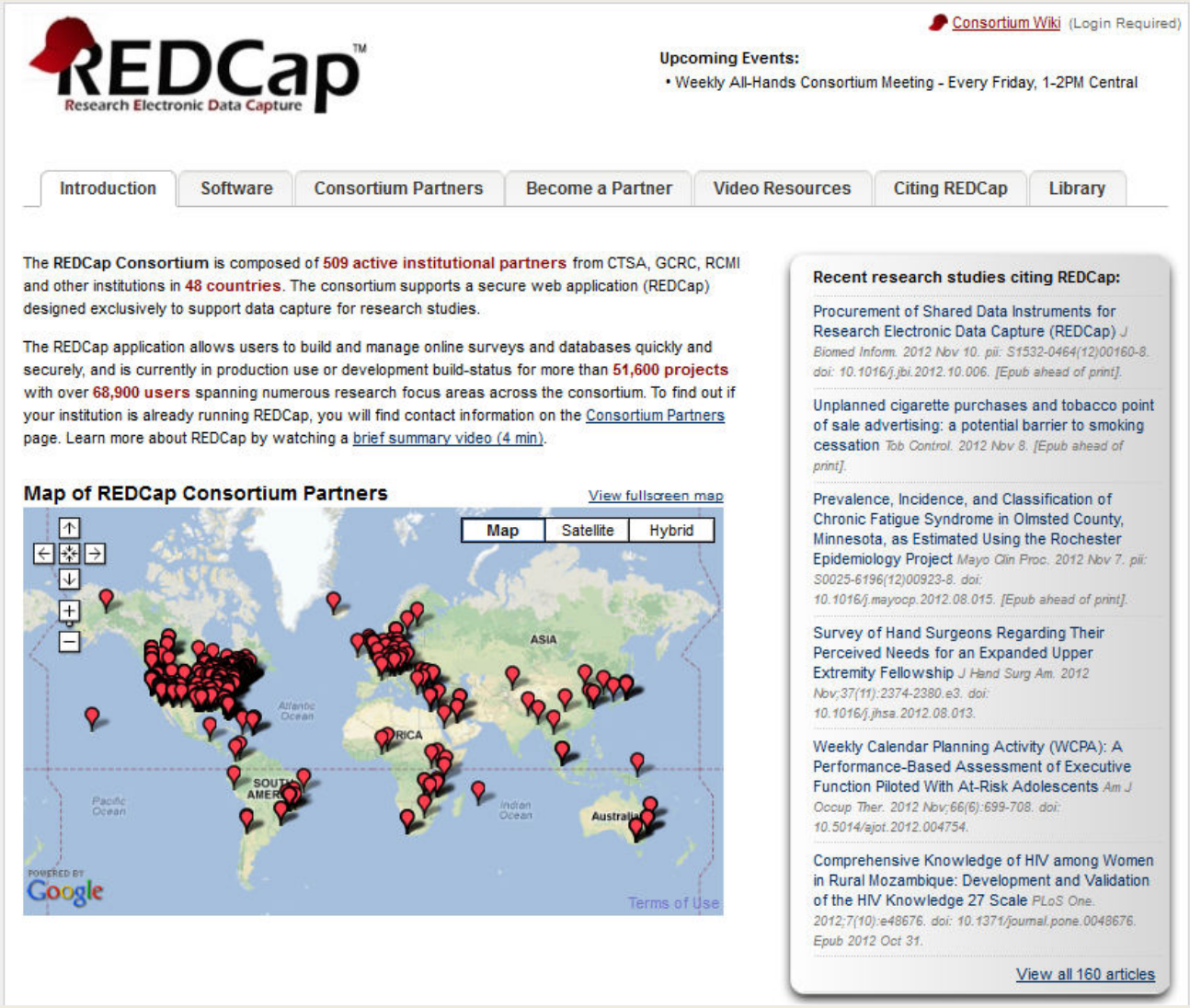
REDCap and DDI

Larry Hoyle
Institute for Policy & Social Research
University of Kansas

REDCap Consortium

http://project-redcap.org/



REDCap Citation

Consortium Wiki (Login Required)

Introduction | Software | Consortium Partners | Become a Partner | Video Resources | Citing REDCap | Library

Citing REDCap

Please cite the publication below in study manuscripts using REDCap for data collection and management. We recommend the following boilerplate language:

Study data were collected and managed using REDCap electronic data capture tools hosted at [YOUR INSTITUTION].[1] REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing: 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources.

[1]Paul A. Harris, Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, Jose G. Conde, Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support, J Biomed Inform. 2009 Apr;42(2):377-81.

Link to article: http://www.sciencedirect.com/science/article/pii/S1532046408001226

