels Hansen og Christen Hansen.</p> <p>Juni 13<sup>o</sup> Niels Grunneblich Niels Sorensen i Nørbye døde. Anne Dorte. 14<sup>o</sup> Grunneblich Niels Hansen, Christen Hansen, Christen Hansen, Christen Hansen + 1890. Marie Hansen.</p> <p>Juni 15<sup>o</sup> April Grunneblich Niels Hansen døde Ole Jørgen Hansen, Christen Hansen + 1890. Niels Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p> <p>26<sup>o</sup> Mai Niels Hansen døde 1<sup>o</sup> juli. Niels Hansen, Christen Hansen, Christen Hansen + 1895.</p> <p>9<sup>o</sup> Juni 1<sup>o</sup> Niels Hansen døde Anne. Christen Hansen, Christen Hansen + 1888. Niels Hansen, Christen Hansen, Ole Pedersen, Niels Hansen.</p> <p>10<sup>o</sup> Juli Grunneblich Peter Hansen døde Ole Hansen i Nørbyen i Nørbyen. Niels Hansen, Christen Hansen, Christen Hansen, Christen Hansen + 1891. Niels Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p> <p>28<sup>o</sup> Maj Jacob Rejchels døde Anne Sophie Magdalene Sønner: Christen Hansen, Niels Nielsen, Hans Hansen Møllentau, Anne Hansen.</p> <p>Edem fættur. Ole Hansen døde 1<sup>o</sup> juli. Else. Sønner: Niels Hansen, Ole Hansen, Tøns Pedersen, Anders Hansen Anne Tøns Pedersen.</p> <p>21<sup>o</sup> Helligd. Et 1<sup>o</sup> juli Grunneblich Ingens døde 1<sup>o</sup> juli. Ellen Jensen 12<sup>o</sup> Sønner: Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p> <p>19<sup>o</sup> p. Trinck: Ole Hansen Hansen Christensens døde 1<sup>o</sup> juli. Marie Sønner: Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p> <p>21<sup>o</sup> p. Trinck: Ole Hansen Hansen Christensens døde 1<sup>o</sup> juli. Marie Sønner: Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p> <p>27<sup>o</sup> p. Trinck: Ole Hansen Hansen Christensens døde 1<sup>o</sup> juli. Ole Hansen Sønner: Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p> <p>4<sup>o</sup> Trinck: Ole Hansen Hansen Christensens døde 1<sup>o</sup> juli. Ole Hansen Sønner: Christen Hansen, Christen Hansen, Christen Hansen, Christen Hansen Jørgen Hansen, Christen Hansen, Christen Hansen, Christen Hansen.</p>
---	---

