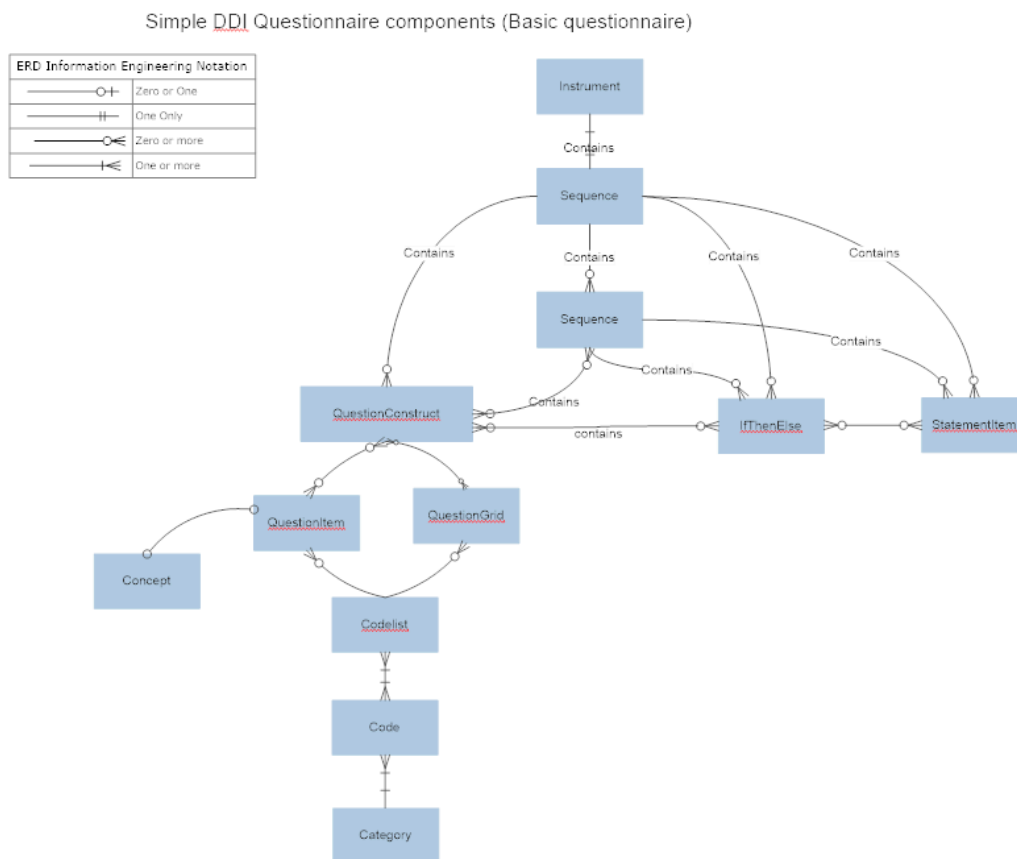


Slide 7 – Documenting Questionnaires in DDI

Section 1 – Building a basic instrument

In DDI a questionnaire is referred to as an Instrument. Instruments can be grouped together in Instrument groups and Instrument Sets and they will generally be contained in a DataCollection object inside a specific Study.

Inside an instrument we can find a series of objects that will form the content of the instrument and the flow of the questions as shown in the image.



To build a basic instrument we will need to understand what these objects are and how they interact with each other so the correct DDI version of an instrument can be built. The basic objects that we will focus on are:

- **Instruments** <Instrument>
- **Control constructs**
 - **Sequence** <Sequence>

- **Statements** <StatementItem>
- **Conditionals** <IfThenElse>
- Loops
- **Question Constructs** <QuestionConstruct>
 - **Question Reference** <r:QuestionReference>
- Repeats
- Whiles
-
- **Questions**
 - **Questions** <QuestionItem>
 - **Code lists** - <CodeList>
 - **Categories** <Category>
 - **Concepts** <Concept>
 - **Question Grids** <QuestionGrid>

Instrument <Instrument>

Defines the type of instrument used for data collection of Capture. At a minimum it should contain:

- DDI URN with agency, unique identifier and version
- Agency
- ID as an unique identifier given to the instrument
- Version
- Standard Name
- Label
- A reference to the Sequence control construct that contains the flow for data collection.

