Corpora Development and Publication
Stephanie Strassel
Andrew W. Cole
University of Pennsylvania, Linguistic Data Consortium
strassel@ldc.upenn.edu
andrew.cole@ldc.upenn.edu
www.ldc.upenn.edu

NOTES
1. Each publication, release, updated version, etc., undergoes both Planning
   and Publications cycle for each event.
2. Quality Control is a continuum and depends on the type of data and level of
   publication. For example, eCorpora, which are released quickly and with
   minimal review have a much different level of Quality Control than a general
   membership publication, which has more comprehensive review.
3. Accepted standards are helpful in understanding the types and Quality
   levels required for a publication.

Planning

Sponsor

Researcher

Independent Data Source

Negotiate with Data Source

Offers Resource for Distribution

Intellectual Property Rights Allow Distribution

Release Schedule Conference/Dissertation

Status of Data and Documentation

Agreement for Resource Creation

Identify Resources

Institutional Review Board (IRB) Approval

Publication

Researchers

Documentation

Data

Agreement between Documentation, Data, and Applicable Standards

Quality Review

Intellectual Property Review

Scheduling

Standards

Project Managers

Media Produced

Licensed, Paid - then Shipped to non-Member Researchers

Licensed then Shipped to Member Researchers

Announced