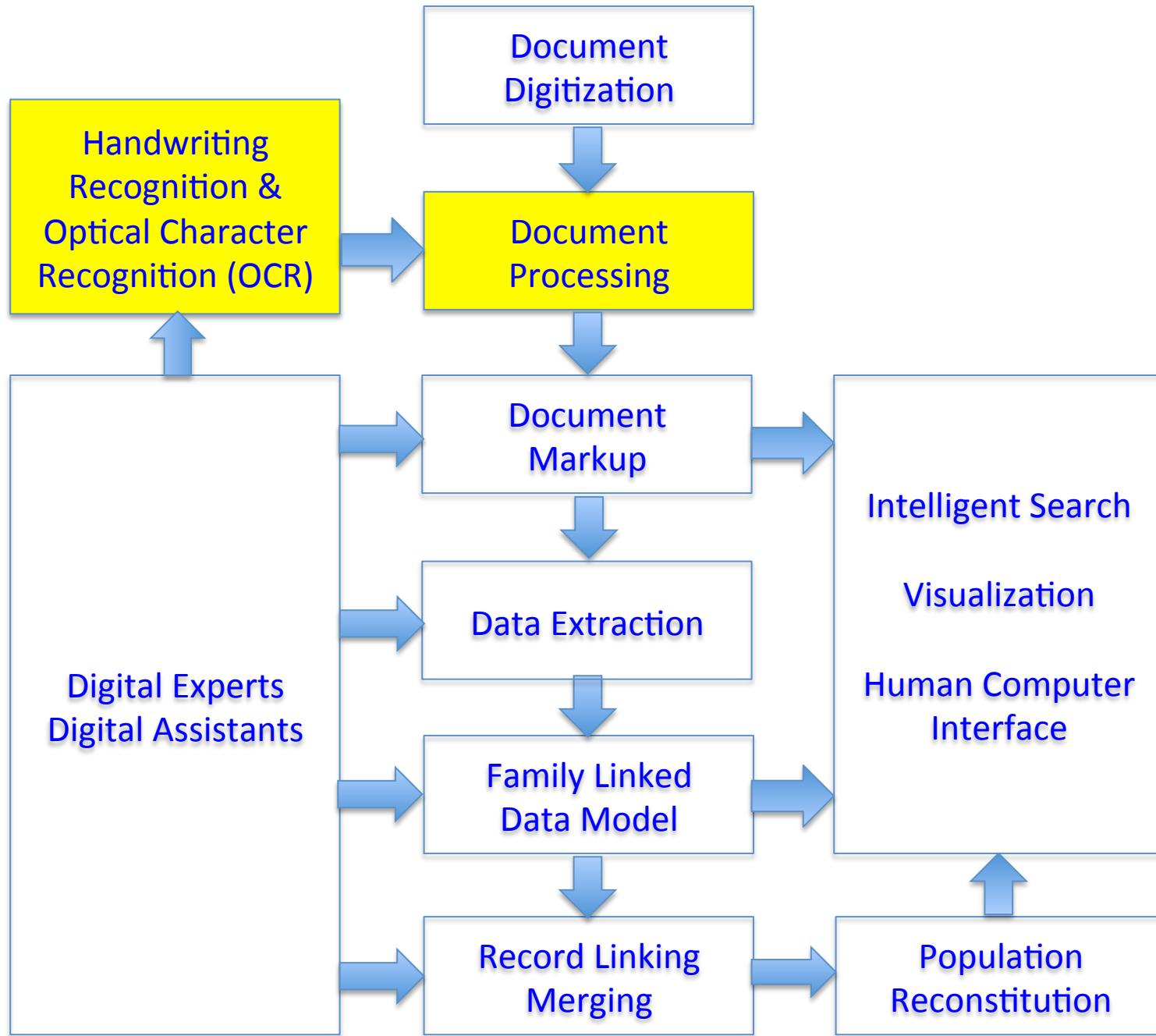


# Document Processing

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

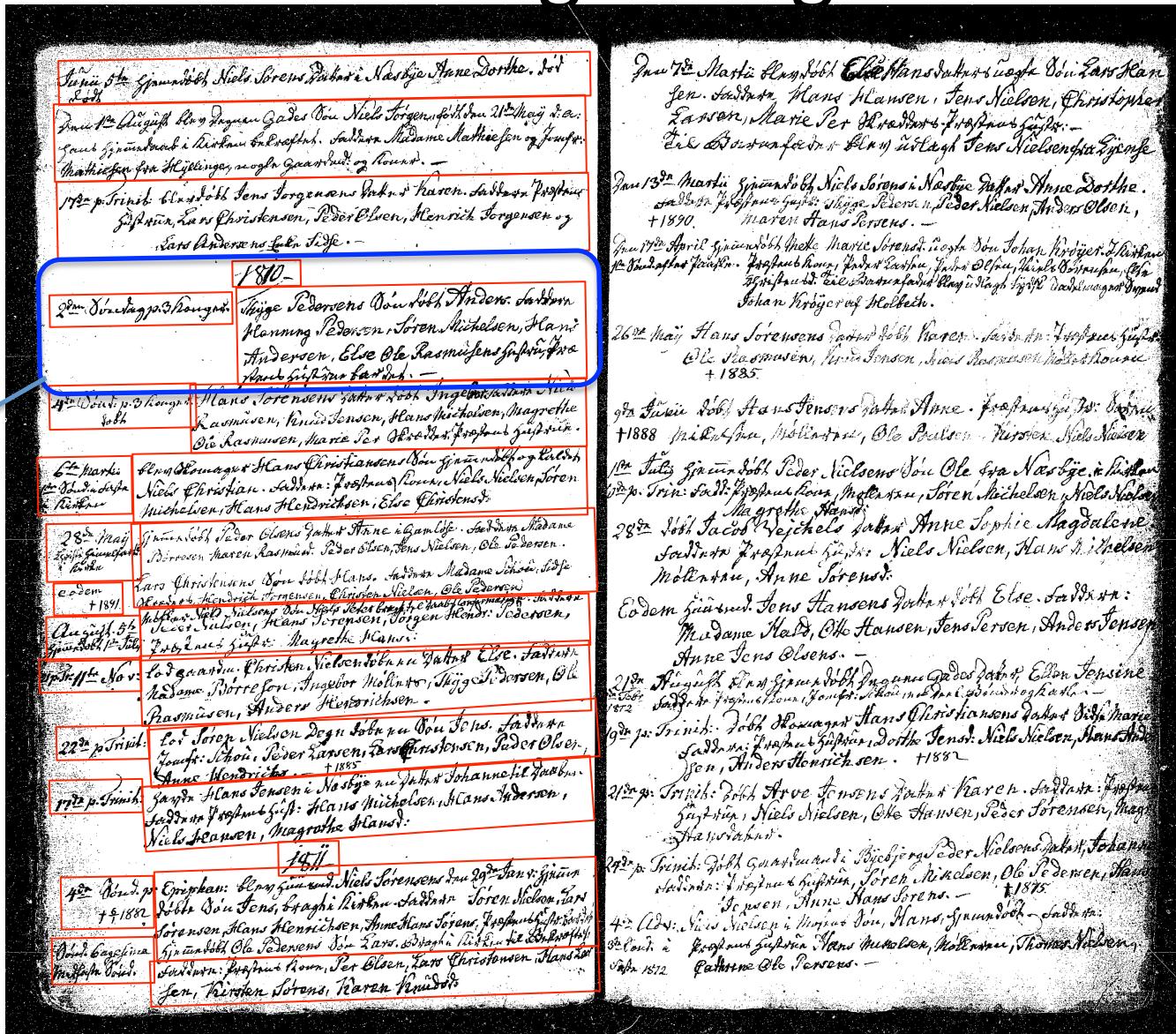
Zoning.

<p>Jan 1<sup>st</sup> 1800 Niels Sorensen datter i Nørreby Anne Dorthe. Døbt 1800</p> <p>Jan 12<sup>th</sup> August 1809 Døgnen gades bøn Niels Sørgen i København 11<sup>th</sup> maj 1809. Guds givne vort i Kristiania biskopshof. Sønnen Nielsme Mathiesen og frue Mathiesen fra Hjelinge, nogen gennemkaldt af Knud.</p> <p>17<sup>th</sup> pr. Trinck bladstikkens Tøns Jørgensens børn i København. Sønnen Peder Jørgensens, Lars Christensen, Peter Olsen, Henrik Jørgensen og Lars Andersen på den side.</p>
1800
<p>Den Torstagsprædikant</p> <p>Hellig Pedersens Døbning i Frederiksberg Kl. 10.00. 1800</p> <p>14<sup>th</sup> Døgn 1<sup>st</sup> Kongens Hånd. Niels Søbh Larsen, Niels Sørensen, Hans Christensen, Magdalene Die Rasmussen, Maria P. Petersen, Peder Jørgensens.</p>
<p>16<sup>th</sup> prædikant i Frederiksberg Kl. 10.00 Hans Christianus Dieg. Sørensen i Frederiksberg Niels Christian. Sørensen: Peder Sørensen, Niels Nielsen Christensen, Hans Hvidrichsen; Else Christensen.</p>
<p>28<sup>th</sup> Mai 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 Peder Sørensen, Niels Nielsen, Ole Hansen, Ole Petersen. Jørgen Hansen, Lars Christensen, Niels Sørensen, Ole Petersen. Ole Hansen, Niels Nielsen, Ole Petersen.</p>
<p>August 5<sup>th</sup> 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 Peder Sørensen, Niels Nielsen, Ole Petersen, Ole Petersen, Jørgen Hansen, Hans Christensen, Søren Hansen, Ole Petersen, Niels Nielsen, Ole Petersen.</p>
<p>28<sup>th</sup> prædikant: Niels Sørensen døbning i Frederiksberg Kl. 10.00 Peder Sørensen, Niels Nielsen, Ole Petersen, Ole Petersen, Anne Hvidrichsen. +1809</p>
<p>17<sup>th</sup> pr. Trinck: Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen, Jørgen Hansen, Lars Christensen, Niels Nielsen, Ole Petersen, Jørgen Hansen, Ole Petersen, Ole Petersen, Ole Petersen.</p>
1801
<p>4<sup>th</sup> Februar: Ole Christensen, Niels Sørensen døbning i Frederiksberg +1802 Bøn bøn Niels Sørensen, Niels Nielsen, Ole Petersen, Ole Petersen, Niels Nielsen, Ole Petersen, Ole Petersen, Ole Petersen.</p>
<p>10<sup>th</sup> Februar: Ole Christensen, Niels Nielsen, Ole Petersen, Ole Petersen, Niels Nielsen, Ole Petersen, Ole Petersen, Ole Petersen.</p>
<p>Jan 7<sup>th</sup> Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen, Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>Jan 13<sup>th</sup> Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen, Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>Jan 17<sup>th</sup> April 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>26<sup>th</sup> Mai: Hans Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>9<sup>th</sup> Jule 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 +1808 Niels Nielsen, Ole Petersen, Ole Petersen, Ole Petersen.</p>
<p>10<sup>th</sup> Jule 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>28<sup>th</sup> Mai Jacob Weychels døbning Anne Sophie Magdalene Sørensen: Peder Sørensen, Niels Nielsen, Ole Petersen, Møllerhus, Anne Sørensen.</p>
<p>Eodem færdig. Niels Hansen døbning i Frederiksberg Kl. 10.00: Niels Nielsen, Ole Hansen, Ole Petersen, Anders Jensen, Anne Ibsen Olsens.</p>
<p>28<sup>th</sup> August 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 Ellen Petersen Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>18<sup>th</sup> September 1809 Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>19<sup>th</sup> pr. Trinck: Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>21<sup>st</sup> pr. Trinck: Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>27<sup>th</sup> pr. Trinck: Niels Sørensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>
<p>4<sup>th</sup> Maj 1810 Ole Christensen døbning i Frederiksberg Kl. 10.00 Jørgen Hansen, Niels Nielsen, Ole Petersen, Ole Petersen.</p>



