Archivist & Mapper

Will Poynter • CLOSER
CLOSER:
Maximise the use, value and impact of the UK’s longitudinal studies, both at home and abroad
CLOSER Discovery

- Began beta testing December 2015
- Production launch January 2017
- Continually increasing content
The problems

- Scale
- Training
- Management
Scale
80

Years of data collection
100,000

Questions to be entered manually
Training
1154

DDI-3.2 Elements
Management
Our team

Metadata Officer

Metadata Assistants
The consortium

- Eight studies
- Five centres
310,778

Items currently in CLOSER Discovery
The solution

Archivist:

- Web application
- Easy to use
- Centralised
- Built for scale
Client

Postgres

Rails 5

Redis

Node.js

Server

Angular.js

Socket.io
Development cycle

A new version every 6 weeks

- December ‘15: Development started
- February ‘16: Version 1: Released for beta testing
- July ‘16: Version 4: Released for production use
- August ‘16: Version 5: Upgraded to Rails 5
- November ‘16: Version 7: Current version
<table>
<thead>
<tr>
<th>ID</th>
<th>Prefix</th>
<th>Study</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1172</td>
<td>hcs_2_clinic_13</td>
<td>HCS</td>
<td>Edit</td>
</tr>
<tr>
<td>1173</td>
<td>hcs_2_clinic_14</td>
<td>HCS</td>
<td>Edit</td>
</tr>
<tr>
<td>1174</td>
<td>hcs_2_clinic_15</td>
<td>HCS</td>
<td>Edit</td>
</tr>
<tr>
<td>1175</td>
<td>hcs_2_clinic_16</td>
<td>HCS</td>
<td>Edit</td>
</tr>
<tr>
<td>1176</td>
<td>hcs_birth</td>
<td>HCS</td>
<td>Edit</td>
</tr>
<tr>
<td>1177</td>
<td>mcs_03_cmq</td>
<td>MCS</td>
<td>Edit</td>
</tr>
<tr>
<td>1179</td>
<td>mcs_03_na</td>
<td>MCS</td>
<td>Edit</td>
</tr>
<tr>
<td>Code Lists</td>
<td>Response Domains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Categories</td>
<td>DateTime Answers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code Lists</td>
<td>Numeric Answers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Text Answers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question Items</td>
<td>Constructs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question Grids</td>
<td>Conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructions</td>
<td>Loops</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequences</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statements</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Question Item

Label
qi_A6

Instruction
Interviewer Instruction

Literal
Have you decided on the names for the child yet?

Response Domains

<table>
<thead>
<tr>
<th>Type</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>ResponseDomainCode</td>
<td>cs_qA6</td>
</tr>
</tbody>
</table>
DDI 3.2
Imports + Exports
Looking forward

Questionnaire capture
Documenting historical questionnaires

Mapping to variables
Identifying source questions for variables

Provenancing questions
Documenting the origin of questions
CLOSER Mapper

- Question to variable mapping
- Derived variable mapping
- Allocate topics:
  - Questions
  - Variables
- Easy user-interface
- Real time validation
<table>
<thead>
<tr>
<th>ID</th>
<th>Qc</th>
<th>Literal</th>
<th>Variables</th>
<th>Topic</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>8365</td>
<td>qc_C4</td>
<td>How many people are there (including your partner) that you can talk about personal problems?</td>
<td>pe223 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8366</td>
<td>qc_C5</td>
<td>How many people (including your partner) talk to you about their personal problems or their private feelings?</td>
<td>pe224 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8367</td>
<td>qc_C6</td>
<td>If you have to make an important decision, how many people (including your partner) are there with whom you can discuss it?</td>
<td>pe225 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8368</td>
<td>qc_C7</td>
<td>How many people are there among your family and friends from whom you could borrow 100 if you needed to?</td>
<td>pe226 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8369</td>
<td>qc_C8</td>
<td>How many of your family and friends would help you in times of trouble?</td>
<td>pe227 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8370</td>
<td>qc_C9</td>
<td>During the last month, how many times did you get together with one or more friends?</td>
<td>pe228 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8371</td>
<td>qc_C10</td>
<td>During the last month, how many times did you get together with one or more of your relatives or your partner's relatives?</td>
<td>pe229 X</td>
<td>Family and social networks</td>
<td></td>
</tr>
<tr>
<td>8372</td>
<td>qc_D1</td>
<td>Do you feel upset for no obvious reason?</td>
<td>pe240 X</td>
<td>Mental health and mental processes</td>
<td></td>
</tr>
<tr>
<td>8373</td>
<td>qc_D2</td>
<td>Do you get troubled by dizziness or shortness of breath?</td>
<td>pe241 X</td>
<td>Mental health and mental processes</td>
<td></td>
</tr>
<tr>
<td>8374</td>
<td>qc_D3</td>
<td>Have you felt as though you might faint?</td>
<td>pe242 X</td>
<td>Mental health and mental processes</td>
<td></td>
</tr>
</tbody>
</table>
Archivist 8

January 2017

- Question to variable mapping
- Derived variable mapping
- Question + variable tagging
- XML Export via API
- Bug fixes
● Longitudinal mapping
● Equivalency mapping
● Greater caching
● Automatic tagging
● More bug fixes
Thank you

Any questions?

Special thank you to

○ ESRC
○ MRC
○ JetBrains

Contact me

○ @willpoynter
○ w.poynter@ucl.ac.uk
○ github.com/wpoynter