Family History Research on the Semantic Web:

Building a Semantic Prototype for Danish Genealogical Research

By
Charla Woodbury and David W. Embley
BYU Computer Science Department
charlajw@cs.byu.edu   embley@cs.byu.edu

Family History Technology Institute
March 24, 2005

Semantic Web
Machine "Understandable" Web

MEANING
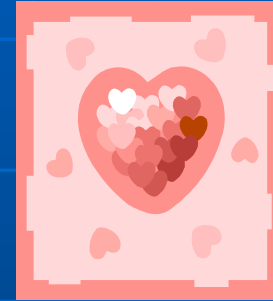
KNOWLEDGE

INFORMATION

DATA

Need for Semantic Web

"The Semantic Web: ... content that is meaningful to computers [and that] will unleash a revolution of new possibilities ... Properly designed, the Semantic Web can assist the evolution of human knowledge ..."

(Tim Berners-Lee, ..., Weaving the Web)

3

Semantic Web
'DATE'

Calendar date

To date an artifact

A fruit

A romantic experience

To go on a romantic experience with someone

4

Also a SURNAME – Mr. C. J. Date**

The semantic web will make it possible for machines to know the difference!

**** Edgar F. Codd and C. J. Date are famous in the area of databases for defining levels of normal forms**

Real Problem

Amtsinddelingen 1793-1970

A person decides to do family history research for the first time on their Danish family lines.

- Where do they go?
- What records do they look for?
- How do they handle records in Danish?
- How can they tell when the records they have match their search family?

Semantic Web
Ideal for Family History

SOLUTION – PROTOTYPE

- **The heart of a one-stop web site for naïve researchers**

- **So many records have been extracted into digitized forms and are often available on the Web**

- **Limited geographically – parish and probate records from Nim District, Skanderborg, Denmark**
  - 100% probates
  - 100% marriages

Semantic Web Prototype

Ontology – semantic model
   (BYU Ontos)

Annotated web pages

(Web Ontology Language OWL proposed W3C Feb 2004)

Solutions for special genealogical problems

Ontology Model
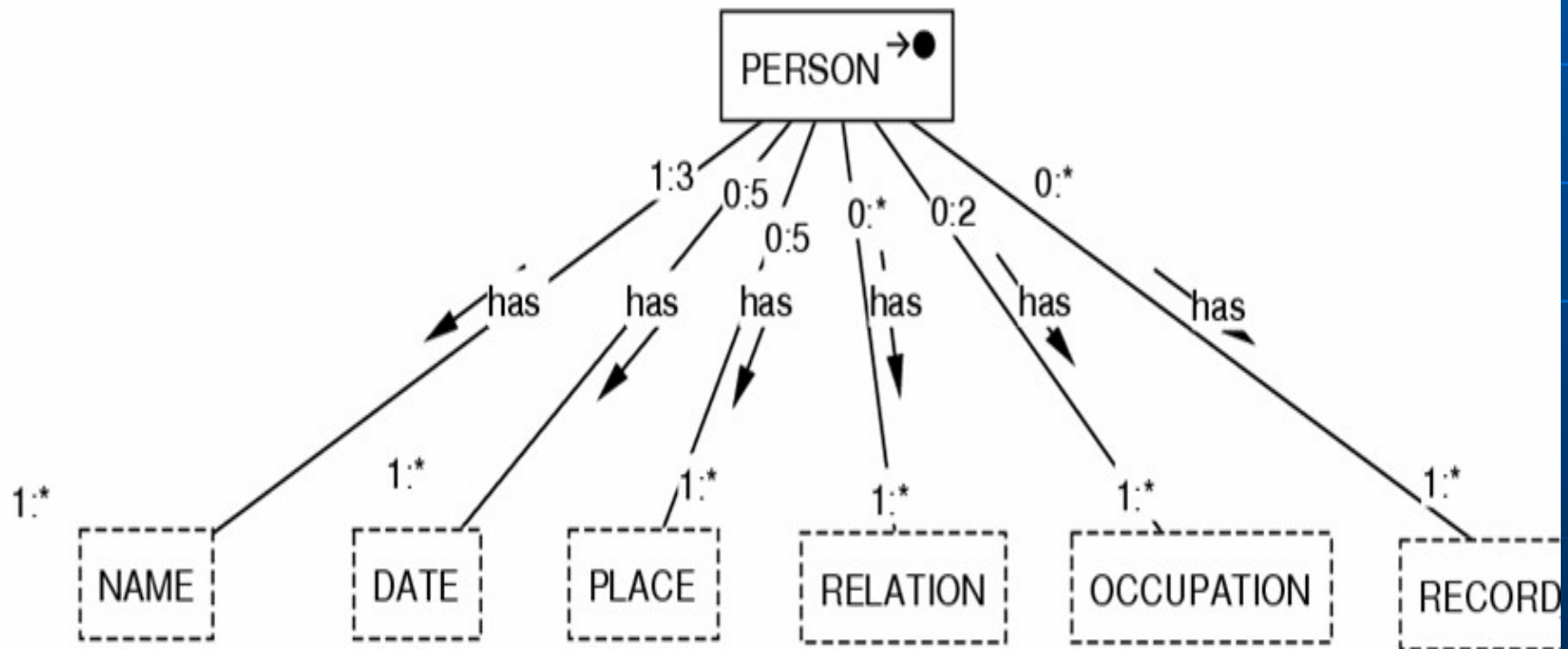
Person Matching
in genealogical research