# Document Processing and Handwriting Recognition

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



# Document Processing and Handwriting Recognition

1810-  
2<sup>den</sup> Søndag p. 3 Konger. Thyge Pedersens Son dobt Anders. Faddern  
Henning Pedersen, Soren Michelsen, Hans  
Andersen, Else Ole Rasmusens hustru, Præ  
stens hustrue bar det.—

[Annotated transcript]

[Christening Record]

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

[Image 112]

[Heading] 1810

[Margin Note] 2<sup>den</sup> Sondag p. 3 Konger.

[Body]

[Line 1] Thyge Pedersens Son dobt Anders. Faddern

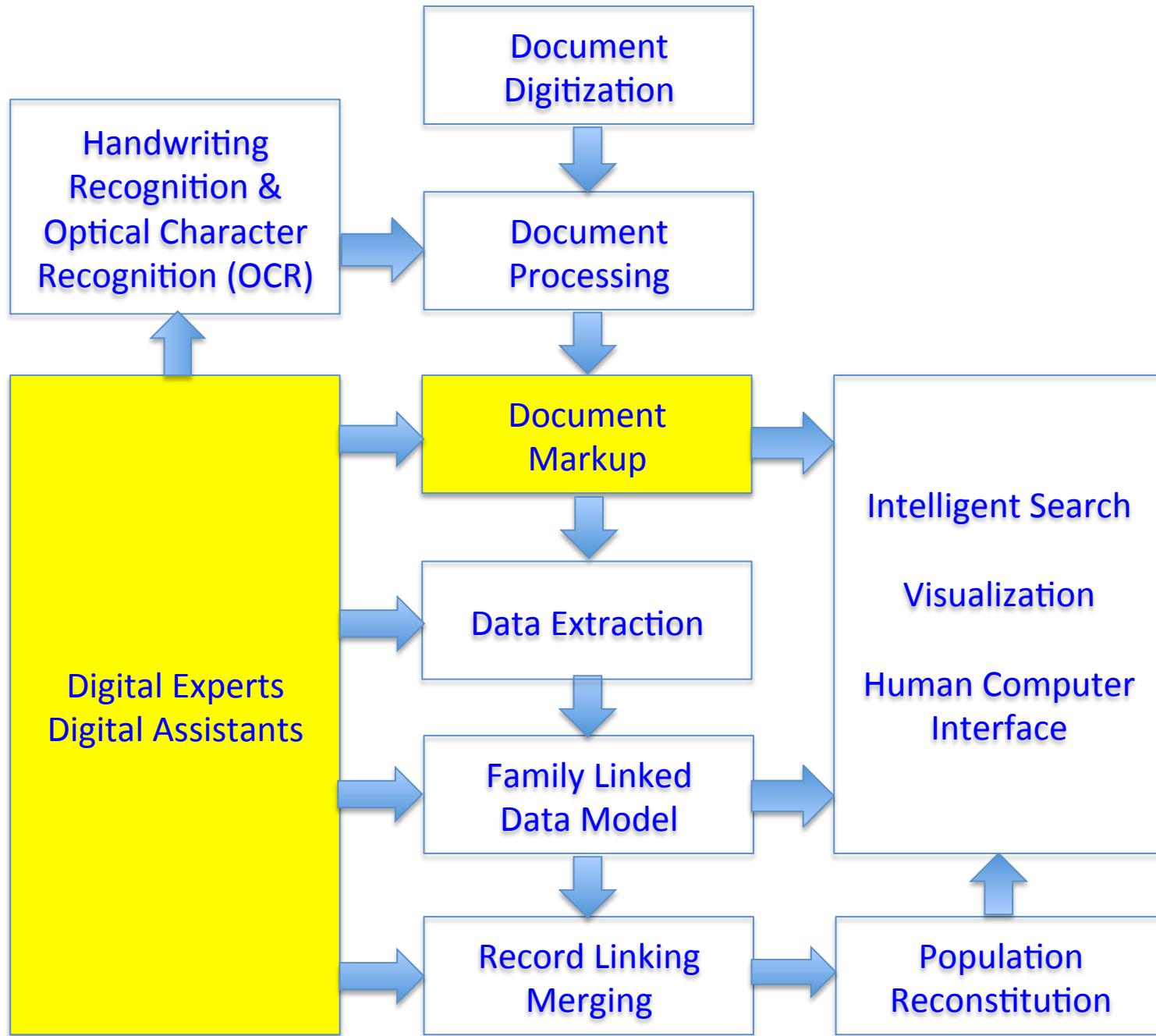
[Line 2] Henning Pedersen, Soren Michelsen, Hans

[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]

[Heading] 1810

[Margin Note] 2<sup>den</sup> 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.—

Categorize text into semantic elements: names, dates, places, etc. Adds markup to the document to convey this information.

