Everyone Has A Story: An Argument for Full Indexing of Records

Bill Barrett

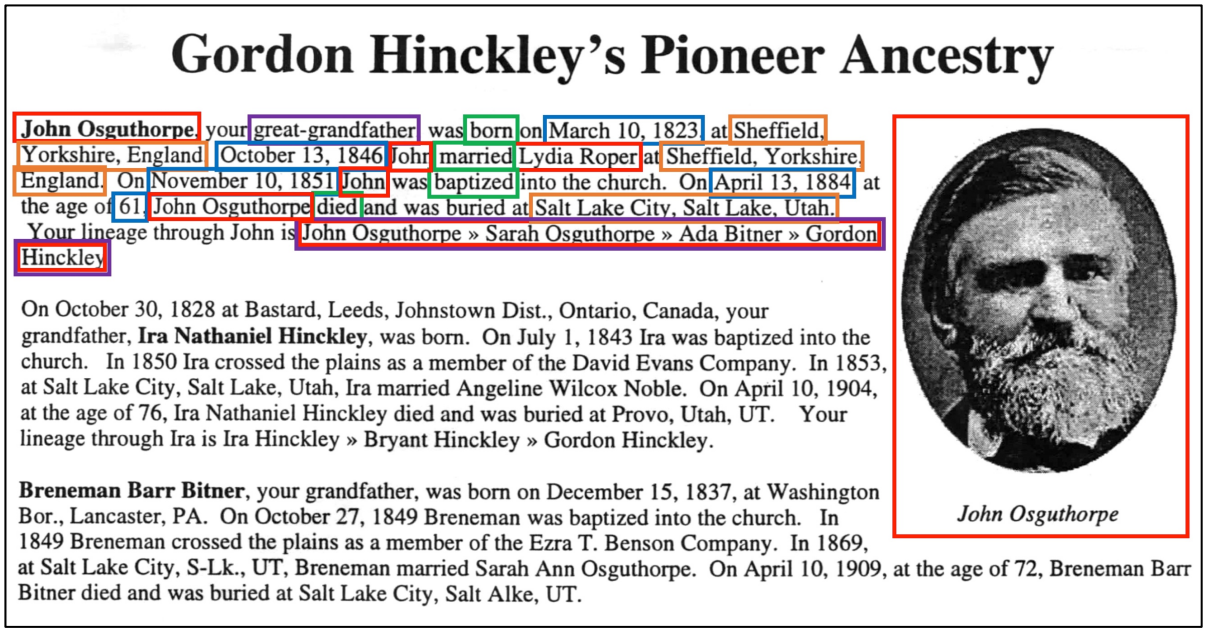
Introduction

In 1997, The Department of Computer Science was invited to participate in a Sesquicentennial Celebration to commemorate the 150th anniversary of the arrival of pioneers into the Salt Lake Valley. Our part was to make it so that people who came to the celebration could find a story and maybe a photo of a pioneer ancestor, and leave with a print-out in their hands.

To approach this, we recruited Evan Ivie, one of our long-term and beloved faculty along with a group of students who we sequestered in a room for six months with an Informix Engine, a Folio image database, crossing-the-ocean and crossing-the-plains indexes from Milt Backman, and the (then) Ancestral File on loan from the Church for the event.

As the date drew near, we found that we were able to find a story and/or photo for only about one out of three test cases. We felt that for it to be a successful event, we would need a much greater hit rate. At that point, one of the students had the bright idea to create stories of pioneer ancestors by narrating between the data (names, dates, places, events). We also decided that if we didn't have a photo of the ancestor, we could at least print out a period image. This made it possible for the large majority of visitors to come away with a story in their hands (see Figure 1).

Stories have the power to move people, turn hearts and create connections. And *everyone has a story*, even if they didn't keep a journal.



Gordon Hinckley's Pioneer Ancestry

John Osguthorpe, your great-grandfather, was born on March 10, 1823, at Sheffield, Yorkshire, England. October 13, 1846 John married Lydia Roper at Sheffield, Yorkshire, England. On November 10, 1851 John was baptized into the church. On April 13, 1884 at the age of 61 John Osguthorpe died and was buried at Salt Lake City, Salt Lake, Utah. Your lineage through John is John Osguthorpe » Sarah Osguthorpe » Ada Bitner » Gordon Hinckley.