Application

Interface

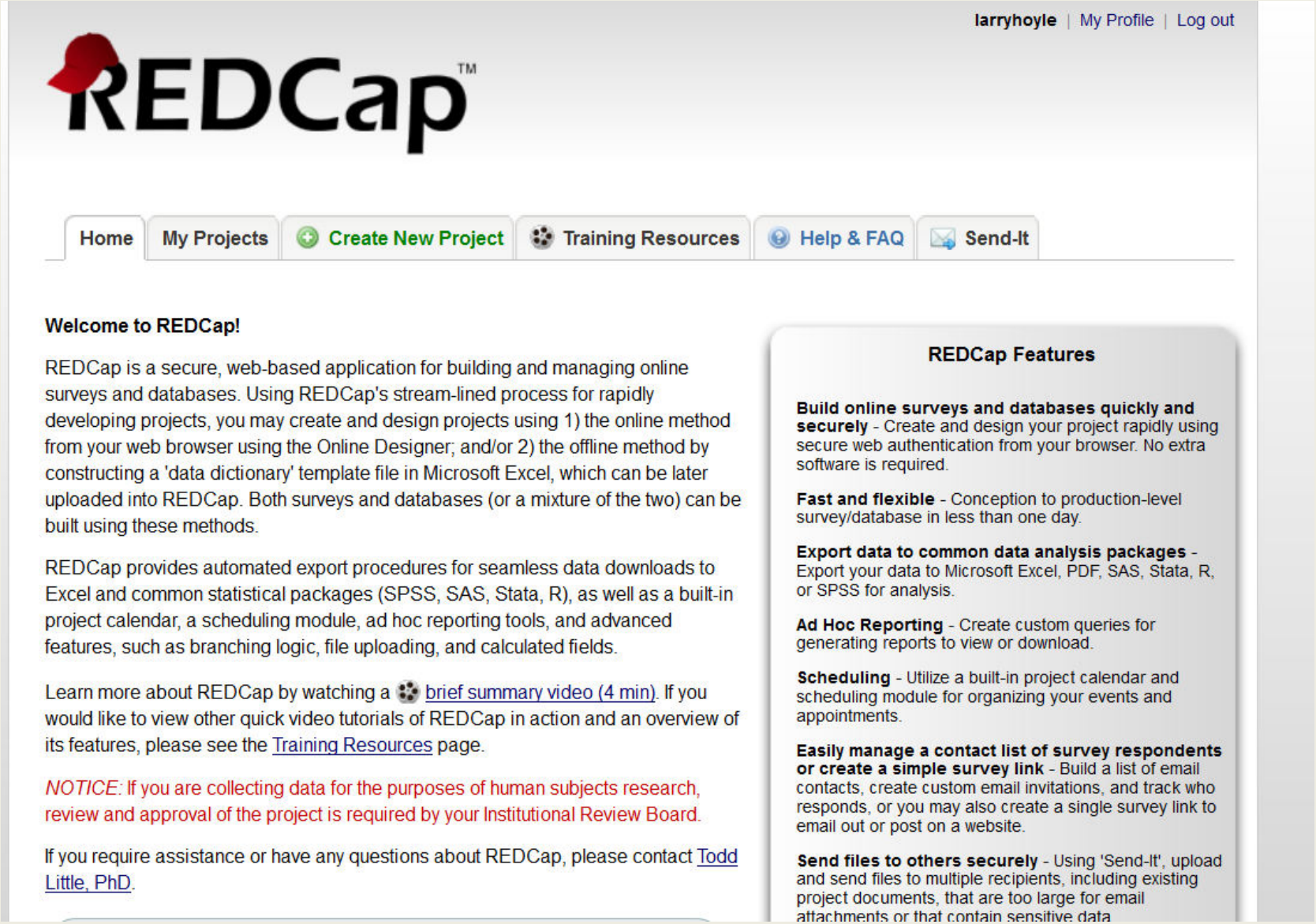
Audit Trails

Export and import

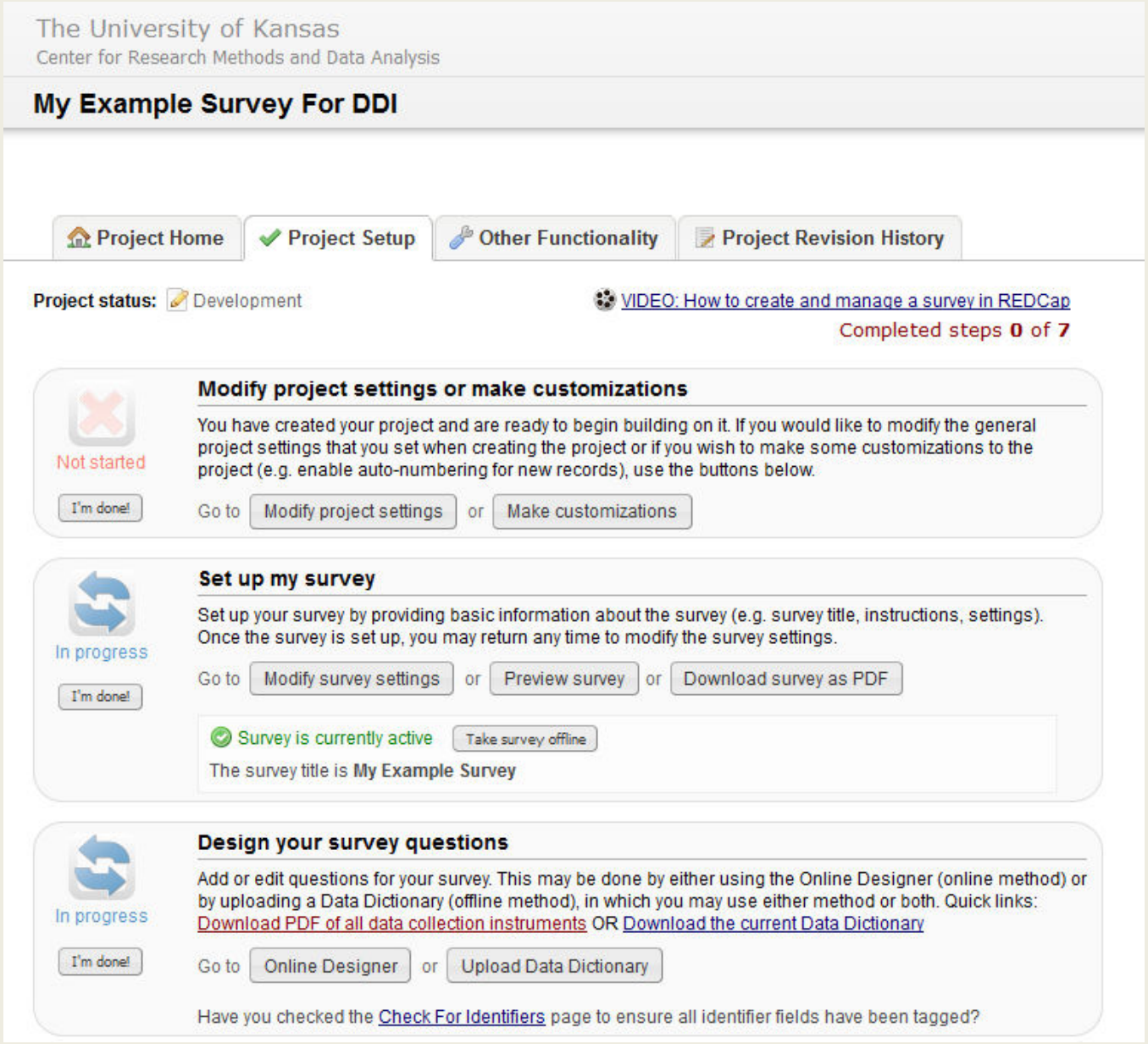
REDCap as an Environment

- Secure
- Web based
- Build and manage online surveys and databases
  - Online Designer
  - Excel template file
- Exports to SPSS, SAS, Stata, and R
- Project management features
  - Calendar, scheduling, reporting, triggers, longitudinal studies, rights and permissions