Uses Date Authority  
Uses Name Authority  
Uses Place Authority

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

(English  
Translation)

[Image 112]

[Heading] 1810

[Margin Note] 2<sup>nd</sup> Sunday past 3 Kings.

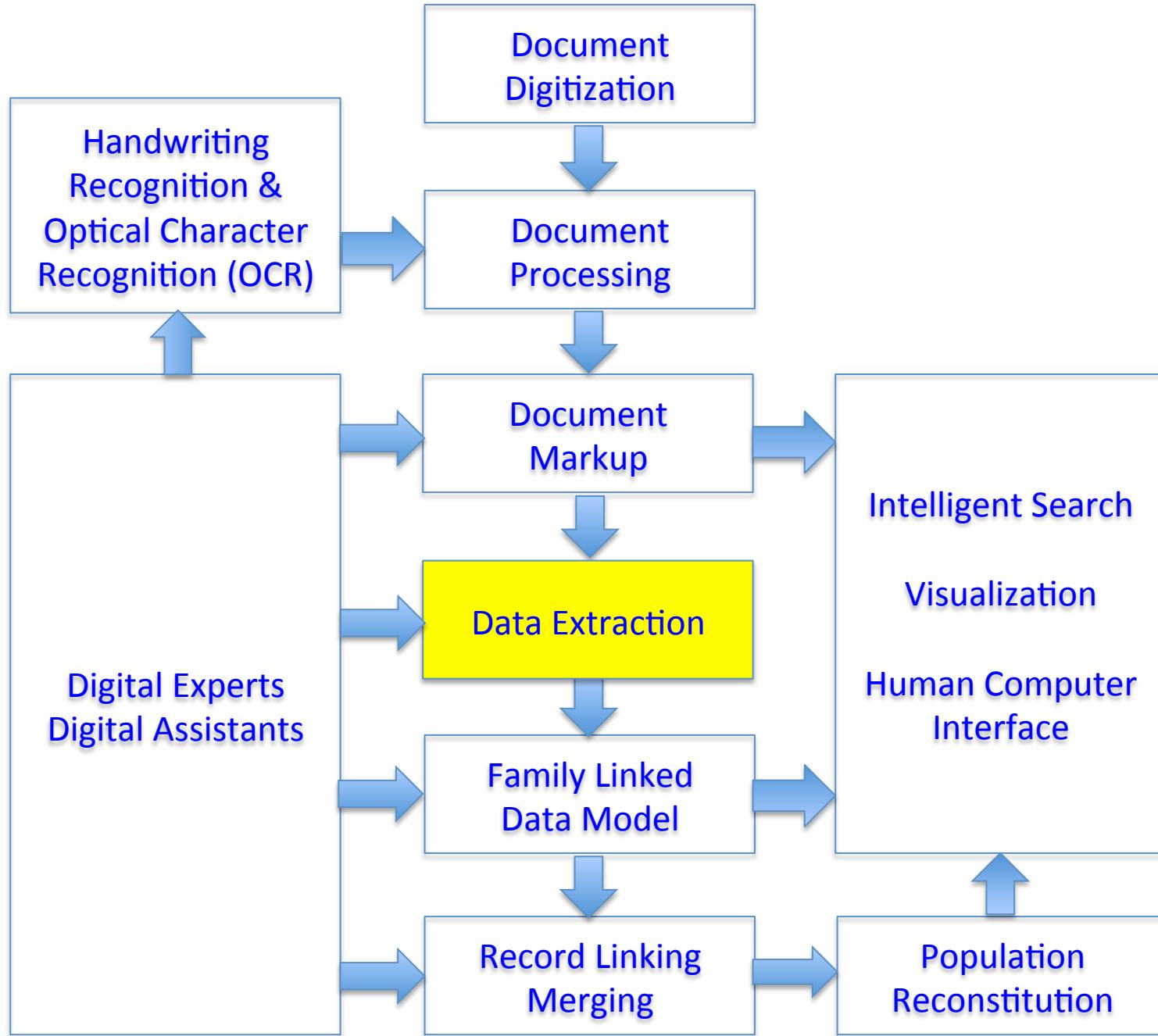
[Body]

[Line 1] Thyges Pedersen's Son christened Anders. Witnesses

[Line 2] Henning Pedersen, Soren Michelsen, Hans

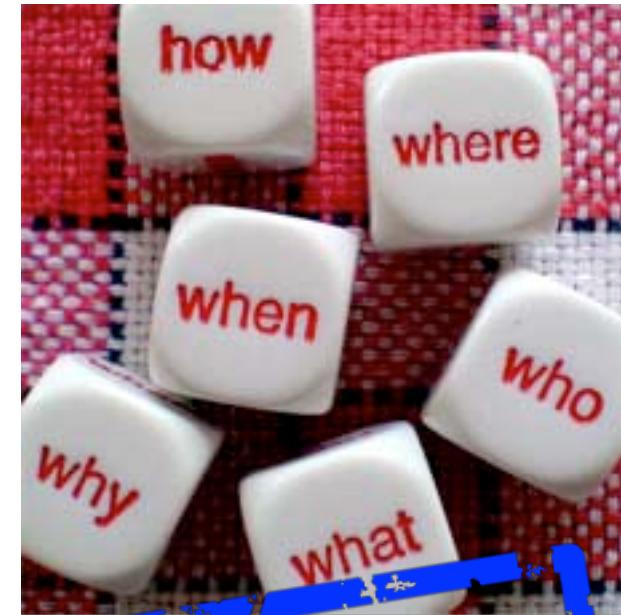
[Line 3] Andersen, Else Ole Rasmusen's wife, the Priest's

[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)

