DDIR: handling social research data standard with R

Yasuto NAKANO

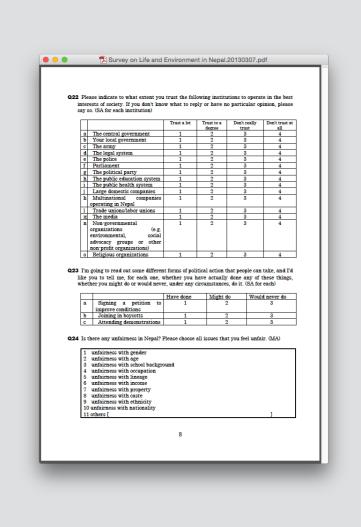


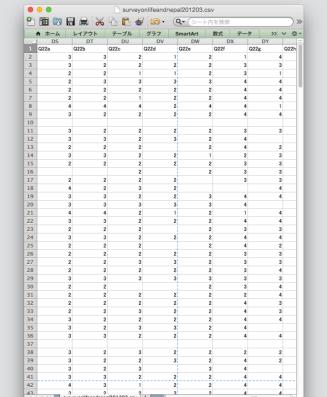


PURPOSE

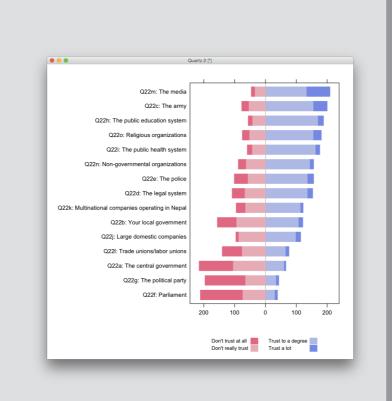
- to propose an environment for social research and its analysispackage DDIR
 - ▶ import & export meta information of research data
 - providing efficient & integrated environment for researcher

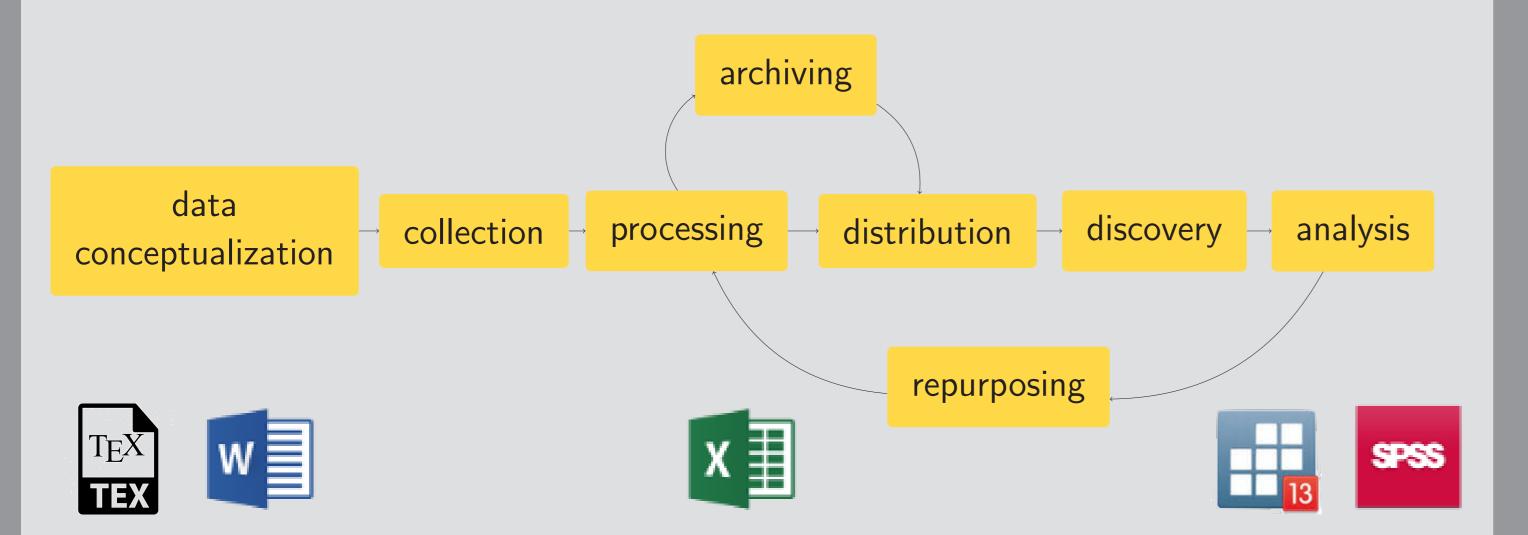
CONFUSIONS in Data Life Cycle of SOCIAL RESEARCH











