

Towards Content-Rich Digital Genealogy with Model-Oriented Approaches: Issues and Standards

Robert B. ALLEN ロバート アレン
Research Center for Knowledge Communities
University of Tsukuba
Tsukuba, Japan

rba@boballen.info

Rich Semantic Descriptions via Modeling

- As we see from the conference, there are a great many facets of interest in digitized newspapers. Can we develop a general framework to support organization, context, and access?
- Rather than developing indexes for specific terms, perhaps we can model complex entities and the what they are doing.
- Modeling would show the network of relationships and could be dynamic to include states, behaviors, and processes.
- Models could be developed for many types of entities such as people, families, multi-families, communities, and cities.
- Here, we focus on communities and the people in those communities since those match the historical newspapers.

THE NORFOLK WEEKLY NEWS.
NORFOLK, NEBRASKA, THURSDAY, MAY 10, 1900.

POPULISTS IN SESSION

Convention Begins Its Work at Sioux Falls.

ISSUE IS THE VICE PRESIDENCY

Delegates differ as to best plan for nominating candidates to run with Bryan, nominate minor and Governor Lee address the delegates.

STOCK FALLS, S. D., May 10.—The national convention of the People's Party began business at 10 p. m. yesterday in the big tent wagon on the top of the hill on the western side of the city of Sioux Falls. There have been larger crowds in other places upon national conventions and possibly there have been questions upon which more enthusiasm has been manifested, but there have been few similar events which have been marked by more evident sincerity of purpose or more pronounced doctrine of behavior. The first text was assigned to full dress array for the reception and entertainment of its guests and the structure proved in every way equal to the service required of it.

There were probably 800 delegates in their seats and surrounding them was a fringe of alternates and visitors numbering from 900 to 1000. Seats had been provided for a much larger number, but the attendance did not appear to be as good as it may be. It is interestingly remarked that what was lost in attendance was compensated for in enthusiasm. From the arrival of the Minnesota delegation bearing Bryan and Tamm banners, there were outbursts of applause at every convenient opening and upon the most pronounced.

Senators Butler, Allen and Hatfield received liberal greeting, while Senator Pettigrew was overwhelmed by the cordiality of the welcome accorded him. There was also generous applause when the name of Mr. Bryan was mentioned.

BOERS CHECK HUTTON

His Command Has Narrow Escape From Being Cut Off.

MANY HIT BY BUSHBER SHELLS.

British forces in Hutton's camp had been driven back to the river by the Boers. Hutton's command had a narrow escape from being cut off.

SMILERS, May 9.—Yesterday General Hutton attempted to seize a Boer conveyance attempting to leave the river. He advanced to the river with mounted infantry, including the Canadians. A long line of wagons was plainly visible. The Boers stopped their retreat and opened fire with 10 guns. They seemed in force and Hutton's retreat was checked. Heavily armed Boers were seen crossing the dry bed of the river. They stood to the right and began to rifle the West Australians, while many of the British were hit by their shells. The position became serious. The Boers sent up reinforcements, placing the Australians in danger of being cut off. Thereupon General Hutton ordered a retreat to Wolgastgen, where he had left his own conveyance. A portion of the forces advanced and engaged the Boers to cover his retreat, which was successfully accomplished. The Boers continued to shell the retreating troops, but showed no desire to press the attack further. Meanwhile the Boer train steamed away northward, blowing up the culverts as they went.

DAYS BEGINS BACK.

LOANERS, May 10.—General Hutton's mounted infantry brigade, including the Canadians, with a part of General Dewar's cavalry, crossed the Great river Tuesday and began to work its way cautiously along the railway northward in the wake of the retreating Boers. About 3,000 horsemen were probably

KNEELING NOT AN EXAGGERATION

Methodist conference, taken from a newspaper, May 10. Kneeling as in the attitude of prayer, has ceased to be a practice in the Methodist church. The book of discipline recognizes the kneeling of the knee, but the city churches have gradually been drifting into the more conventional kneeling of the head. The officers of Wesleyan method for the most part only in the country. Delegate Tynhill of Nebraska is a believer in the former kneeling form, and yesterday sought to correct the tendency of the latter day Methodists in departing from the ancient custom by introducing a resolution in the general conference emphasizing the direction in the discipline. The conference refused to declare "we are as much as ever in favor of kneeling" the kneeling attitude.