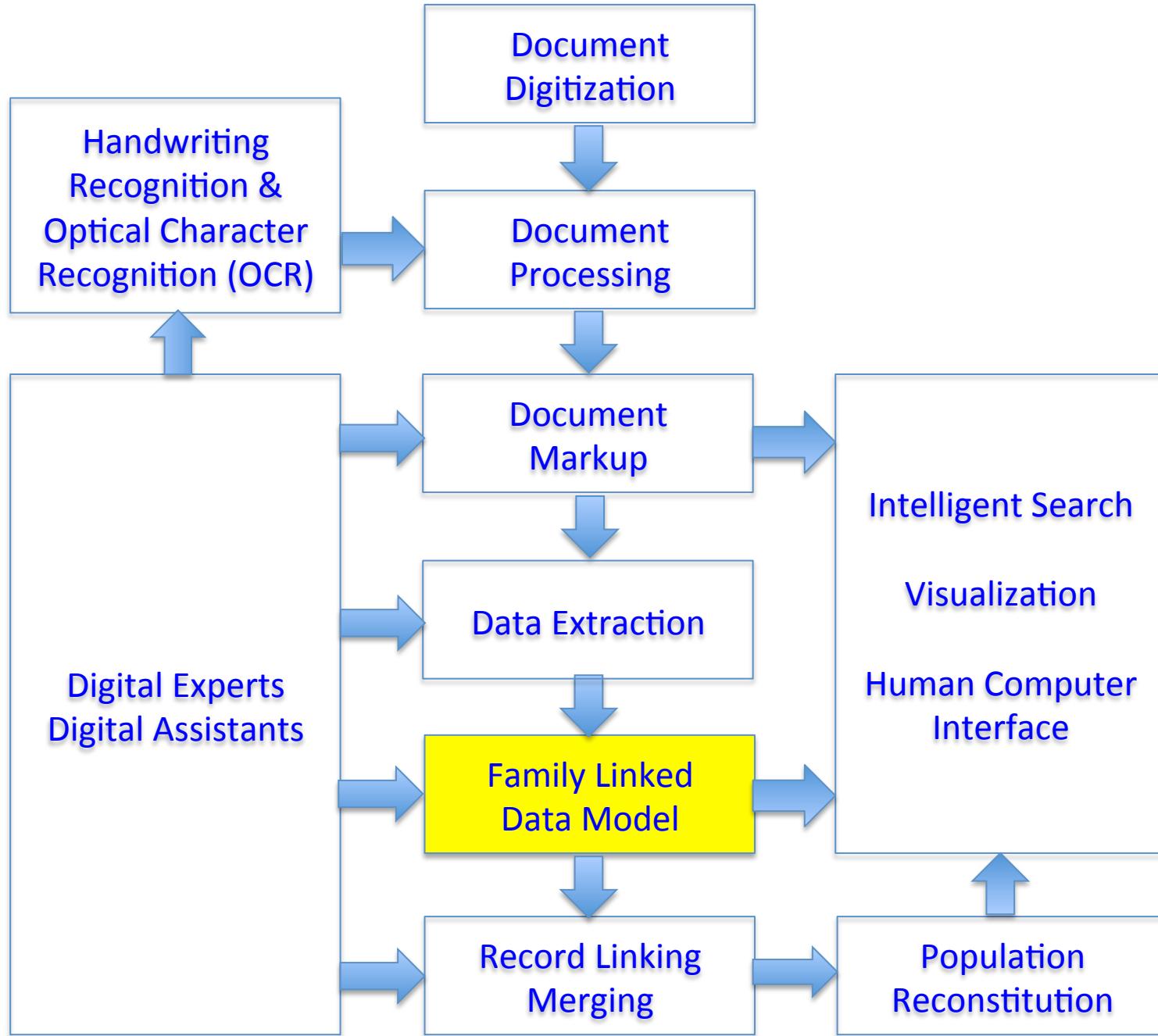


# Data Extraction

## 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]
- 2<sup>nd</sup> Sunday past 3 Kings Day in 1810 is January 14, 1810. [Date Authority]
- Anders Thygesen was christened
- Thyges Pedersen was the father of Anders Thygesen
- Henning Pedersen was a witness to the christening of Anders Thygesen
- Soren Michelsen was a witness to the christening of Anders Thygesen
- Hans Andersen was a witness to the christening of Anders Thygesen
- Else was a witness to the christening of Anders Thygesen
- Else was the wife of Ole Rasmusen
- The Priest's wife was a witness to the christening of Anders Thygesen

(12 Facts from this simple Christening record.)