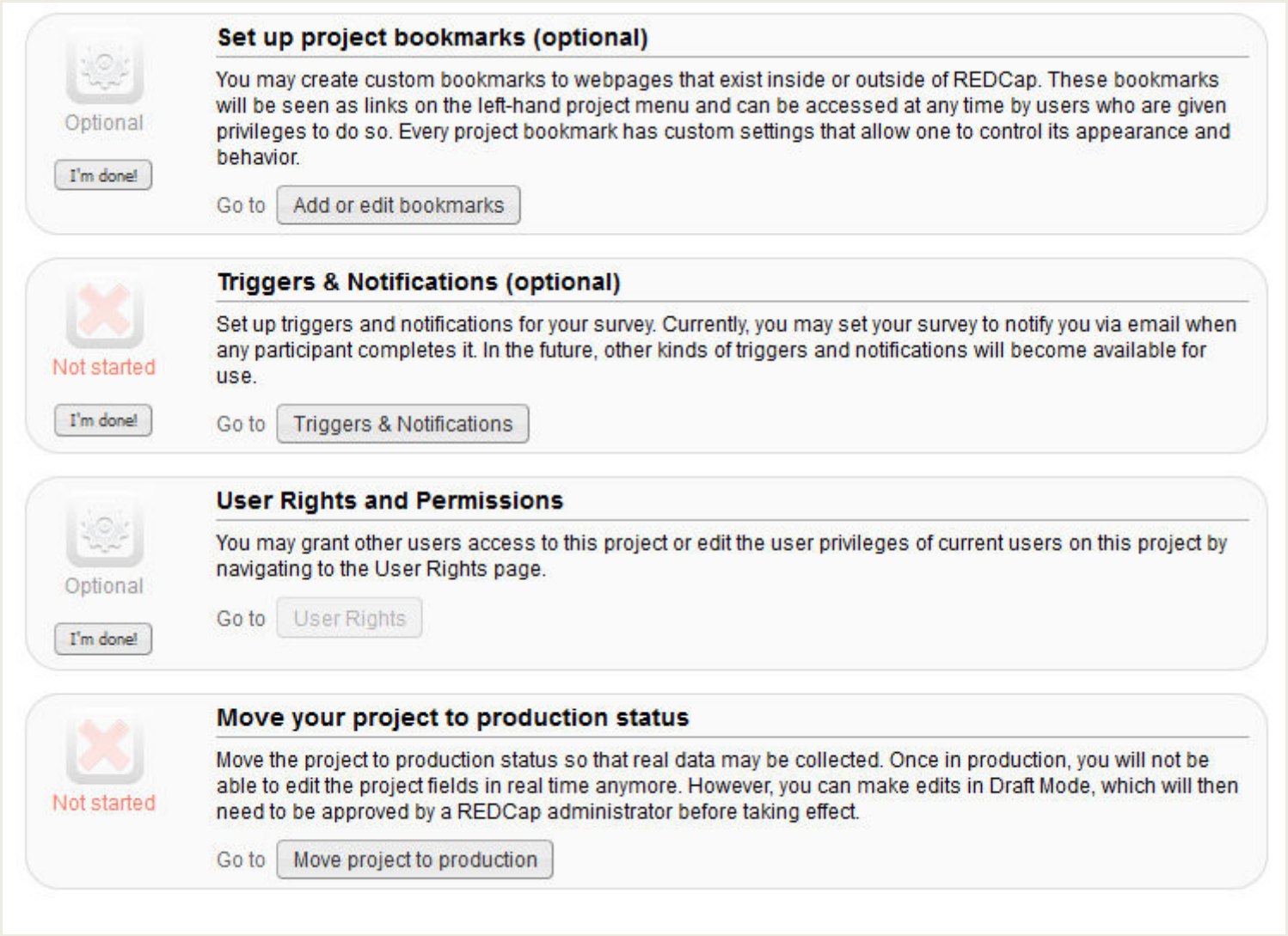
Start Screen



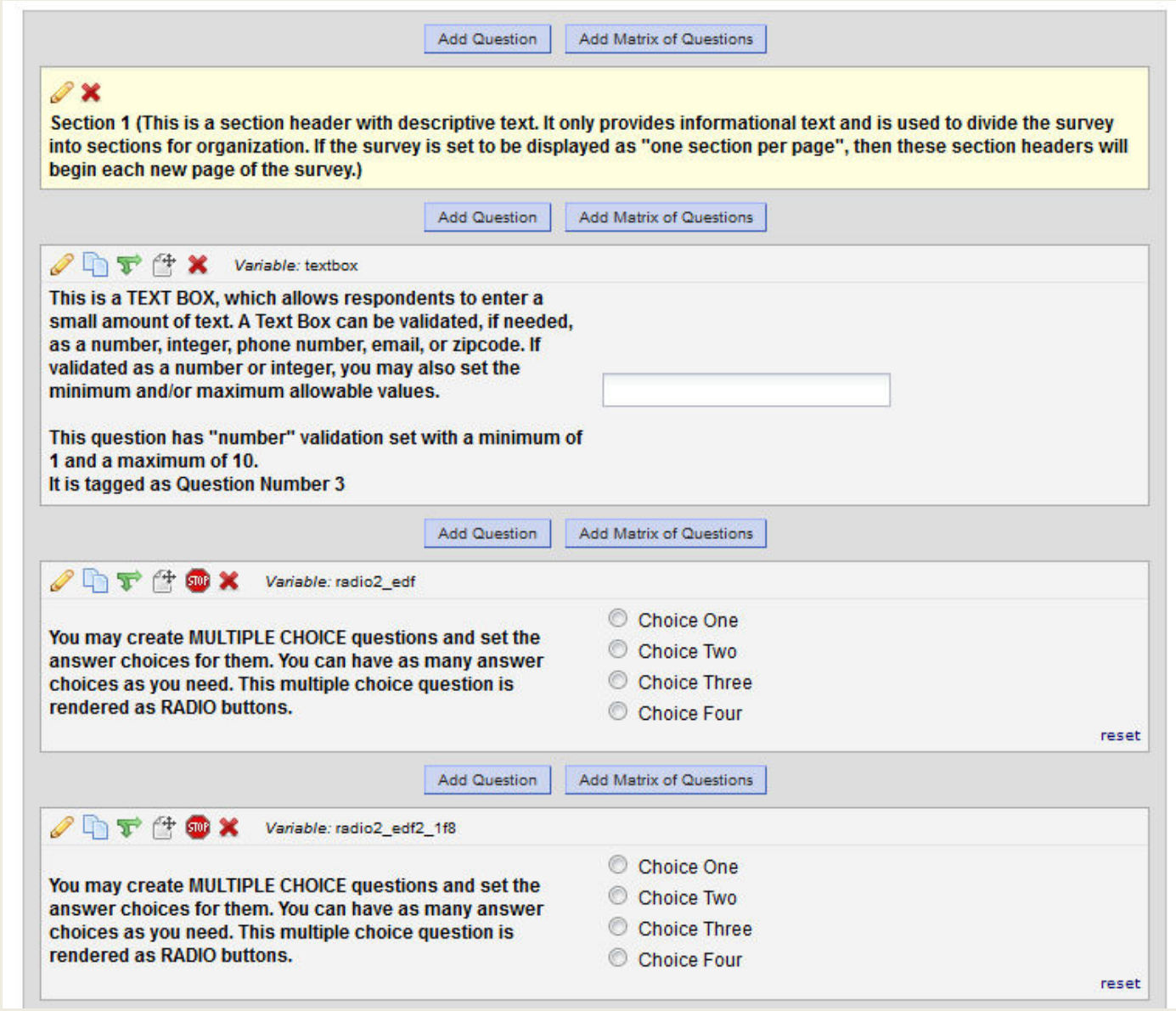
