DDI Version 3 and Instrument Documentation

Karl Dinkelmann
Survey Research Center
Institute for Social Research
University of Michigan

IASSIST
Ann Arbor, MI
May 23 - 26, 2006
Acknowledgments & Participation

- **DDI Instrument Documentation Workgroup**
  - Cavan Capps, Census Bureau
  - Dan Smith, University of Minnesota
  - Ed Ross, OpenSurvey Project; Tripe-S
  - Gina-Qian Cheung\(^1\), ISR, SRC, University of Michigan
  - Karl Dinkelmann\(^1\), ISR, SRC, University of Michigan
  - Peter Sparks, ISR, SRC, University of Michigan
  - Jeremy Iverson, University of Wisconsin-Madison
  - John Ladds\(^1\), Statistics Canada
  - Karsten Stephen
  - Lon Hofmann, Statistics Netherlands
  - Mari Kleemola\(^1\), Finnish Social Science Data Archive
  - Peter Joftis\(^1\), ICPSR
  - Tom Piazza\(^{1,2}\), University of California, Berkeley
  - **WSG**: SRG\(^1\), Reto Hardon\(^1\), and many more behind the scenes

\(^1\)At IASSIST

\(^2\)WG Facilitator
Overview

- Background
- Motivation
- Proposal
- DDI V.3 instance
- Instrument Documentation (ID) home in DDI V.3
- DDI:ID model
- Notable expansions
- Issues / omissions
- Next steps
DDI-ID: Background

- DDI-ID working group Feb 2005 ~ Present
- Weekly conference calls
- Have day jobs...
- Submitted DDI:ID proposal to SRG (10/2005)
- Currently reviewing DDI V.3 schema
DDI-ID: Motivation (summary)

- Extending DDI to encompass elements describing the CAI process.
- Machine-readable, Machine-processable, & Machine-actionable
- Develop a model allowing the flexibility to document as little or as much detail as one prefers.
- Added complexity means automating the process is key for complex instruments
• Addressed the following types of information:
  – Questions in a survey instrument
  – Structure, flow, and logic of an instrument
  – Relationship of instrument data to analysis data
  – Reasoning behind asking specific questions
  – Other metadata such as the methodology and process used to administer the instrument.
  – Derived/Constructed variables
Data Collection Module Concepts

• Methodology
  – Data collection
  – Time methods
  – Sampling procedures
  – Deviations from the sample design

• Collection Events
  – Collector
  – Data source
  – Frequency of data collection
  – Mode of collection
  – Collection situation
  – Actions to minimize data loss

• Instrument
  – Questions (structured)
  – Instrument (structured)

• Processing
  – Control operations
  – Cleaning operations
  – Weighting
  – Data appraisal
  – Coding
Notable Expansions: Question Sequence
Notable Expansions: Universe
Notable Expansions <ControlConstruct>
Notable Expansions <ControlConstruct> <IfThenElse> (1)
Notable Expansions <ControlConstruct> <IfThenElse> (2)
DDI:ID Issues, etc?

• Recursion & the Control Construct?
• What have we left out?
  – Preload variables
  – Alien Procedures (calls to external routines)
  – Alien Routers (calls to external programs)
  – Associated paradata
  – Response ordering
  – Consistency checks
  – Display issues
    • web surveys (example…)

• Others?
Experimental Designs in Web Surveys

**condition 1**

<table>
<thead>
<tr>
<th>Maßnahmen zur Steigerung der Sicherheit im Internet</th>
<th>sehr wichtig</th>
<th>sehr unwichtig</th>
<th>weiß nicht</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mails verschlüsseln</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Festplattendaten verschlüsseln</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Javascript deaktivieren</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Popups verhindern</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cookies verbieten</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verschlüsselte Webseiten (https)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viele verschiedene Passwörter nutzen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwierig zu knackende Passwörter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-Virenprogramm nutzen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firewall nutzen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eigene IP-Adresse verschleim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auslesen der Bildschirmauflösung verhindern</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angaben zum Browser verfälschen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name des Betriebssystems zurückhalten</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verwendete Suchmaschine verheimlichen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vorhergehende Seite verheimlichen</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*data documentation initiative*
*a project of the social science community*
Next Steps...

- Document schema
- Create examples
- Proof of concept
- Where do all of you come?
  - Public release (of draft)...
  - Offer feedback
  - Your opinion will influence the final version of DDI V.3!
- Expand Data Collection module to include other aspects earlier in the lifecycle (new work group?)
Contact

Questions, comments, concerns please email:

ddi-id@icpsr.umich.edu

Thank-you!