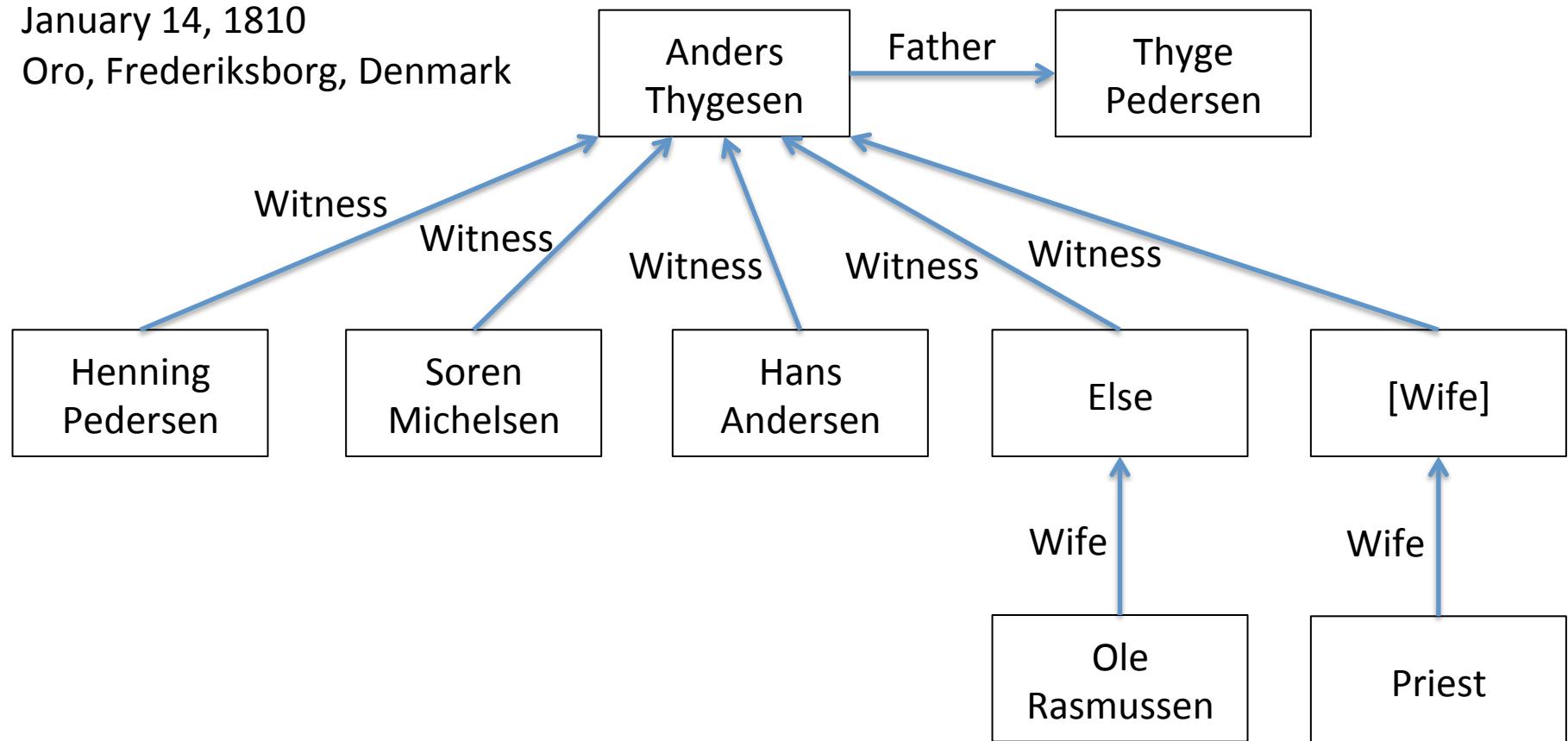


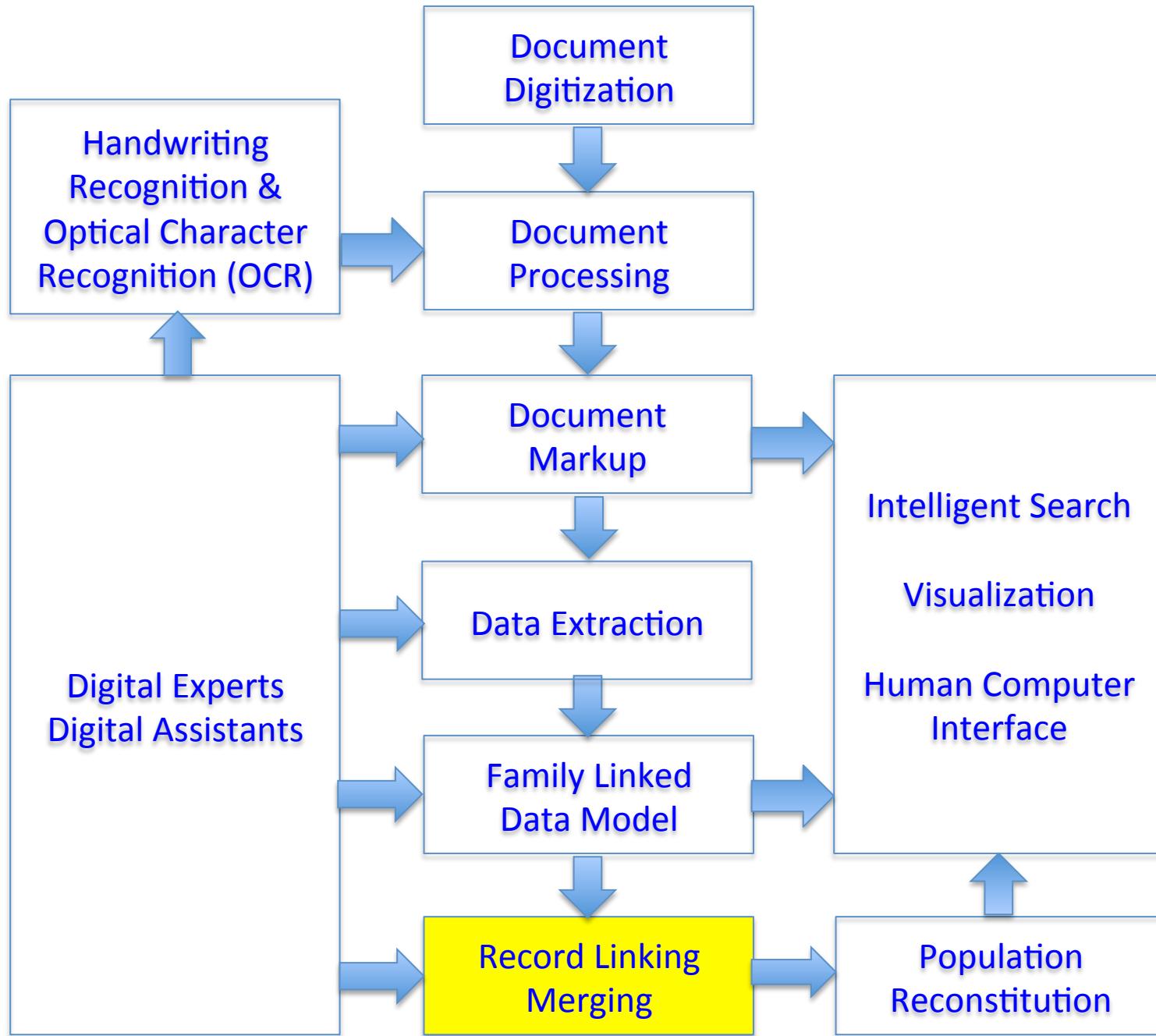
# Family Linked Data Model

Christening Event

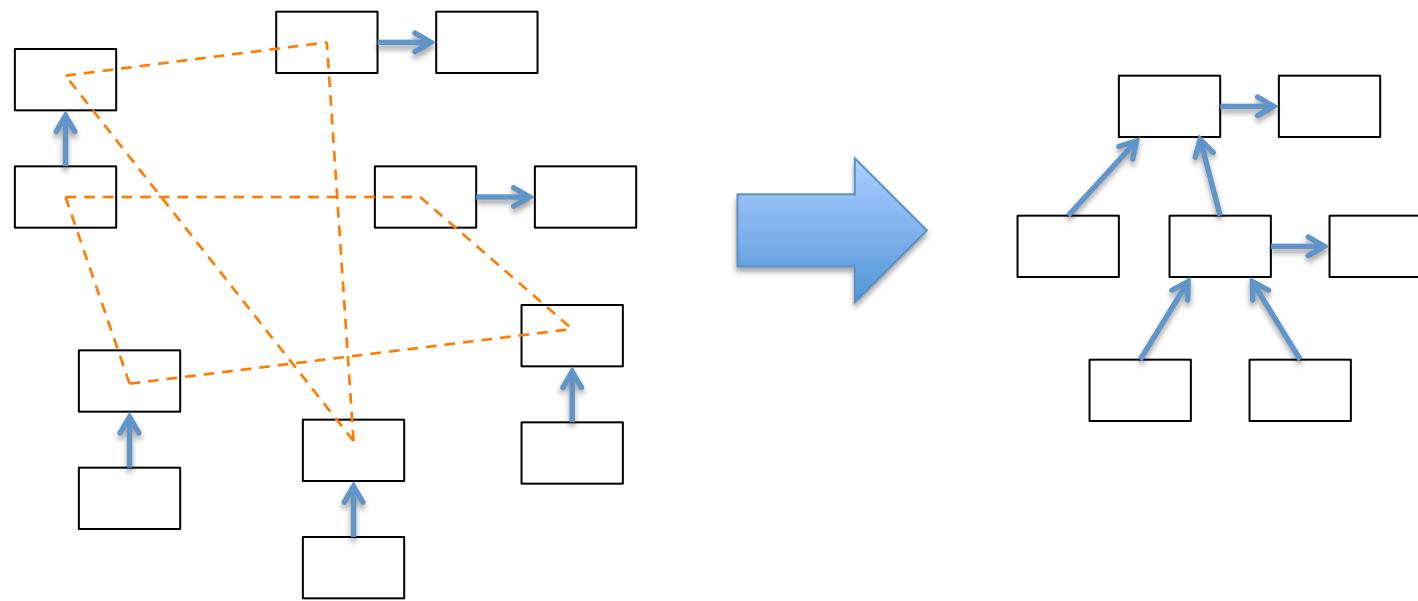
January 14, 1810

Oro, Frederiksborg, Denmark



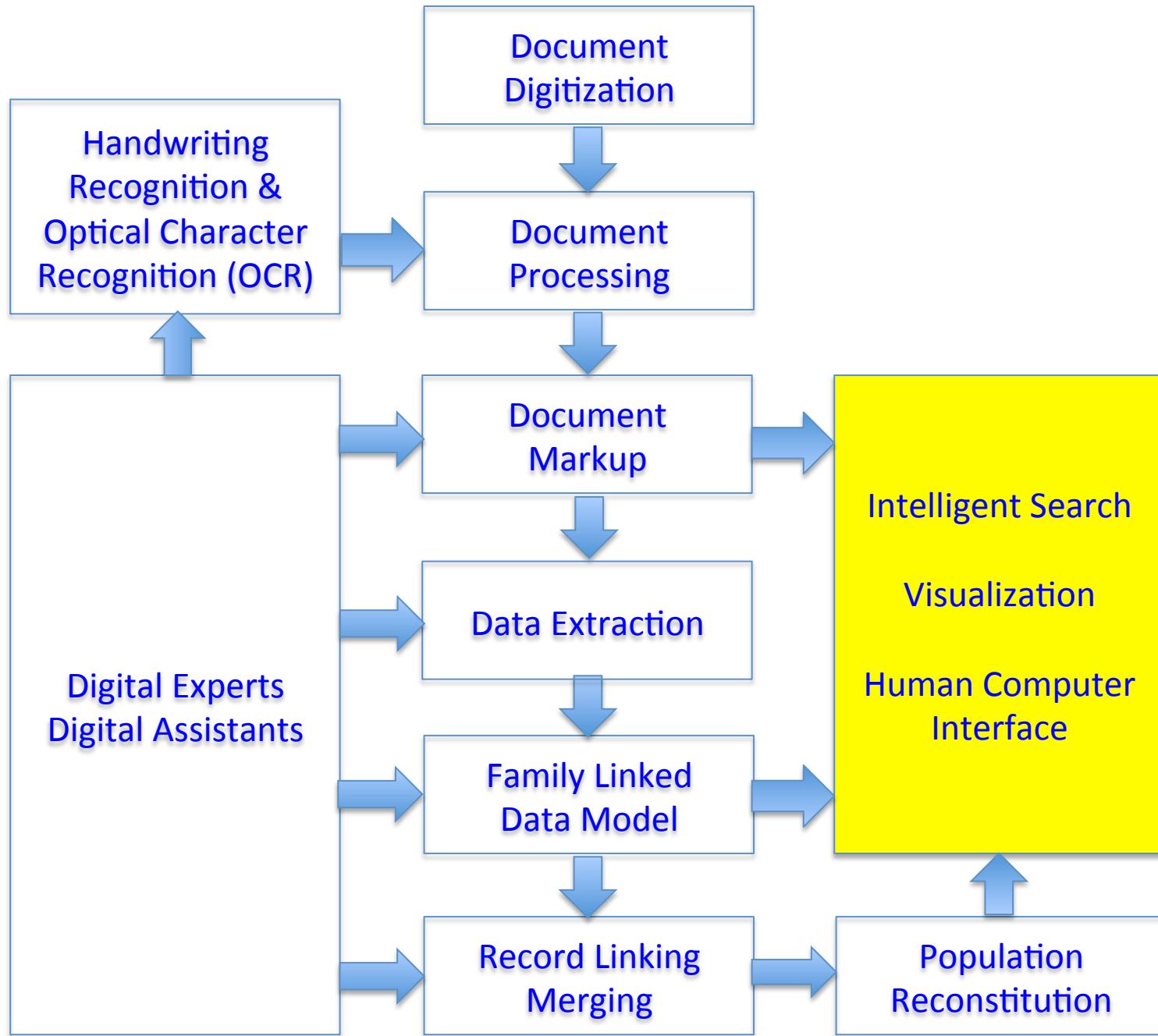