Online Designer – Project Setup



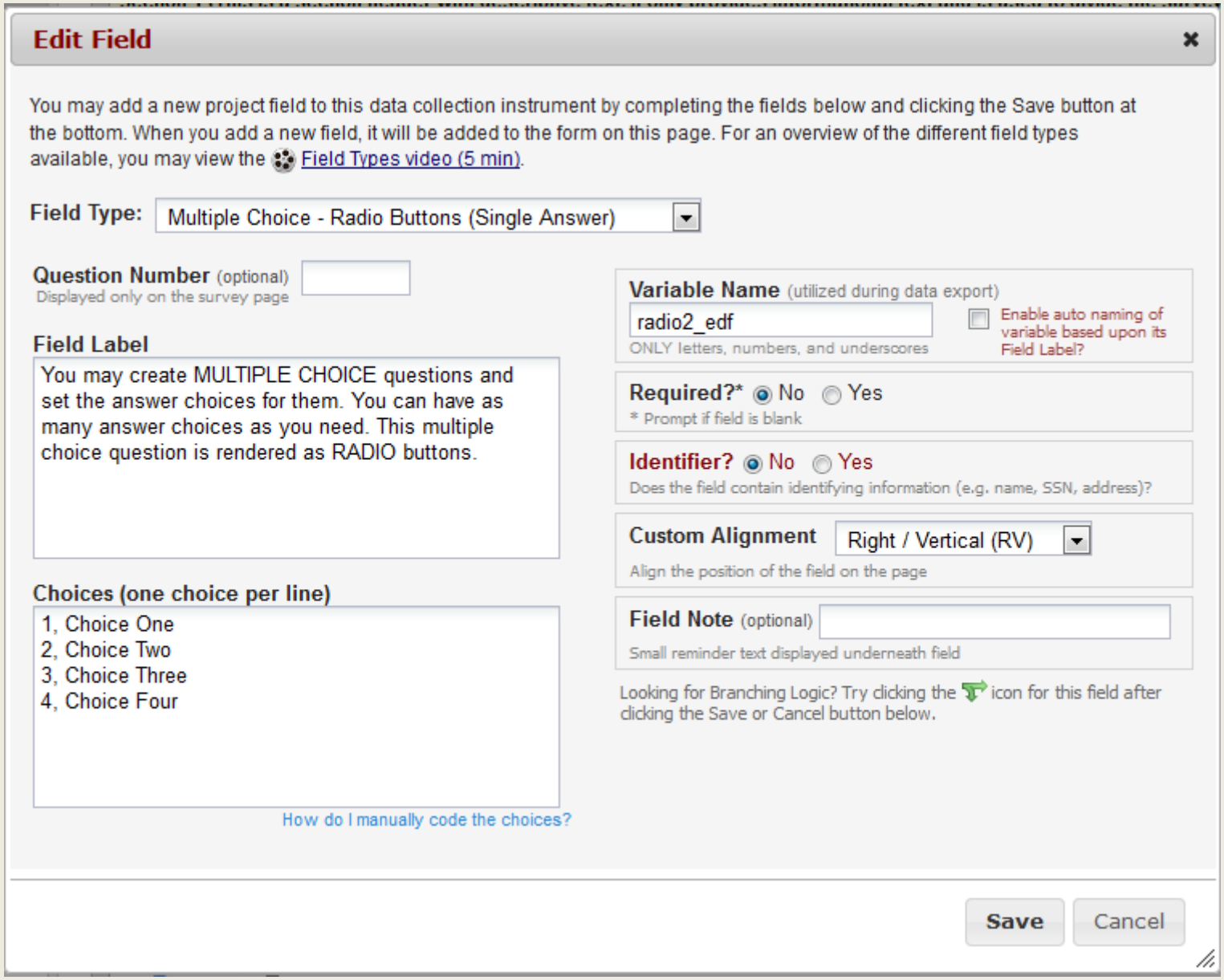
Project Setup (continued)



