GOALS, DIRECTIONS, POSSIBILITIES, THOUGHTS

A: Family Reconstitution & Patron Indexing & Scan to Submit with Implicit Indexing (research)

- Family Reconstitution (work toward having a prototype that adds genealogies to LLS)

- code auto-tree-generation pipeline (much done; finish:SW&DE)

- write CoMoDiH Wkshp paper (done); write Journal paper

- resolve issues regarding actual ingest of generated gedcomx into LLS (DE&JM)

- Patron Indexing (work toward enabling others to process a book)

- complete COMET revision (done; tweak:SL)

- build GreenQQ patron indexer interface (Web:SL, csv:DE, python:GN&DE&DL)

- move testing and PRF & summary reporting pipeline to dithers (DE&SL)

- find and download a collection of documents for patron field testing (DE&JM)

- Scan to Submit (work toward having a prototype for FamilySearch evaluation)

- create Tree-Ready COMET interface (done; tweak:SL)

- define input file specification and code its transformation from COMET-json (DE&JM)

- integrate the generation of COMET-json into FamilySearch keyword search (DE&JM)

- implicit indexing (work toward creating a “Green” feedback loop for implicit indexing)

- implement initial name selection (form fill & capture of implicit indexing info)

- brainstorm issues (use GreenQQ & GreenML, update of search repository, …)

B: Family Reconstitution & Patron Indexing & Scan to Submit with Indirect Indexing (tech transfer)

- Image Capture & Preprocessing: eventually multilingual OCR and HWR (multilingual COMET)

- User Search: hybrid keyword and semantic search (HyKSS)

- Form Fill: automatic initialization; document-specific form fill; user correction (pipeline & COMET)

- Error Correction: adjustment to search repository (rerun of downstream pipeline)

- Tree Import: check constraints; check for duplicates; resolve duplicates; post information

(DeepChecker, standardization & inference, D-Dup-like interface, source documentation)

C: Miscellaneous

- tools: OntoES, GreenFIE, ListReader + (for implicit indexing) Green-ML & Green-NLP

- papers: CoMoDiH+ for journal; if invited: GreenFIE (w/ Tae Woo) & ListReader (w/ Thomas)

- grand challenges (to which we may be able to contribute):

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

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

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