<pre> <Instrument isUniversallyUnique="true" versionDate="2025-05-09T09:08:39.8071334Z" xmlns="ddi:datacollection:3.3"> <r:URN>urn:ddi:int.example:d92c9b98-38a8-4963-a0b9-b2156beaa179:21</r:URN> <r:Agency>int.example</r:Agency> <r:ID>d92c9b98-38a8-4963-a0b9-b2156beaa179</r:ID> <r:Version>21</r:Version> <r:UserAttributePair> <r:AttributeKey>x:Languages</r:AttributeKey> <r:AttributeValue>["en-IE"]</r:AttributeValue> </r:UserAttributePair> <r:VersionResponsibility>CSOCORK\portillos</r:VersionResponsibility> <r:VersionRationale> <r:RationaleDescription> <r:String xml:lang="en-IE">Attached TOVT_CSO concept to turnover question</r:String> </r:RationaleDescription> </r:VersionRationale> <InstrumentName> <r:String xml:lang="en-IE">MSI SUR</r:String> </InstrumentName> <r:Label> <r:Content xml:lang="en-IE">Monthly Services Inquiry Form</r:Content> </r:Label> <ControlConstructReference> <r:Agency>int.example</r:Agency> <r:ID>62b587de-3f9c-4e51-8136-9cd5e058abe7</r:ID> <r:Version>19</r:Version> <r:TypeOfObject>Sequence</r:TypeOfObject> </ControlConstructReference> </Instrument> </pre>	
URN Agency, ID & Version	
Standard Name	
Label	
Reference to Sequence ControlConstruct via the Fo the SequenceID	

Note that objects are not referenced directly from within another object, rather a second type of object is used called a “ControlConstructReference” where the agency, ID, version and type of object of the referenced object are specified.

The “TypeOfObject” specified inside this ControlConstructReference can be any of the other objects that make the questionnaire.

Every Instrument will contain a Control Construct Reference to a single Sequence, which will be the top level sequency under which all other objects are organised to create/document the questionnaire.

Section 2: building a more complex instrument

- **Instruments** <Instrument>
 - Instruments
 - Instrument groups
 - Instrument Sets
- **Questions**
 - **Questions** <QuestionItem>
 - **Code lists** - <CodeList>
 - **Categories** <Category>
 - **Concepts** <Concept>
 - **Question Grids** <QuestionGrid>

- Question Blocks
- Question Groups
- Question Sets
- Interviewer Instructions
 - Interviewer Instructions
 - Interviewer Instruction Groups
 - Interviewer Instruction Sets
- **Control constructs**
 - **Sequence** <Sequence>
 - Measurements
 - **Statements** <StatementItem>
 - Actions
 - **Conditionals** <IfThenElse>
 - Loops
 - **Question Constructs** <QuestionConstruct>
 - **Question Reference** <r:QuestionReference>
 - Repeats
 - Whiles
 - Control Construct Sets
- Data Processing
 - Processing events
 - General Instructions
 - Generation Instructions
 - Processing event groups
 - Processing Instruction Groups
 - Processing Event sets
 - Processing Instruction Sets

DataCollection

A maintainable module containing information on activities related to data collection/capture and the processing required for the creation of a data product. This section covers the methodologies, events, data sources, collection instruments and processes which comprise the collection/capture and processing of data. Metadata regarding the methodology of the data collection process including, determining repetition patterns, sampling, collection modes, etc. Collection Event specifies data sources, collection instruments, questions and question flow, and data processing activities. This module houses Processing Instructions (General Instructions and Generation Instructions) which may be referenced by variables or comparison maps. The module is described by a name, label, and description, provides spatial, temporal, and topical coverage information on the activities covered by the module, and references to external material related to objects in the module using OtherMaterial. The content of the module is organized within the following sections; Methodology, DataCaptureDevelopment, Collection Event, QuestionScheme (in-line or by reference), ControlConstructScheme (in-line or by references) containing the flow of a questionnaire or data capture process, InterviewerInstructionScheme (in-line or by reference), InstrumentScheme (in-line or by reference), ProcessingEventScheme (in-line or by reference), SamplingScheme (in-line or by reference) and DevelopmentActivityScheme.