- **NAMES**

- **DATES**

- **PLACES**

- **RELATIONS**

Ontology Entities

FIND and MARK UP relevant web pages by:

- NAME   \<NAME\>
- DATE   \<DATE\>
- PLACE   \<PLACE\>
- RELATIONSHIP   \<RELATION\>
- OCCUPATION   \<OCCUPATION\>
- RECORD_TYPE   \<RTYPE\>
- SOURCE   \<SOURCE\>

Partial Danish GIVEN NAME LEXICON

MALE
- And.
- Anders
- Andreas
- Christen
- Christian
- Eric
- Erik
- Gregers
- Hans
- Ib
- Jacob
- Jens
- Jep

FEMALE
- Ane
- Anna
- Anne
- Birthe
- Birte
- Bodil
- Caroline
- Dorte
- Dorthe
- Elene
- Ellen
- Elisabeth
- Elsbeth

Partial DATE Lexicon

(actual lexicon is a single list in alphabetic order)

MONTHS

- January –Jan –Januar -11br
- Februrary –Feb –Februar -12br
- March –Mar –Marts
- April – Apr –Apl
- May –Mai
- June –Jun –Juni
- July –Jul –Juli -5br
- August –Aug –Augst -6br
- September –Sep –Sept -7br –Septembre
- October –Oct -8br –Octobre
- November –Nov -9br –Novembre
- December –Dec -10br -Decembre

TIME

- Year –yr –aar –år
- Month –mo –maaned –måned –m.
- Week –uge –ug.
- Day –dag –dg.
- Hour – h. –hr.

FEAST DATES (partial)

- Easter – Paaske –Påske –Paasche –Påsche
- Pentecost – Pent –Pinse -Pin
- Trinity –Tr –Trin –Trinitatis

DAYS OF WEEK

- Sunday –Dominico –Dom.
- Monday –Mondag –Mond.
- Tuesday –Tirsdag –Tirsd.
- Wednesday -Onsdag –Onsd.
- Thursday –Tørsdag –Tørsd.
- Friday –Fredag –Fred.
- Saturday –Lørsdag –Lørs.

Original Record
FHL Film#052,236 Tvilum Parish

Web Page

- SOURCE URL -Tvilum Sogne Kirkebog

- [PAGE HEADER]   Fødde 1751 3

- [BODY] Truust Dom. 23 p: Trinit: laest over Niels Baches SØREN fadd. Johannes Michelsens og Niels Mollers hustruer af Søebyevad, Peder Rasmussen af Søebyevad, Jens Bachis søn Peder og Niels Thylkes s. Peder af Truust

Ontology Entities

FIND and MARK UP relevant web pages by:

- NAME   <NAME>
- DATE   <DATE>
- PLACE   <PLACE>
- RELATIONSHIP <RELATION>
- OCCUPATION   <OCCUPATION>
- RECORD_TYPE <RTYPE>
- SOURCE<SOURCE>

Colors only represent OWL annotation mark-ups automatically placed in the web page using the ontology

Annotated Web Page

- SOURCE -Tvilum Parish Register

- [PAGE HEADER]   Fødde 1751 3

- [BODY] Truust Dom. 23 p: Trinit: laest over Niels Baches SØREN fadd. Johannes Michelsens og Niels Mollers hustruer af Søebyevad, Peder Rasmussen af Søebyevad, Jens Bachis søn Peder og Niels Thylkes s. Peder af Truust