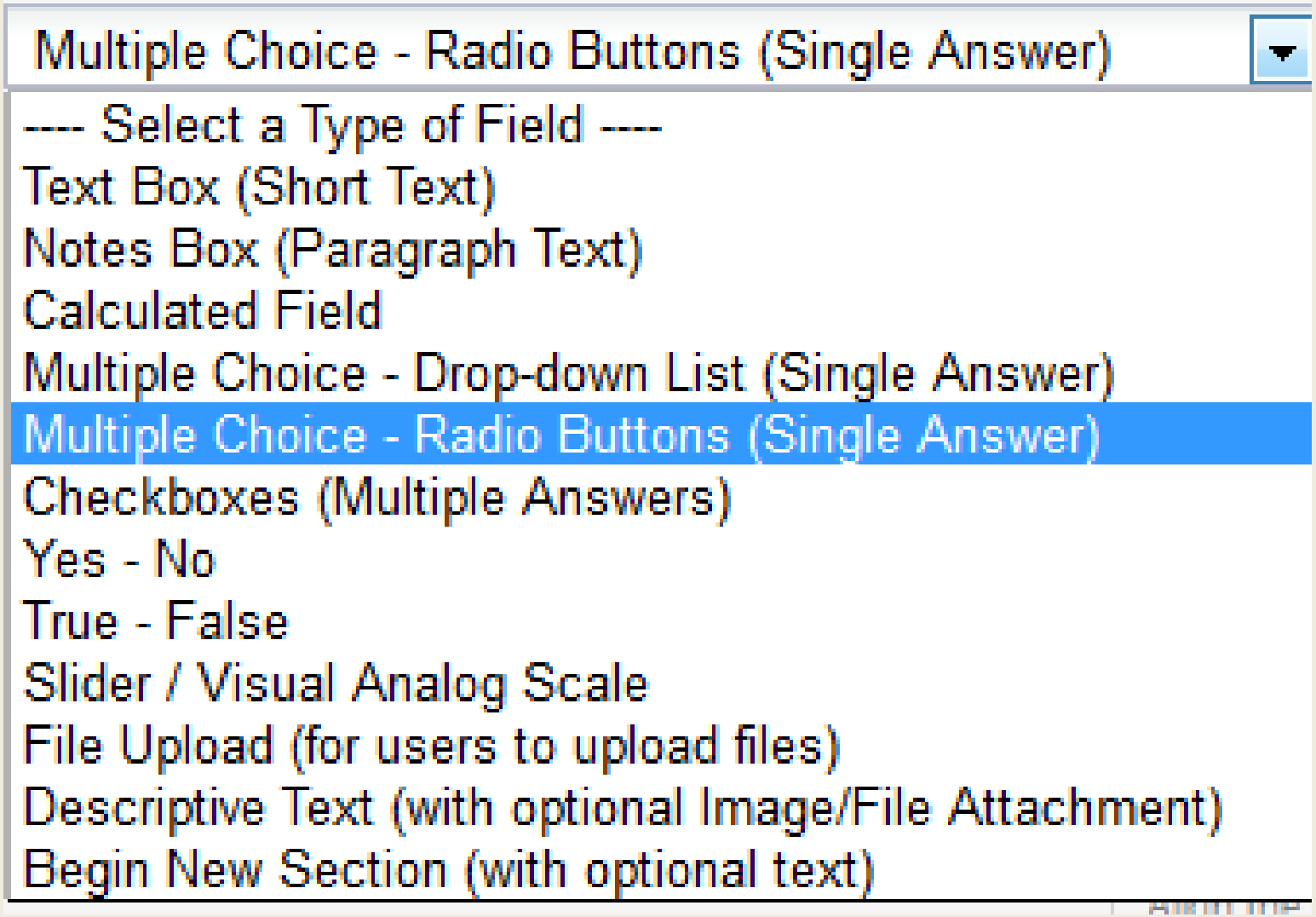
Online Designer – Question Sequence



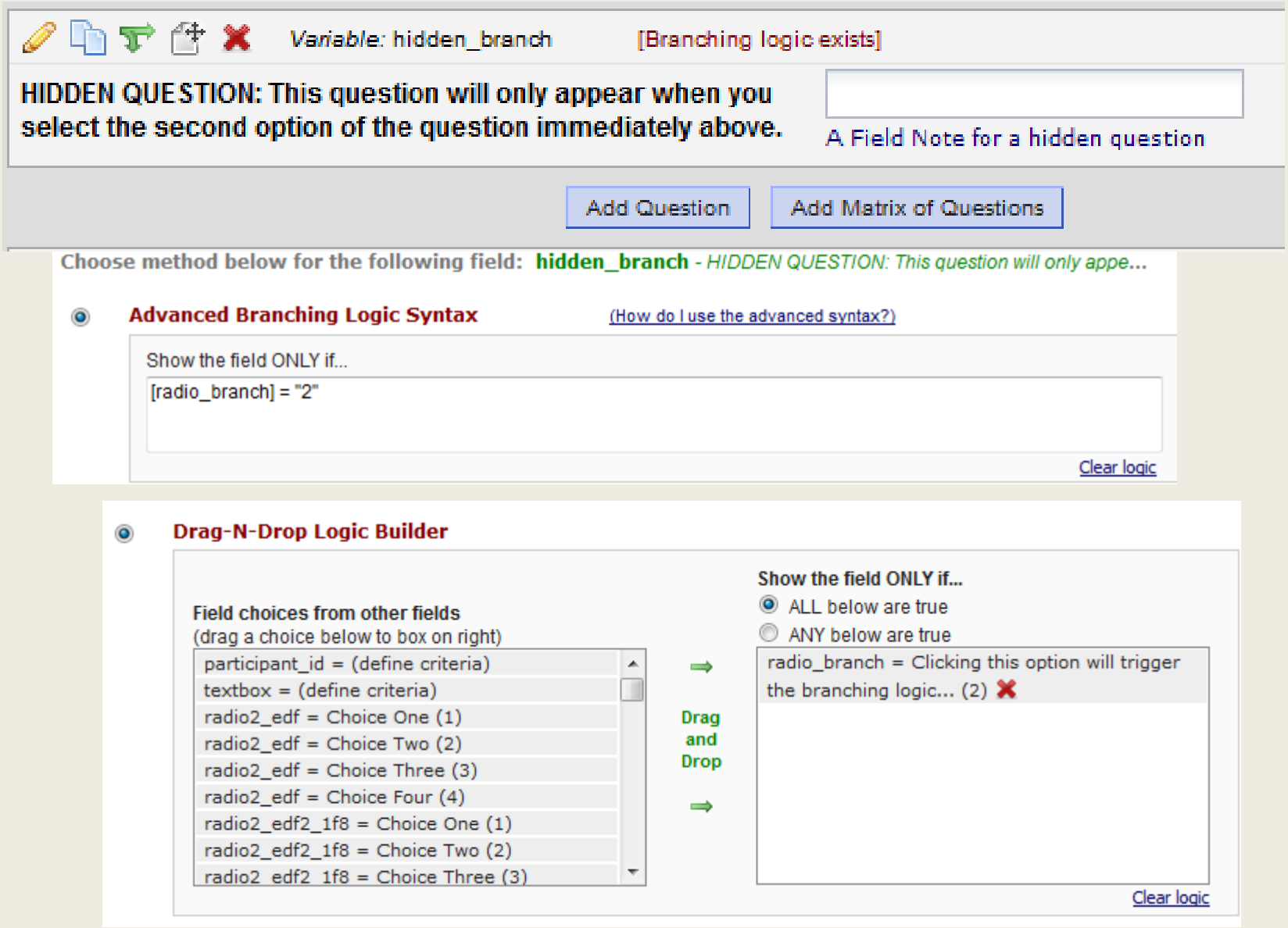
Individual Question



Field Types


Multiple Choice - Radio Buttons (Single Answer)
---- Select a Type of Field ----
Text Box (Short Text)
Notes Box (Paragraph Text)
Calculated Field
Multiple Choice - Drop-down List (Single Answer)
Multiple Choice - Radio Buttons (Single Answer)
Checkboxes (Multiple Answers)
Yes - No
True - False
Slider / Visual Analog Scale
File Upload (for users to upload files)
Descriptive Text (with optional Image/File Attachment)
Begin New Section (with optional text)

Branching


Variable: hidden_branch
[Branching logic exists]
HIDDEN QUESTION: This question will only appear when you select the second option of the question immediately above.
A Field Note for a hidden question
Add Question
Add Matrix of Questions
Choose method below for the following field: hidden_branch - HIDDEN QUESTION: This question will only appe...
Advanced Branching Logic Syntax
(How do I use the advanced syntax?)
Show the field ONLY if...
[radio_branch] = "2"
Clear logic
Drag-N-Drop Logic Builder
Field choices from other fields
(drag a choice below to box on right)
participant_id = (define criteria)
textbox = (define criteria)
radio2_edf = Choice One (1)
radio2_edf = Choice Two (2)
radio2_edf = Choice Three (3)
radio2_edf = Choice Four (4)
radio2_edf2_1f8 = Choice One (1)
radio2_edf2_1f8 = Choice Two (2)
radio2_edf2_1f8 = Choice Three (3)
Drag and Drop
Show the field ONLY if...
ALL below are true
ANY below are true
radio_branch = Clicking this option will trigger the branching logic... (2)
Clear logic

A REDCap survey description can be exported to a csv file. An R program reads the csv and writes a DDI 3.1 file.

REDCAP TO DDI

One Row, CSV File (Excel Template)

- Variable / Field Name
- Form Name
- Section Header
- Field Type
- Field Label
- Choices, Calculations, OR Slider Labels
- Field Note
- Text Validation Type OR Show Slider Number
- Text Validation Min
- Text Validation Max
- Identifier?
- Branching Logic (Show field only if...)
- Required Field?
- Custom Alignment
- Question Number (surveys only)
- Matrix Group Name

TASK

From:
One row per question
(with Header and Note)

To:
Hierarchical DDI
structure

Example: text field containing a number

CSV Column Name	CSV value
Variable / Field Name	**textbox**
Form Name	survey
Section Header	Section 1 (This is a section header with descriptive text. It only provides informational text and is used to divide the survey into sections for organization. If the survey is set to be displayed as "one section per page", then these section headers will begin each new page of the survey.)
Field Type	**text**
Field Label	This is a TEXT BOX, which allows respondents to enter a small amount of text. A Text Box can be validated, if needed, as a number, integer, phone number, email, or zipcode. If validated as a number or integer, you may also set the minimum and/or maximum allowable values. This question has "number" validation set with a minimum of 1 and a maximum of 10. It is tagged as Question Number 3
Choices, Calculations, OR Slider Labels	
Field Note	
Text Validation Type OR Show Slider Number	**number**
Text Validation Min	**1**
Text Validation Max	**10**
Identifier?	
Branching Logic (Show field only if...)	
Required Field?	
Custom Alignment	
Question Number (surveys only)	
Matrix Group Name	