# Record Linkage/Merging

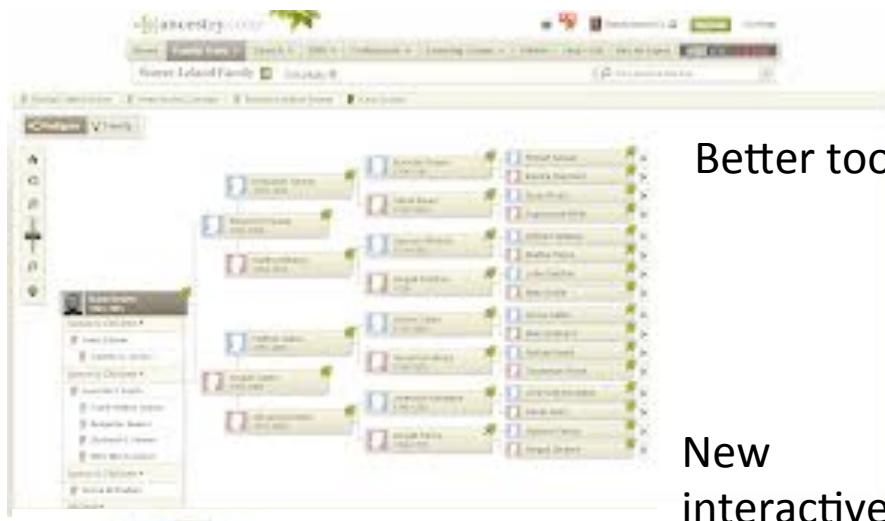


Record linking determines when two or more different records refer to the same individual.

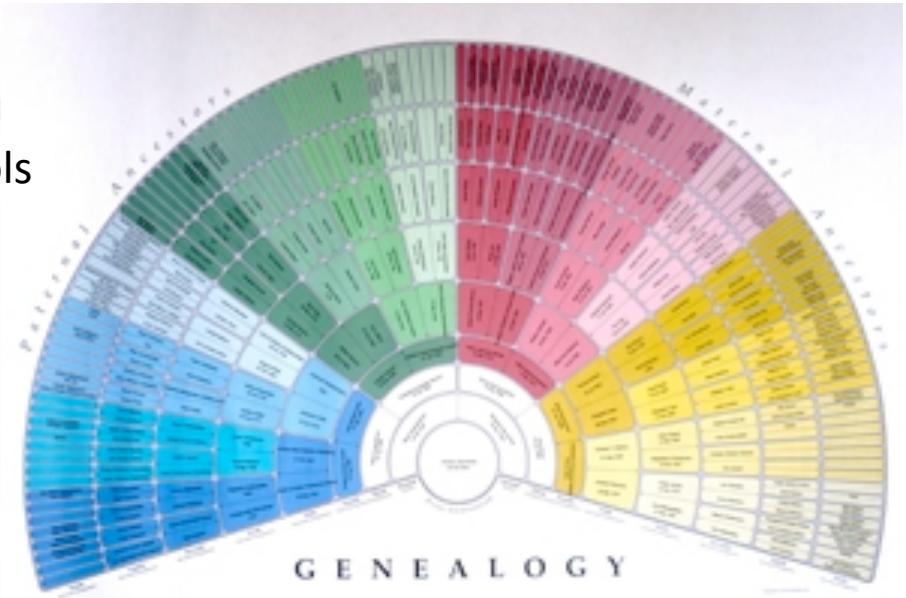
Duplicates are merged together to build lineage-linked families.



