NEPS
National Educational Panel Study

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Referencing NEPS data with DOIs

Knut Wenzig
knut.wenzig@uni-bamberg.de
NEPS Data Center

Metadata and Persistent Identifiers
for Social and Economic Data

7–8 May 2012, Berlin

Three questions

1. What is the structure of the DOI?

2. What kind of metadata are expected and in which form?

3. What should the landing page look like?

NEPS data

Data should be known, if DOI structure should depend on data.

- 6 panels and 2 additional studies
- Each main release integrates data from last main release.

- 3 different versions according to degree of anonymization resp. data access platform
- 2 languages (German/English)
- 2 software platforms (Stata, SPSS)

Release Strategy and Versions

Three-digit system to identify versions

- 1st digit: main release

  (each wave)

- 2nd digit: major updates

  (affecting data structure, necessary to update syntax files)

- 3rd digit: minor update

  (affects only content of cells, bugfixes)

Structure of DOI – granularity choice

Decision: deliver systematic DOI

- Included, seperated by colons:
  - "NEPS" – our name in the DOI
  - Key for cohort (starting cohort 1–6, TH, G89)
  - Version
- Not included:
  - Different flavours
    (software platforms, languages, access platforms)
  - Due to *Principle of Functional Granularity*: "it should be possible to identify an entity whenever it needs to be distinguished." Norman Paskin (2003)
    http://www.dlib.org/dlib/january03/paskin/01paskin.html

The result

doi:10.5157/NEPS:SC6:1.0.0

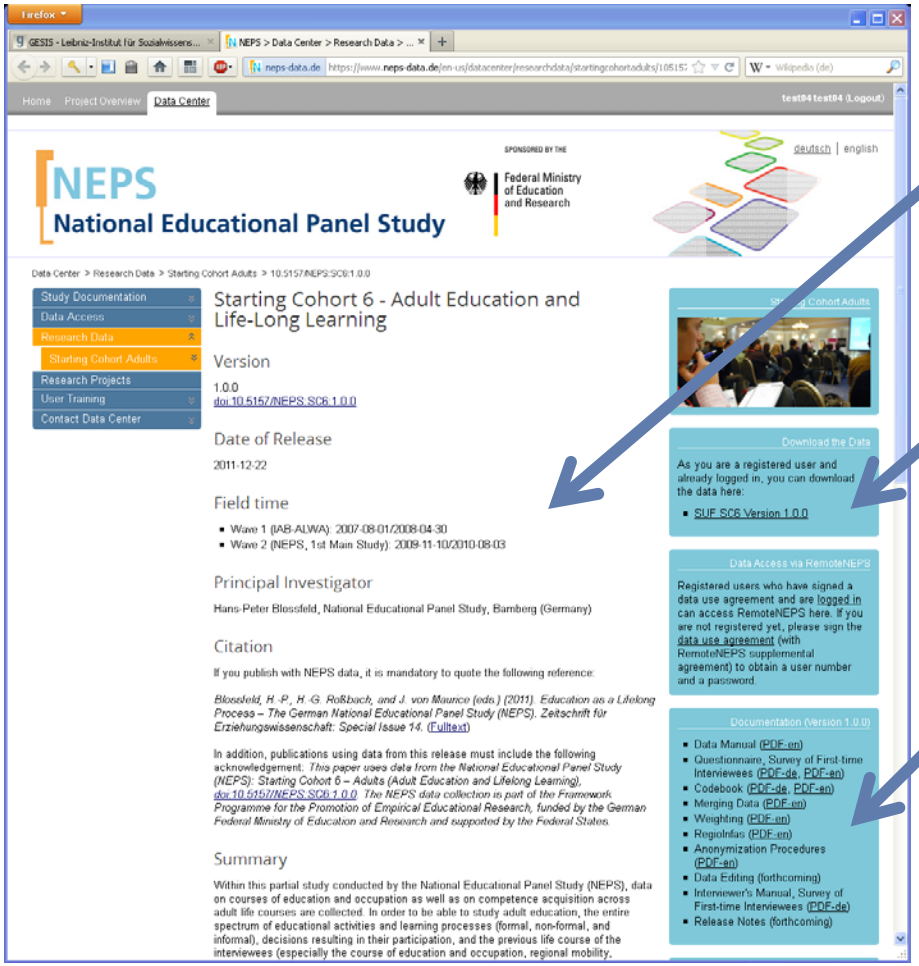
Question 2: What kind of metadata is expected and in which form?

Metadata via XML

- Three different specifications:
  - dara's metadata scheme – short version (PDF) "Metadatenschema 1.0 (Kurzversion)"
  - dara's metadata scheme - long version (XML/XSD)
  - DataCite's Metadata Schema v 2.2 (XML)
- Some inconsistencies and contradictions

Question 3: What should the landing page look like?

Arrival at NEPS data via DOI



1. information from DOI's metadata
2. access to data with the next click
3. compressed documentation for related SUF