GOALS, DIRECTIONS, POSSIBILITIES, THOUGHTS

Fe6 Project

A: Quarterly goals (2018-1)

- FS-ingest: enable others (missionaries, students, …) to process a book

- CMS: complete book import, OntoES configuration and testing, and non-COMET lines

- Pipeline: push complex annotations through 2&3; check and adjust 4&5&6 components

- COMET: enhance (if needed) for initial production work

- ConstraintEnforcer: papers/demos for FHTW, RootsTech, ER

- FROntIER:

- GreenDDA/GreenML: test relationships

- GreenFIE: submit paper

- GreenQQ: set-up python; resolve record template groupings; design interface

- ListReader:

- OntoES: do backend coding for CMS configuration and testing

- OntoSoar:

A: Ingest into FS

- Possibilities (?):

- FamilyTree (Knox County temple submissions -- survey of patrons ??)

- Obituaries (GedcomX specs -- looks like it could work, but not encouraged)

- Scanned books (HyKSS-like search -- discussions with Ty Davies, Jon Morrey, and Jonathan

Cranford without a favorable resolution, but they’re at least thinking about it)

- LLS and CDS (initially, not hopeful, but the “soon” to be LLS Tree looks promising)

- Jon Morrey’s suggestion: run 1,000 books through the system

B: FS future projects (wish list, to continue to encourage and help realize):

- Highlighted, document-first output by search system

- HyKSS-like search engine over all content

- Patron ingest via COMET, Ddupe-like entity resolution, and semi-automatic FamilyTree update

- Green system of the whole – constant improvement while doing real work

Research

A: (me, from Thomas) follow up on TKDE paper (Dec 17 resubmission)

- Name/Place/Date-level text abstraction (w/ DL & SL) and ? for GreenQQ w/ GN

- high-level nested pattern discovery in a second pass with record-level text abstraction

A: (Liddle) GreenFIE paper (DKE)

- forms to mini ontology and eventual integration into backbone ontology

- extraction by layout for forms & diagrams (especially for Jaipu)

A: (Lonsdale)

- GreenDDA (w/ GN)

- GreenML & GreenOTS

- OntoSoar: preprocess with named-entity recognizer (does this give a significant boost?)

- long-term directions: anaphor resolution, XNL parser; declarative rule specification; rule learning

A: (Nagy) journal paper

- extend GreenQQ to work with our set of data

- create a COMET-based interface for GreenQQ

A: (Woodfield) FHTW’18 paper, RootsTech demo, ER’18 paper

- quality assessment (of merge, of an individual’s genealogy, of FamilyTree)

- Bayesian reasoning for assessing quality

C: Grand Challenges

- “Green Interaction” (systems that improve while being used for real-world applications)

- “Teaching Computers to Read” (cognitive computing grand challenge): EMISA paper

- “Web of Knowledge” (WoK vision with FamilySearch as an example)