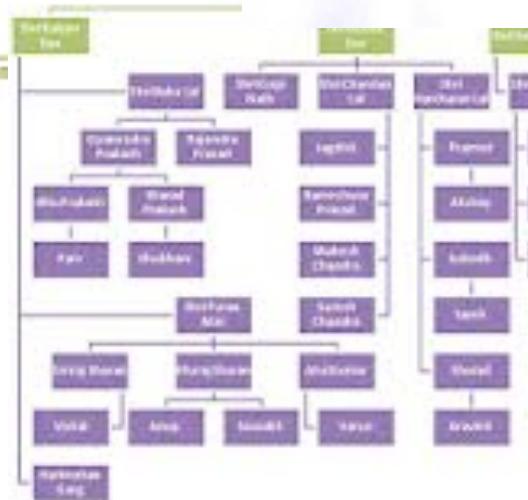
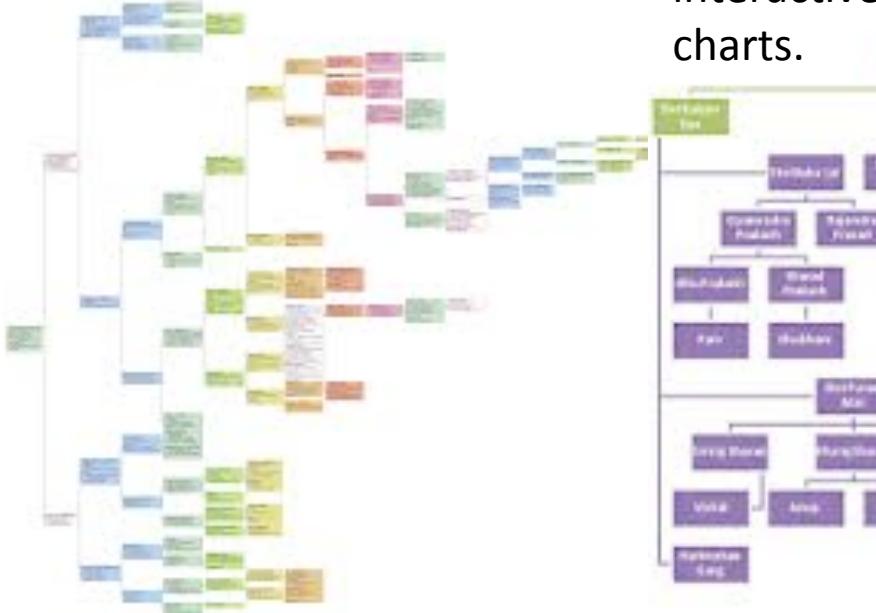
# Search, Visualization & Computer-Human Interface



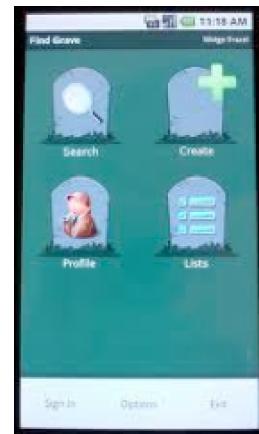
Better tools



New  
interactive  
charts.



Mobile apps



# 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.