The effective argument was presented by a city minister, who said: "We would have to rebuild our churches. Five cents is too contracted to permit of kneeling."

The Nebraska preacher was to urge the adoption of his views, but the conference adjourned for a vote, and the more easy kneeling of the body toward the lack of the pose was adopted.

TRUCKS FOR TRAMPS YESTERDAY.

DEER MOVING, May 10.—The federal grand jury of the southern district of Iowa in session here, is of the opinion that it is a crime under the federal statute for a tramping stevedore to ride on a passenger train carrying mail. Henry H. Miller and C. H. Cherry of Lucas county were yesterday indicted for obstructing the United States mail. The indictment states that they loaded on the tracks of a car on a Chicago, Hastings, and Quincy mail train and that it was necessary to stop the train to get them off.

ST. LOUIS, May 10.—The Iowa Republican state convention meets here today. The convention will name 22 district delegates and four delegates at large to the Chicago-National convention.

THESE FOUR HATS ARE LATE PARIS PATTERNS.



See copies of them ...at... **INSKEEP'S.**

ST. LOUIS' BIG CAR STRIKE.

Every line in the city except the In Transit.

ST. LOUIS, May 10.—The second day of the great street railway strike was as quiet and uneventful as the first was turbulent and stormy. The St. Louis

BARNES & TYLER.

J. J. Barnes & S. O. Tyler, Attorneys at Law.

Norfolk, Nebraska

Start with Entity-based Descriptions (e.g., People)

- There are several emerging but limited standards for describing people:

- FOAF – Friend of a Friend
- GEDCOM – GENEALOGICAL DATA COMMUNICATION

- People Australia <http://peopleaustralia.anu.edu.au/>

Awards

Birth town

Birth state

Birth country

Cause of death

Death town

Death state

Death country

Educational institution (higher)

Educational institution (school)

Ethnicity

Events

Gender, Groups

Military service

Occupations

Organizations

Places

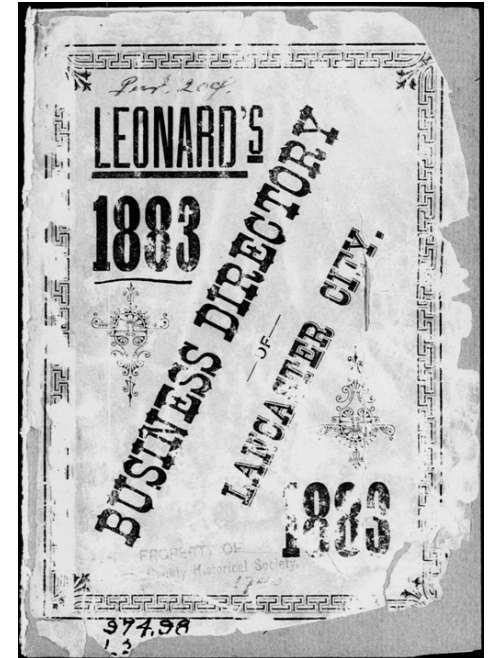
Properties

Religion

Workplaces

- “Interactive directory” for Norfolk, Nebraska

Allen, R.B., Toward an Interactive Directory for Norfolk, Nebraska: 1899-1900, IFLA Newspaper and Genealogy Section Meeting, Singapore, Aug 2013. arXiv:1308.5395



Models that include Behaviors

- We need models which describe behavior, situations, and contexts of entities.
- We have explored using object-oriented models implemented with Java for descriptions of people and communities.
- Object-oriented models are useful because they have:
 - Classes (objects) model entities with attributes as a package
 - Behaviors are implemented with methods which are part of the classes
 - Java encourages simple approaches to abstraction and instantiation
- We can use the models for simple descriptions or we can develop executable models which unfold as they are run.