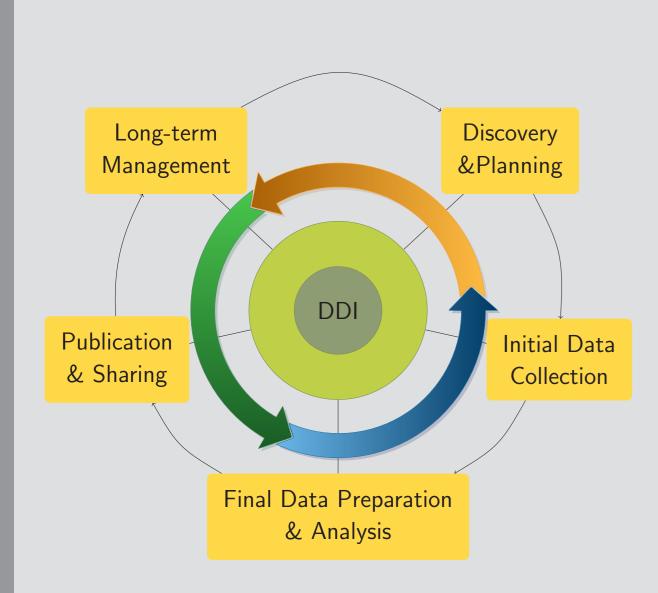
- ► There are several phases to create social research data
 - □ data conceptualization, collection, processing, distribution, discovery, analysis, repurposing, and archiving
- Different materials/tools for different phases
 - ▶ text files, word files
 - ▶ excel files
- ▷ SPSS files etc.

HOW TO AVOID CONFUSION

- consistent data source/tool throughout data life cycle
 - ▷ 'DDI' could be a solution

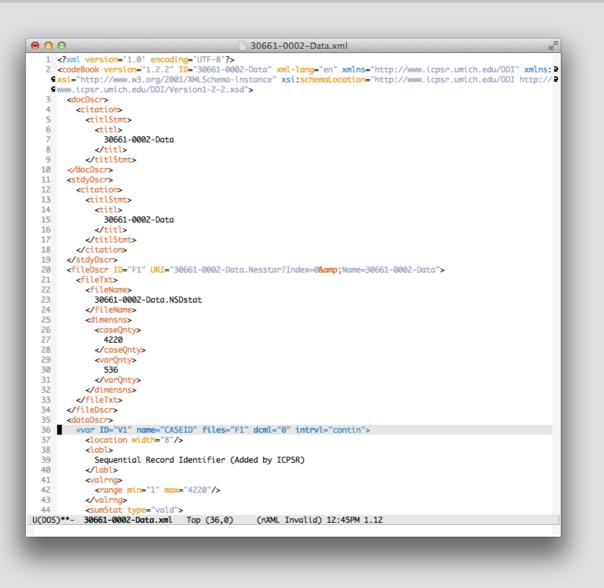


- http://www.ddialliance.org/
- DDI is an international standard protocol for describing data from the social, behavioral, and economic sciences.
- ▶ There are several international research projects which use this protocol as a standard format.



- ▶ DDI is expressed in XML.
- ▶ DDI supports all phases of social research with one simple, standardized and integrated DDI files (a meta data XML file and a data CSV file).
- ➤ You can retrieve/add informations from/to DDI files troughout entire data life cycle.
 - > stocking ideas of questions and theoretical concepts behind them
- memorizing research design
- publishing questionnaire sheet
- > storing codebook including variable informations
- publishing summary report
- ▶ R could be a tool for analysis and publication with DDI