Results Listing

TARGET – **Jens Pedersen Bach**
Truust, Tvilum Parish, Gjern District, Skanderborg
Date Range - born 1693 to died 1778

Name	Date	Place	Relation	Occupation	Record Type	Source (URL)
Jens Bachis	Dom. 23 p: Trinit: 1751 (14 Nov 1751)	Truust	fadd:		Fødde	Tvilum Parish Register

SOURCE -Tvilum Parish Register
[PAGE HEADER]   Fødde 1751 3
[BODY] Truust Dom. 23 p: Trinit: laest over Niels Baches SØREN
fadd. Johannes Michelsens og Niels Mollers hustruer af Søebyevad,
Peder Rasmussen af Søebyevad, **Jens Bachis** søn Peder og
Niels Thylkes s. Peder af Truust

Conversion Functions inside the ontology

- Compute birthdate from age at death

  Death – 22 Mar 1743

  Age - 23 yr 2 m

  ## -> BIRTH Jan 1720

- Compute dates from feast dates

  Sunday 23rd after Trinity 1751

  ## -> 14 Nov 1751
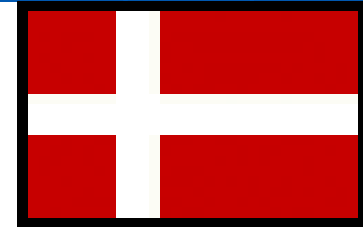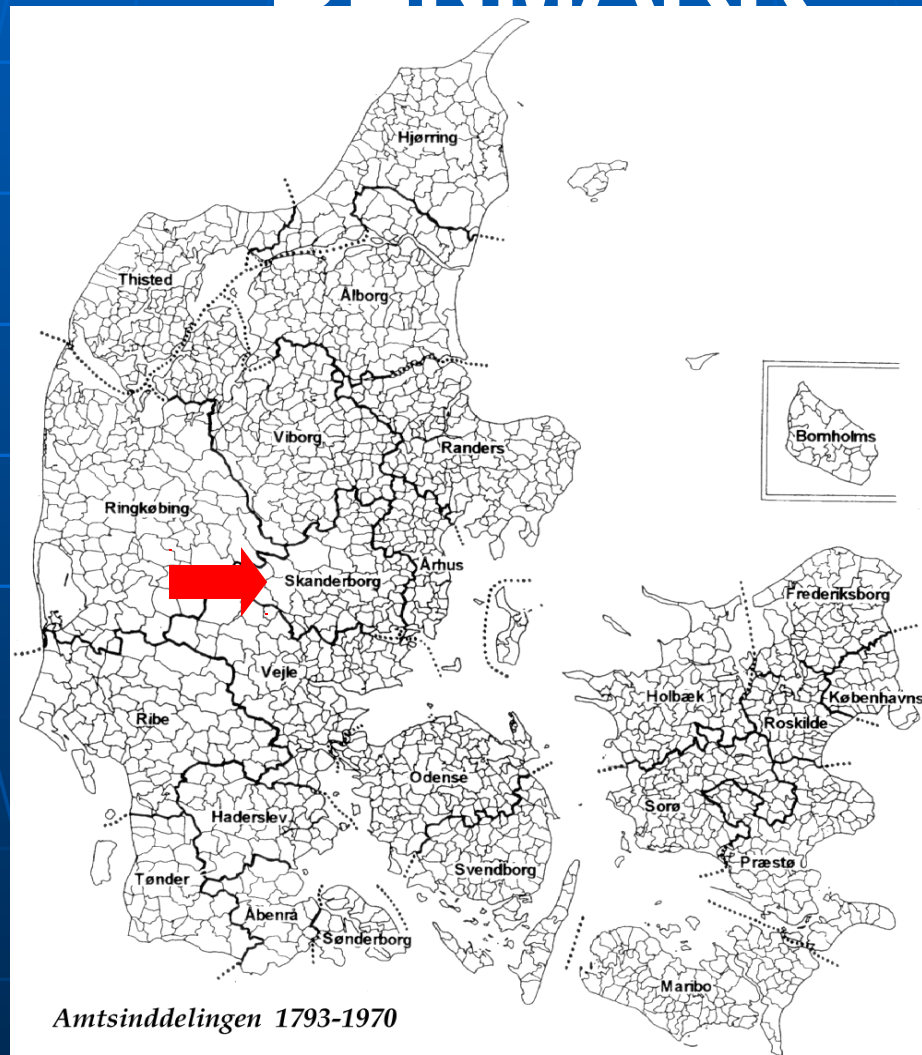
Solutions for Special Problems