Using FrameNet Frames to Structure Java Classes for Modeling Texts

- We have lots of rich text descriptions from cultural descriptions. Could we use that? After all, the text descriptions are representations.
- One approach to modeling would be to model the text with FrameNet (<https://framenet.icsi.berkeley.edu/fndrupal/>)
 - Essential concepts in natural language described with frames. These frames are related to semantic roles.
 - FrameNet is based on cognitive principles but we use it as a language resource for modeling.
 - We are particularly interested in verb frames because they describe transitions in attributes.
 - About 700 verb frames.
 - Example: Frame for “Release”

A Captor ends the captivity or inhibition of the motion of a Theme from the Location_of_confinement. The release is in accord with the plans of the Captor.

Example Text from Wikipedia

“Early history (1750-1820) of Minneapolis, Minnesota”

- Used textbook or Wikipedia-level texts
 - These are relatively straight-forward, with simple past tense
 - By comparison, primary sources have many difficulties. Full of slang, complex constructions, un-grammatical, and often incorrect statements.
 - Some massaging is still required
- Includes a passage about “releasing” captives.
- Verb frames can be modeled via methods in Java class.
- So far we have a “proof of concept” but:
 - Limitations of frames
 - Not always a perfect match
 - Limitations of Java
 - Does not support multiple inheritance

Building Comprehensive Person Models and Community Models

- While Frames focus on text descriptions, additional information about the community may be more readily available from other sources (census, diaries, oral histories).
- By combining sources of information at different levels, we can develop “community models” and “person models”. Many communities are relatively closed systems. This makes them more tractable than indexing, say, cities.
- Community models must also include Person models
 - Explanations based on theories of mental events and human motivation.
 - Interact with community models.
 - Communities structure human a lot of behavior. This structure help us to describe a person’s behavior
 - Can show internal dynamics such as;
 - Time course of emotion
 - Maturation across years
- These models need complex structures for entities and processes.
 - Classification (inheritance) hierarchies
 - Partonomies
 - Grouping like-objects
 - Hierarchy of parts
 - Systems and functionality

Discourse Models to Support Presentations and Interaction

- In addition to the content model, we want ways to support interaction and presentations.
- Discourse overlay
 - Explanations
 - Tutorial-like overviews of people and communities
 - Argumentation
 - Structured comparison of generalizations and of evidence
- Supporting history and genealogy researcher
 - Interactive interfaces for working with community histories
- Managing the content model
- Visualizations, Narratives, and Games
 - Interactive historical re-enactors, games, and cyber-dramas

Status and Standards

- We have:
 - Shown first steps to developing structured models of communities with rich semantics
 - Shown it is possible to model natural language texts with FrameNet. However, there are often gaps and we to develop descriptions that cross many levels.
- Standards:
 - Need to develop additional structures for:
 - Complex entities such as people and organizations
 - FrameNet frames generally works well but they need to be extended.
 - Several aspects of incorporating natural language are not yet well explored.
 - Need community involvement in setting standards

Issues and Future Directions

- Broader effort to develop model-oriented information organization
 - Extend model-oriented information organization to museum objects and informatics
 - The benchmark for semantics is human cognition. To what extent should our models be based in cognitive modeling?
 - Modeling nuances of natural language such as representing the relationship of clustered terms like: Baker – Bakery - Baking
- Extend community models to:
 - Multi-family genealogies
 - Networks of communities
 - Modeling cities, nations

Towards Content-Rich Digital Genealogy with Model-Oriented Approaches: Issues and Standards

Robert B. ALLEN ロバート アレン
Research Center for Knowledge Communities
University of Tsukuba
Tsukuba, Japan

As of March 1, 2014
Robert B. ALLEN 로버트 알렌
Department of Library and Information Science
Yonsei University
Seoul, Korea

rba@boballen.info