

Aristar-Dry (1998) to see how these topics have been part of linguistic discussion for some time.

3 <http://www.sil.org/computing/toolbox>

4 <http://www.sil.org/computing/fieldworks>

5 <http://www.lat-mpi.eu/tools/elan>

6 See Bird and Simons (2003) for a detailed discussion of the principles underlying the creation of good linguistic data from fieldwork.

7 <http://www.language-archives.org>

8 <http://www.mpi.nl/IMDI/documents/documents.html>

9 So-called 'backslash' codes or standard field markers are a simple way to delimit fields in a text document. MDF, or Multi-Dictionary Formatter, is the system used by Toolbox to create dictionaries as styled RTF documents from a structured Toolbox file (cf. <http://www.sil.org/computing/shoebox/mdf.html>), and it provides a set of over 100 standard 'backslash' codes.

10 <http://www.stanford.edu/~popolvuh/field-checklist.htm>

11 Memory cards of various kinds are used in many portable devices like digital recorders and cameras.

12 OTG USB devices are those that can act as hosts to other USB devices and also interact with a computer; in this case, a device to which you can upload files from your recorder's SD card for storage until you can transfer them to a computer.

13 Temporarily restricting access may also be necessary if your recordings include sections with local gossip or other content that a speaker subsequently requests to have suppressed. Given the problem of 'data deluge', this necessary work can become more and more onerous as the collection grows, and is not a task that current digital tools provide much help in automating. Restricting access to all primary materials contributed to the archive until they can be reviewed and edited offers an immediate solution to issues of preservation and privacy, but it does not necessarily address the bottleneck created by this kind of review, and potentially poses difficulties for later open access. Furthermore, it could be controversial to archive language materials outside of the speaker communities involved in documentation, especially in cases where past research relationships have contributed little of appreciable benefit to the speaker communities involved, or even proven hurtful in the treatment of language materials which may have been removed from those communities. This is a matter that often requires more discussion between all partners in documentation to arrive at a solution that ensures long-term preservation while respecting the concerns of all involved.

14 A temporary solution might be to append the date, such that IMG_0086 could become 20110102IMG_0086 (for 2 January 2011).

15 <http://www.doi.org/>

16 <http://www.handle.net/>

17 <http://www.lat-mpi.eu/tools/abil/>

18 <http://www.mpi.nl/IMDI/>

19 <http://www.paradisec.org.au/downloads.html>

20 <http://saymore.palaso.org/>

21 <http://www.fieldhelper.org/>

22 ISO 8601, http://www.iso.org/iso/ue_detail?csnumber=40874

23 ISO 3166, http://www.iso.org/iso/country_codes.htm

24 ISO 639, <http://www.sil.org/iso639-3/codes.asp>

25 <http://dublincore.org/>

26 <http://www.openarchives.org>

27 http://en.wikipedia.org/wiki/Regular_expression, <http://www.regular-expressions.info/>, or <http://etext.virginia.edu/services/helpsheets/unix/regex.html>.

28 This example is from Clark (2009).

29 <http://sweet.artsrn.ualberta.ca/cdcox/cuped/>

30 <http://www.eopas.org>

31 <http://www.lat-mpi.eu/tools/annex/>

32 <http://www.lat-mpi.eu/tools/elan/>

33 <http://trans.sourceforge.net/en/presentation.php>

34 http://www.exmaralda.org/en_index.html

35 It is best to avoid the less desirable commercial transcription software options targeted to the legal and medical professions.

36 http://en.wikipedia.org/wiki/Open_science_data

37 <http://www.rnld.org>

38 <http://paradisec.org.au/blog>

39 <http://www.hrelp.org/languages/resources/orel/tech.html>