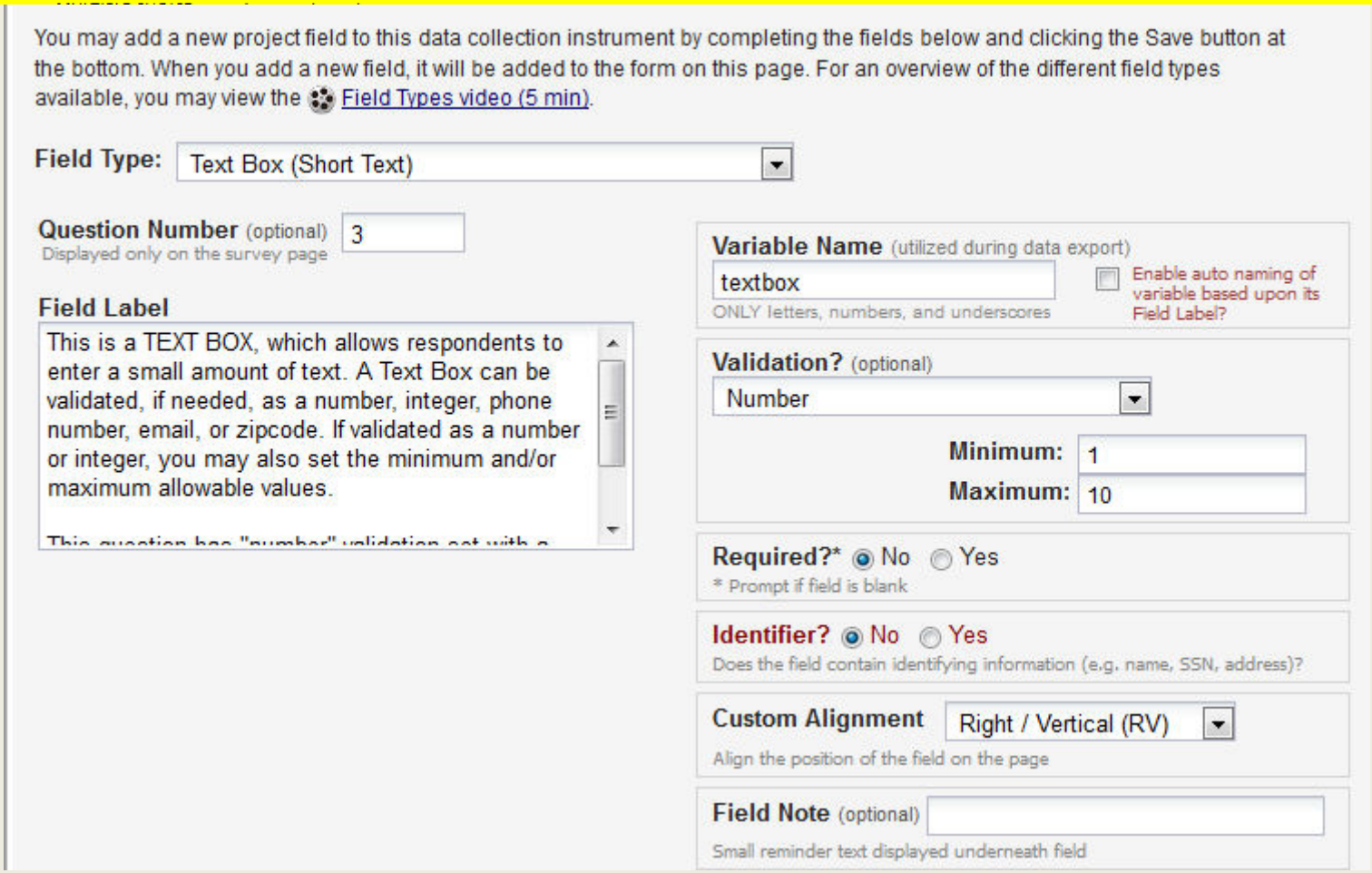
Field Type: **text**

Text Validation Type OR Show Slider Number: **number**

Extrema

Example: text field containing a number

Field Type	Field Label	Choices, Calculations, OR Sli	Field Note	Text Validation Type OR Show Slider Number	Text Validation Min	Text Validation Max
text	This is a TEXT BOX, which allows respondents to enter a small amount of text. A Text Box can be validated, if needed, as a number, integer, phone number, email, or zipcode. If validated as a number or integer, you may also set the minimum and/or maximum allowable values. This question has "number" validation set with a minimum of 1 and a maximum of 10. It is tagged as Question Number 3			number	1	10



DDI – Question w/ Numeric Domain

Field Type	Field Label	Choices, Calculations, OR Sli Field Note	Text Validation Type OR Show Slider Number	Text Validation Min	Text Validation Max
text	This is a TEXT BOX, which allows respondents to enter a small amount of text. A Text Box can be validated, if needed, as a number, integer, phone number, email, or zipcode. If validated as a number or integer, you may also set the minimum and/or maximum allowable values. This question has "number" validation set with a minimum of 1 and a maximum of 10. It is tagged as Question Number 3		number	1	10

```
<d:QuestionItem id="525CE623-5CF6-4E99-AF5B-807FCD0DB13F" version="1.0.0">
 <d:QuestionItemName>textbox</d:QuestionItemName>
 <d:QuestionText>
 </d:QuestionText>
 <d:NumericDomain type="Double">
  <r:NumberRange>
   <r:Low>1</r:Low>
   <r:High>10</r:High>
  </r:NumberRange>
  <r:Description>REDCap Field Type:  text  .</r:Description>
 </d:NumericDomain>
</d:QuestionItem>
```

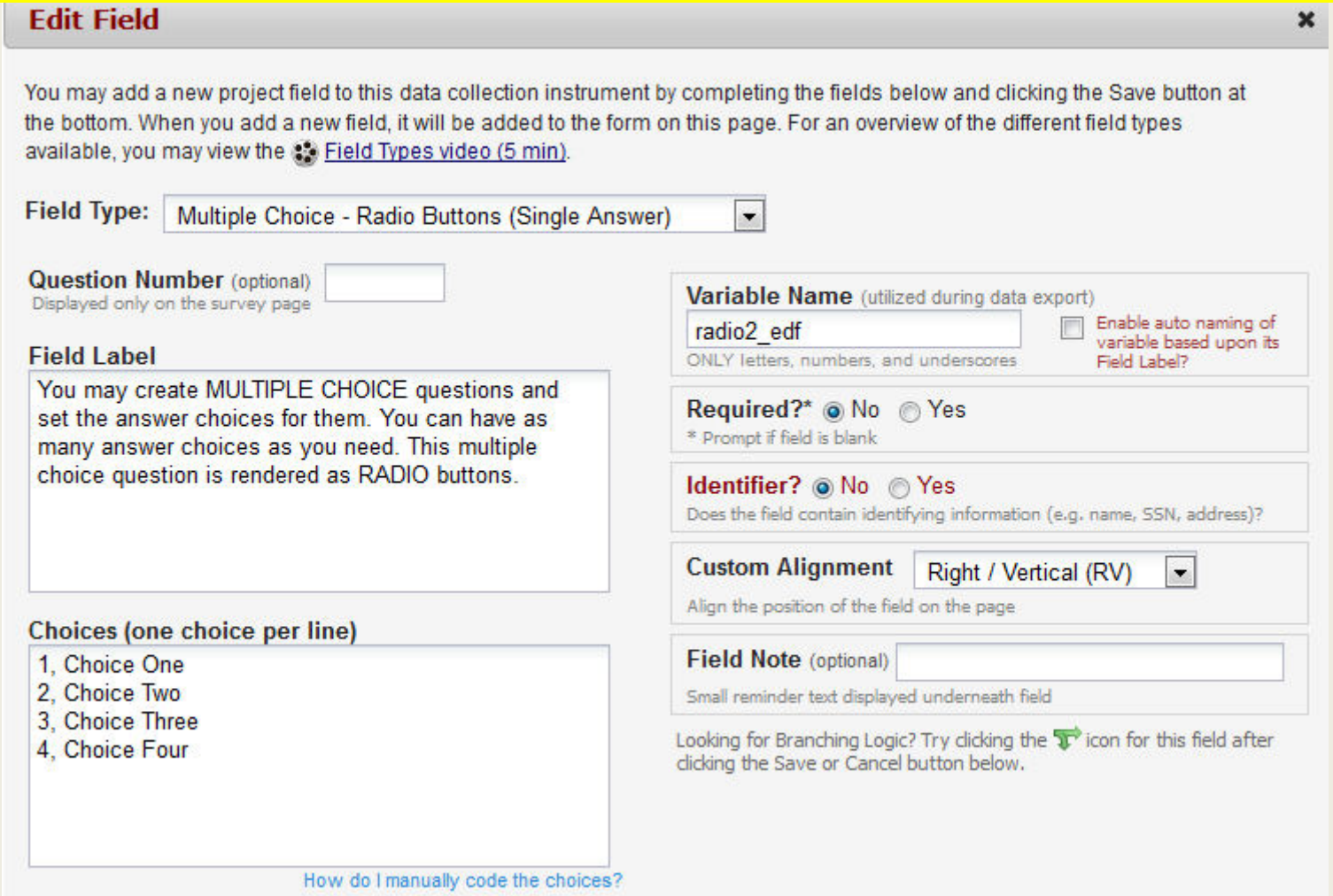
Section Header – in same row as data field

Variable / Field Name	Form Name	Section Header	Field Type	Fi
				Th
				ar
		Section 1 (This is a section header with		nu
		descriptive text. It only provides informational		as
		text and is used to divide the survey into		m
		sections for organization. If the survey is set		
		to be displayed as "one section per page",		Th
		then these section headers will begin each		1
textbox	survey	new page of the survey.)	text	It

```
<d:StatementItem id="467E8690-9399-4CFA-9528-288CF89AF310" version="1.0.0">
 <d:ConstructName>SectionHeaderForQuestion1</d:ConstructName>
 <r:Label>Section Header preceding question:  1 ,  textbox</r:Label>
 <d:DisplayText>
  <d:LiteralText>
   <d:Text>Section 1 (This is a section header with descriptive text. It only
  </d:LiteralText>
 </d:DisplayText>
</d:StatementItem>
```

Example: coded variable w/ radio buttons

Field Type	Field Label	Choices, Calculations, OR Slider Field Note
radio	You may create MULTIPLE CHOICE questions and set the answer choices for them. You can have as many answer choices as you need. This multiple choice question is rendered as RADIO buttons.	1, Choice One | 2, Choice Two | 3, Choice Three | 4, Choice Four