John Osguthorpe

On October 30, 1828 at Bastard, Leeds, Johnstown Dist., Ontario, Canada, your grandfather, **Ira Nathaniel Hinckley**, was born. On July 1, 1843 Ira was baptized into the church. In 1850 Ira crossed the plains as a member of the David Evans Company. In 1853, at Salt Lake City, Salt Lake, Utah, Ira married Angeline Wilcox Noble. On April 10, 1904, at the age of 76, Ira Nathaniel Hinckley died and was buried at Provo, Utah, UT. Your lineage through Ira is Ira Hinckley » Bryant Hinckley » Gordon Hinckley.

Breneman Barr Bitner, your grandfather, was born on December 15, 1837, at Washington Bor., Lancaster, PA. On October 27, 1849 Breneman was baptized into the church. In 1849 Breneman crossed the plains as a member of the Ezra T. Benson Company. In 1869, at Salt Lake City, S-Lk., UT, Breneman married Sarah Ann Osguthorpe. On April 10, 1909, at the age of 72, Breneman Barr Bitner died and was buried at Salt Lake City, Salt Alke, UT.

Figure 1. Story of President Gordon B. Hinckley's great grandfather John Osguthorpe, created automatically by narrating between names (red), dates (blue), places (orange), events (green) and relationships (purple).

Story Telling Using Deep Learning
Our use of textual glue to narrate between pioneer names, dates, places and events 25 years ago is primitive compared to today's Deep Learning Networks that are trained on copious amounts of data to formulate fluid narratives by predicting the next word that comes in a text sequence. For example, GPTs (Generative Pre-trained Transformers, such as GPT-3.5) use a stimulus-response perception system with large language models (LLMs) and reinforcement learning aided by human feedback to create concise textual summarizations and narratives for interactive conversational use. Such networks could also be used to create stories of individuals when trained on data, stories and photos such as those found in memories of FamilySearch.

In recent years, significant advances have been made in automated extraction of family history information from textual sources such as Family Histories, Books, Obituaries etc. into semantically rich ontologies and entity relationship models [1-2]. Further strides have demonstrated how machine learning and computer vision can be used to automatically link historical records and leverage related information from photos [3-4]. As we contemplate the technologies used to extract data from a growing volume of family history information, we might want to consider how we can exploit that data to reverse-engineer a story of an individual's life. As a concrete example, can a deep learning network used to extract information from obituaries [2] also be used to recreate the obituary and/or a life sketch?

Full Indexing for Richer Story Telling and Search
A major source of family history information is Census Records which have mostly been used to extract only names, ages, dates and places, creating only a partial index of the record. However, there is much more that can be extracted such as education, profession or occupation, address, names and nationality of parents and neighbors, where they lived and where they and their parents came from, all of which can be used to create a full(er) index and a richer story of the person's life.

What is true of Census Records is likewise true of Birth, Marriage and Death Records, hand-written letters or the like, which contain information on parentage, country of origin, cause of death with associated names and places, with activities and events and historical social networks [5], all of which enrich the underlying story of each individual. As named entities are added to the persona, the surface area increases, providing additional input nodes to a GPT-style neural network to create a fluid and compelling narrative and life story, rather like a rolling stone gathering more mass and momentum until we not only know, but begin to actually understand the life story of this person. Finally, those same kinds of nodes can also be used in Deep Learning networks to broaden the search for additional instances and information about each individual and thereby continue to expand the narrative.

1. Embley, D. "Linking Families with Enriched Ontologies," FHTW 2020.
2. Schone, P and Nielson, H. "'Robokeying' of Born-Paper Obituaries," FHTW 2017.
3. Price, J. "Combining Family History and Machine Learning to Link Historical records," FHTW 2020.
4. Kennard, D. "Faces, Places, and Kin: Applying Computer Vision to Family History," FHTW 2020.
5. Kennard, D. Kent, A. and Barrett, W. "Linking the Past: Discovering Historical Social Networks from Documents and Linking to a Genealogical Database" HIP'11.