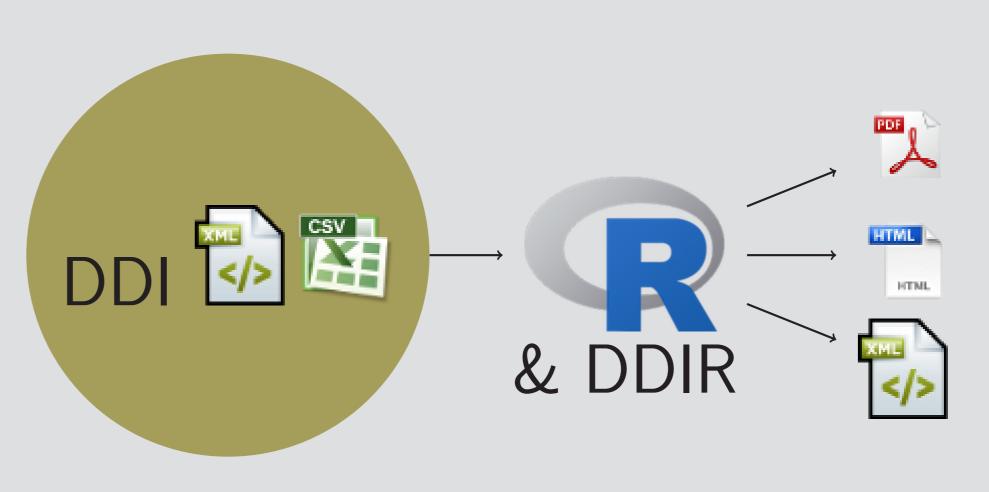
DDI xml file



- retrieve informations
 required to data analysis
 variable names, values
 lavels, data formats, etc.
- restore infromations
 resulting from analysis
 frequencies, summary
 statistics, etc.

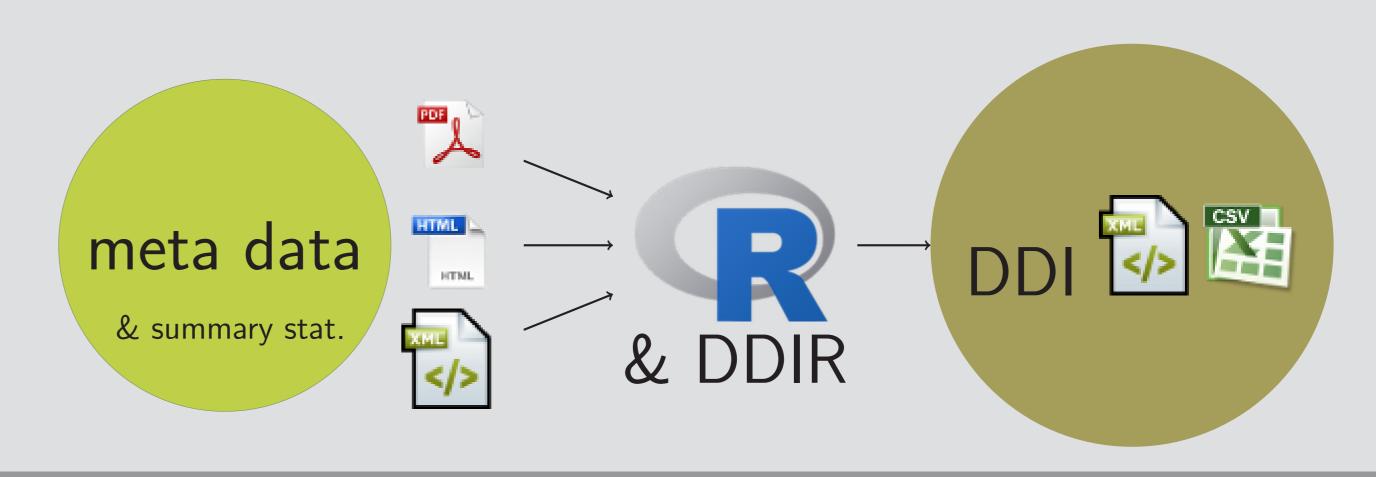
CONCEPT OF the package DDIR

- ► We are developing a new package DDIR that import/export a DDI xml file from/to R.
- imorting information
 - ▶ Variable names, value labels, and questionnaire sentences etc. are imported from a DDI xml file



exporting information

> meta information and summary statistics are exported as a DDI xml file



FUNCTIONS OF the package DDIR

- read.codebook(ddi.xml.file,csv.data.file)
 - > retrieving informations from a ddi.xml file
 - combining these informations with data
 - returning a data frame object
- ▶ the object with variables, variable labels, value labels, question text in the questionnaire, etc.
- show.question(ddiobject\$varname)
- show.varlabel(ddiobject\$varname)
- show.valuelabel(ddiobject\$varname)
- write.codebook(filename,docDscr,stdyDscr,fileDscr,dataDscr)
 - > exporting metadata and related informations in DDI xml format
 - docDscr : information on the meta file itself
 - stdyDscr: information on the researchfileDscr: information on related files
 - dataDscr: information on the data

REFERENCES

- ► "An Integrated Environment for Social Research Analysis," *R user conference 2014* (Los Angels, USA), 1-3/Jul/2014.
- ► "DDIR and dlcm: integrated environment for social research data analysis," *R user conference 2015* (Aalborg, Denmark), 01/Jul/2015.
- "Integrated Environment for Social Research Data Analysis: DDIR," IASSIST 2016 (Bergen, Norway), 02/Jun/2016.
- ► "DDIR: R package for handling DDI file," *EDDI16* 8th Annual European DDI User Conference (Cologne, Germany), 2016.12.7.