REDCap Code and Category Strings Repeat

Field Type	Field Label	Choices, Calculations, OR Sli Field Note
radio	You may create MULTIPLE CHOICE questions and set the answer choices for them. You can have as many answer choices as you need. This multiple choice question is rendered as RADIO buttons.	1, Choice One | 2, Choice Two | 3, Choice Three | 4, Choice Four
radio	You may create MULTIPLE CHOICE questions and set the answer choices for them. You can have as many answer choices as you need. This multiple choice question is rendered as RADIO buttons.	1, Choice One | 2, Choice Two | 3, Choice Three | 4, Choice Four

```
<l:CategoryScheme id="7723A911-9A92-4C27-80BD-295D10B2DE35" version="1.0.0" ag
 <l:CategorySchemeName>CatScheme3</l:CategorySchemeName>
 <l:Category id="32F3F66F-836B-4594-8C3B-10A034104AB1" version="1.0.0">
  <r:Label>Choice One </r:Label>
 </l:Category>
 <l:Category id="0FCAA8C9-A448-4C8B-BA26-4F526A5A9AD8" ve
  <r:Label>Choice Two </r:Label>
 </l:Category>
 <l:Category id="EB152730-D18D-4660-BECF-6EAEFF2B41FD" ve
  <r:Label>Choice Three </r:Label>
 </l:Category>
 <l:Category id="CD5239C6-B3B1-49A0-B660-DCB0D19853EF" ve
  <r:Label>Choice Four</r:Label>
 </l:Category>
</l:CategoryScheme>
```

```
<l:CodeScheme id="FCCD19BC-726D-454E-9D01-6D9178838905" ver
 <l:CodeSchemeName>CodeScheme3</l:CodeSchemeName>
 <l:CategorySchemeReference>
  <r:ID>7723A911-9A92-4C27-80BD-295D10B2DE35</r:ID>
  <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
  <r:Version>1.0.0</r:Version>
 </l:CategorySchemeReference>
 <l:Code>
  <l:CategoryReference>
   <r:Scheme>
    <r:ID>7723A911-9A92-4C27-80BD-295D10B2DE35</r:ID>
    <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
    <r:Version>1.0.0</r:Version>
   </r:Scheme>
   <r:ID>32F3F66F-836B-4594-8C3B-10A034104AB1</r:ID>
   <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
   <r:Version>1.0.0</r:Version>
  </l:CategoryReference>
  <l:Value>1</l:Value>
 </l:Code>
```

In DDI, only one Code and Category Scheme to be used by reference

Lots of References

Question references CodeScheme

```
<d:QuestionItem id="00776E9E-8DC0-4A21-BFEC-ADA427EB18C1" version="1.0.0">
 <d:QuestionItemName>radio2_edf</d:QuestionItemName>
 <d:QuestionText>
  <d:LiteralText>
   <d:Text>You may create MULTIPLE CHOICE questions and set the answer choic
  </d:LiteralText>
 </d:QuestionText>
 <d:CodeDomain classificationLevel="Nominal">
  <r:CodeSchemeReference>
   <r:ID>FCCD19BC-726D-454E-9D01-6D9178838905</r:ID>
   <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
   <r:Version>1.0.0</r:Version>
  </r:CodeSchemeReference>
  <r:Description>REDCap Field Type:  radio  .</r:Description>
 </d:CodeDomain>
</d:QuestionItem>
```

```
<l:CategoryScheme id="7723A911-9A92-4C27
 <l:CategorySchemeName>CatScheme3</l:Cat
 <l:Category id="32F3F66F-836B-4594-8C3B
  <r:Label>Choice One </r:Label>
 </l:Category>
 <l:Category id="0FCAA8C9-A448-4C8B-BA26-4F526A5A9AD8" version="1.0.0">
  <r:Label>Choice Two </r:Label>
 </l:Category>
 <l:Category id="EB152730-D18D-4660-BECF-6EAEFF2B41FD" ve
  <r:Label>Choice Three </r:Label>
 </l:Category>
 <l:Category id="CD5239C6-B3B1-49A0-B660-DCB0D19853EF" ve
  <r:Label>Choice Four</r:Label>
 </l:Category>
</l:CategoryScheme>
```

CodeScheme references CategoryScheme

```
<l:CodeScheme id="FCCD19BC-726D-454E-9D01-6D9178838905" ver
 <l:CodeSchemeName>CodeScheme3</l:CodeSchemeName>
 <l:CategorySchemeReference>
  <r:ID>7723A911-9A92-4C27-80BD-295D10B2DE35</r:ID>
  <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
  <r:Version>1.0.0</r:Version>
 </l:CategorySchemeReference>
 <l:Code>
  <l:CategoryReference>
   <r:Scheme>
    <r:ID>7723A911-9A92-4C27-80BD-295D10B2DE35</r:ID>
    <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
    <r:Version>1.0.0</r:Version>
   </r:Scheme>
   <r:ID>32F3F66F-836B-4594-8C3B-10A034104AB1</r:ID>
   <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
   <r:Version>1.0.0</r:Version>
  </l:CategoryReference>
  <l:Value>1</l:Value>
 </l:Code>
```

Note: the program generates random style UUIDs

DateTime fields

Field Label	Choices, Calculations, OR Sli	Field Note	Text Validation Type OR Show Slider Number
DATE questions are also an option. If you click the calendar icon on the right, a pop-up calendar will appear, thus allowing the respondent to easily select a date. Or it can be simply			date_ymd
DATETIME: This Datetime filed is required and has a minimum and maximum		A note to point out that this is required	datetime_ymd
TIME: a time field			time

Text Validation Type OR Show Slider Number: **date_ymd**

```
<d:QuestionItem id="ED0ACBE7-8FC5-4907-8C90-61B6C2CE7CEA" version="1.0.0">
 <d:QuestionItemName>date</d:QuestionItemName>
 <d:QuestionText>
  <d:LiteralText>
   <d:Text>DATE questions are also an option. If you click the calendar icon
  </d:LiteralText>
 </d:QuestionText>
 <d:DateTimeDomain type="Date" format="ymd">
  <r:Description>REDCap Field Type:  text  .</r:Description>
 </d:DateTimeDomain>
</d:QuestionItem>
```

Sliders

- Return a value between 0 and 100
- End and mid points labeled
- Optionally display the numeric value

- How to represent?
  - Numeric ?
  - Categories for some values (0, 50 , 100)?
    - Recorded in the NumericDomain Description

Slider as Numeric question/variable

A SLIDER is a question type that allows the respondent to choose an answer along a continuum. The respondent's answer is saved as an integer between 0 (far left) and 100 (far right) with a step of 1.


```
<d:QuestionItem id="97C40F2C-B7DA-4388-AC42-54CCB513D5A0" version="1.0.0">
 <d:QuestionItemName>slider</d:QuestionItemName>
 <d:QuestionText>
  <d:LiteralText>
   <d:Text>A SLIDER is a question type that allows the respondent to choose a
  </d:LiteralText>
 </d:QuestionText>
 <d:NumericDomain type="Integer">
  <r:NumberRange>
   <r:Low>0</r:Low>
   <r:High>100</r:High>
  </r:NumberRange>
  <r:Description>REDCap Field Type:  slider  .  NOTE: This question has a min
 </d:NumericDomain>
</d:QuestionItem>
```

This Description added to the NumericDomain

REDCap Field Type: slider .

NOTE: This question has a minimum of 0 and a maximum of 100.