RULES FOR

- [Matching different name forms](#)
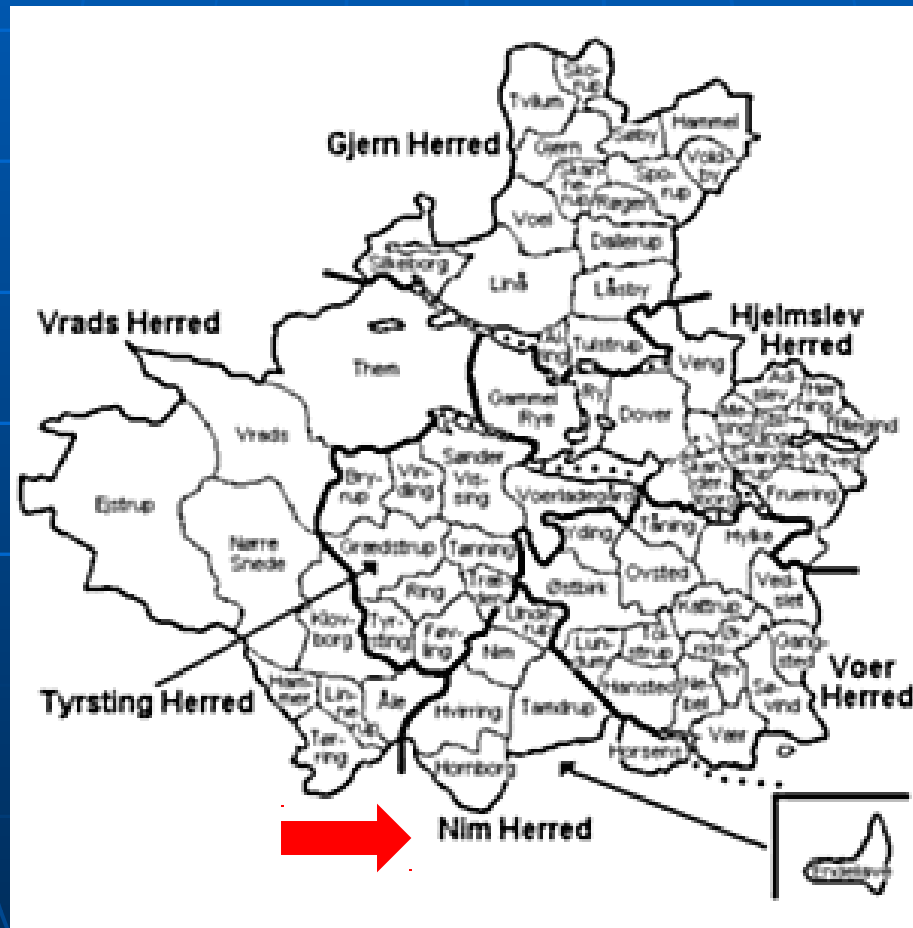
- [Matching place names to appropriate records](#)

RULE - Match different name forms as ONE PERSON

- JENS PEDERSEN

- JENS PEDERSEN BACH

- JENS BACH

- JENS BACHIS

PLACES - County Map of DENMARK

Amtsinddelingen 1793-1970

Parish and District Map of SKANDERBORG

Road Map
www.expedia.com

Matching Places to Records

Farm name	Parish	District	County	Record Links
Molger	Tamdrup	Nim	Skanderborg	**PARISH** Tamdrup 1684-1912 **PROBATE** Nim Herred Provisti Rask Skanderborg Rytterdistrikt
	Tamdrup	Nim	Skanderborg	**List of URL's** **Includes Molger URL's** **Adds Parish specific records**
		Nim	Skanderborg	**List of URL's** **Includes Tamdrup URL's** **Adds District specific records**
			Skanderborg	**List of URL's** **Includes all district URL's** **Adds County specific records**

MAJOR CONTRIBUTIONS

- **First genealogical prototype for the semantic web**
  - FOCUS on primary records
  - Not just an index of the records

- **Practical demonstration of the superiority of the semantic web for research**

- **Portal for family history research that could be easily expanded:**
  - Maps
  - Look-ups
  - Helps
  - Research training
  - Other countries and states

QUESTIONS?