The end and middle points were labeled as:

0, You can provide labels above the slider | 50, Middle label | 100, Right-hand label

Could be Slider as Coded variable


```
<d:QuestionItem id="55930335-0902-4739-9506-CAD738709E41" version="1.0.0">
 <d:QuestionItemName>slider</d:QuestionItemName>
 <d:QuestionText>
  <d:LiteralText>
   <d:Text>A SLIDER is a question type that allows the respondent to choose an answer
  </d:LiteralText>
 </d:QuestionText>
 <d:CodeDomain classificationLevel="Nominal">
  <r:CodeSchemeReference>
   <r:ID>757C6819-5850-4AE3-BACC-364790D08935</r:ID>
   <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
   <r:Version>1.0.0</r:Version>
  </r:CodeSchemeReference>
  <r:Description>REDCap Field Type:  slider  .  NOTE: The REDCap slider field returns
 </d:CodeDomain>
</d:QuestionItem>
```

This Description added to the CodeDomain

REDCap Field Type: slider .

NOTE: The REDCap slider field returns a numeric value between 0 and 100, but may have labels for the end and mid points.

CSV values for a conditional question

CSV Column Name	CSV value
Variable / Field Name	hidden_branch
Form Name	survey
Section Header	Section Header immediately before question hidden_branch
Field Type	text
Field Label	HIDDEN QUESTION: This question will only appear when you select the second option of the question immediately above.
Choices, Calculations, OR Slider Labels	
Field Note	A Field Note for a hidden question
Text Validation Type OR Show Slider Number	
Text Validation Min	undefined
Text Validation Max	undefined
Identifier?	
Branching Logic (Show field only if...)	[radio_branch] = "2"
Required Field?	
Custom Alignment	
Question Number (surveys only)	
Matrix Group Name	

The header and note will also be displayed conditionally

The Question and surrounding text will appear when this REDCap expression is true

Branching Logic (Show field only if...) **[radio_branch] = "2"**

hidden_branch,survey,"Section Header immediately before question hidden_branch ",text,"HIDDEN QUESTION: This question will only appear when you select the second option of the question immediately above.",,"A Field Note for a hidden question",,undefined,undefined,,"[radio_branch] = ""2""",,,,

ControlConstruct chain to build the Instrument

CSV Column Name	CSV value
Variable / Field Name	hidden_branch
Form Name	survey
Section Header	Section Header immediately before question hidden_branch
Field Type	text
Field Label	HIDDEN QUESTION: This question will only appear when you select the second option of the question immediately above.
Choices, Calculations, OR Slider Labels	
Field Note	A Field Note for a hidden question
Text Validation Type OR Show Slider Number	
Text Validation Min	undefined
Text Validation Max	undefined
Identifier?	
Branching Logic (Show field only if...)	[radio_branch] = "2"
Required Field?	
Custom Alignment	
Question Number (surveys only)	
Matrix Group Name	

DDI d:StatementItem

DDI d:StatementItem

DDI d:QuestionConstruct

DDI d:Sequence

DDI d:IfThenElse

DDI d:Sequence
(master sequence)

DDI d:Instrument

```
<d:Instrument id="52B1FE0E-094F-437A-96AC-778CC7C3DD2E" version="1.0.0" age
 <d:InstrumentName>REDCap Questionnaire</d:InstrumentName>
 <r:Software id="2234B4FA-D1EE-4BEE-A051-C05FF9897F91">
  <r:Name>REDCap Survey Questionnaire</r:Name>
  <r:Name>????</r:Name>
  <r:Name>REDCap Online Survey Questionnaire designer</r:Name>
 </r:Software>
 <d:ControlConstructReference>
  <r:Scheme>
   <r:ID>3ED16C54-7799-4597-9C23-82B73E6B30DD</r:ID>
   <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
   <r:Version>1.0.0</r:Version>
  </r:Scheme>
  <r:ID>895C3AA6-5EBE-4B05-AACD-3E45DEA56060</r:ID>
  <r:IdentifyingAgency>example.org</r:IdentifyingAgency>
  <r:Version>1.0.0</r:Version>
 </d:ControlConstructReference>
</d:Instrument>
```

More complex Instrument issues

- REDCap can set an option that ends the session upon entry of a particular value (a "stop action")
  - This could be represented with IfThenElse of a Sequence
    - Not done in this version
- REDCap has a Matrix Group construct
  - This is partly a layout issue but could be better documented with a sub-sequence than in this version of the program

Going the other way – defining a REDCap survey by transforming a DDI 3.1 file to a REDCap csv template.

DDI TO REDCAP

Hierarchical DDI to one row per Question

- REDCap uses Code/Category schemes by repetition instead of by reference
- DDI Instruments may be much more complex than a simple sequence with a conditional display of a question.
  - Combining REDCap calculated fields with conditional display might allow Loop, RepeatUntil etc. function.
  - Parsing Instrument and ControlConstructs not done in this version.
    - E.g. no Section Headers or Field Notes

Hierarchical DDI to one row per Question

- URNs not evaluated
- Coded questions treated as "radio" fields.
- "checkbox" (multiple response)
  - multiple variables not generated
  - placed note in CodeDomain Description (not machine actionable)
  - treated as radio
- Calculated fields not generated
  - Could look for d:ComputationItem
    - Would require parsing arbitrary programming code and generating REDCap syntax

The metadata models for REDCap and DDI are not exactly the same. Where do they not match up?

ISSUES

Issues

- Full DDI complexity not representable in REDCap
- REDCap co-mingles layout with content metadata
  dropdown, radio, slider, matrix, checkbox
- REDCap field type of “file” – file upload action
- REDCap CSV does not include project management features that it has (email triggers, scheduling.
- DDI 3.1 doesn’t handle numeric variables with some labeled values well (like slider fields)

Issues

- REDCap allows minimum and maximum for DateTime types
- DDI allows multiple low, high pairs for numeric variables – REDCap has just one pair
- DDI 3.1 has no Description for a QuestionItem, no good place to put notes about display options
- REDCap can mark a Question as "required“
  - DDI RepeatUntil?
- REDCap can mark a Question as having personally identifiable information.

Issues

- ResourcePackage vs. Studyunit
  - Where do Questions, Variables, Categories, Codes go?
  - Used ResourcePackage (more compatible with Colectica, which I used for validation)

Future

- REDCap has an api – develop extensions that would work directly in REDCap?
  - Capture (and import?) project management level metadata?

Contact

- Larry Hoyle

  Senior Scientist

  Institute for Policy & Social Research,

  University of Kansas

  1541 Lilac Lane, Blake 607

  Lawrence, KS 66045-3129

  LarryHoyle@ku.edu

  http://www.ipsr.ku.edu

R DETAILS

Used R XML Package

- Used the R package "XML"
  - http://cran.r-project.org/web/packages/XML/XML.pdf

    Author Duncan Temple Lang

    Maintainer Duncan Temple Lang <duncan@r-project.org>

    Title Tools for parsing and generating XML within R and S-Plus.

    Depends R (>= 1.2.0), methods, utils

R tools, Redcap to DDI

- Used Hash Tree representation: xmlHashTree()
  - Allowed non-sequential addition to elements
- Fixed framework added first and then elements extended
  - CategorySchemes
  - CodeScheme
  - VariableSchemes
  - QuestionScheme
  - ControlConstructs and Instrument
- Hash table to find the index number of the code and category schemes

R tools, DDI to Redcap

- Used internal tree representation from reading DDI with xmlTreeParse()
- Built code and category string in REDCap format e.g.

  1, bad, or worse | 2, good, or better

- Built a URN like string for complete category identifiers – used this as a hash key to the category value
- Did the same for CodeSchemes with the value being the REDCap code and category string