EDDI 2025 Q2 workshop content starts here!

Instrument

In DDI, a questionnaire is referred to as an Instrument. An Instrument is a group of questions that allow you to gather information on concepts of interest. As well as questions, Instruments contain other constructs such as sequences, statements and conditions - these constructs are described below. In DDI, instruments are defined as: *“Defines the type of instrument used for data collection or capture. In addition to the standard name, label, and description contains a classification of the type of instrument, a reference to an external instance of the instrument (such as an image of a questionnaire or programming script) and a reference to the Sequence control construct that contains the flow for data collection or capture.”*

Sequence

Sequences group the elements of a questionnaire and have a clear start and end. They can contain questions, statements, conditions, loops, or even other sequences. Sometimes Sequences are subject to routing (i.e. only a subset of respondents answer the questions). In DDI, sequences are defined as *“A member of the ControlConstruct substitution group. Provides a sequence order for operations expressed as control constructs. The sequence can be typed to support local processing or classification flags and alternate sequencing instructions (such as randomize for each respondent).”*

QuestionItem

A question in DDI is represented as a QuestionItem. A QuestionItem contains a label, question text, a response, and sometimes an instruction. QuestionItems should be designed for reusability, because you can use the same QuestionItem across multiple Instruments. You can't use QuestionItems in isolation - they need to be used with QuestionConstructs (see below). In DDI, a question item is: *“Structure a single Question which may contain one or more response domains (i.e., a list of valid category responses where if “Other” is indicated a text response can be used to specify the intent of “Other”). The structure provides detail on the intent of the question, the text of the question, the valid response options and the number of allowed responses, references to external aids and instructions, and an estimation of the time needed to respond to the question. Note that the QuestionItem is a reusable format for use in any number of applied uses. External aids, instructions, response sequencing etc. should contain information consistent with the general use of the QuestionItem. Additional materials and information can be added within the QuestionConstruct which is the applied use of a question.”*

QuestionConstruct

A QuestionConstruct is a wrapper for a QuestionItem - it holds the QuestionItem within the flow of the questionnaire. Unlike QuestionItems, QuestionConstructs are instrument-specific. Any customisation of a question in an Instrument (e.g. question label) should be done via the QuestionConstruct. A member of the ControlConstruct substitution group. A construct which

ties question content to the programmatic logic of the control constructs. Contains a reference to a QuestionItem, QuestionGrid or QuestionBlock, can set response or dimension sequence for use in a specific application, identifies the response unit, analysis unit, and universe. May provide an estimate of the number of minutes needed to respond.

QuestionGrid

QuestionGrids collect information on multiple related questions in a single table. DDI-L: <https://docs.ddialliance.org/DDI-Lifecycle/3.2/model/item-types/QuestionGrid/>

QuestionBlock (Jon? Someone else?)

Instruction

An Instruction is information that helps the respondent to answer the question. They are usually “doing” things (e.g. “Please select one option only”).

Provides the content and description of a single instruction. In addition to the standard name, label, and description, an InParameter can be designated to specify information needed to process the dynamic content of the instruction, an image can be associated with the instruction, and the instruction text provided using DynamicText. Note that when using Dynamic Text, the full InstructionText must be repeated for multi-language versions of the content. Different languages may handle the dynamic portions in different locations and/or with different content. Therefore languages cannot be mixed within a dynamic text except when the full text itself has multiple language sections, for example, a foreign language term in a text. The InstructionText may also be repeated to provide a dynamic and plain text version of the instruction. This allows for accurate rendering of the instruction in a non-dynamic environment like print.

Response domains

The format in which you collect your responses (e.g. text, numeric, codes).

CodeList

A multiple choice response with predetermined categories (e.g. “0 = Yes, 1 = No”). DDI-L: <https://docs.ddialliance.org/DDI-Lifecycle/3.2/model/item-types/CodeList/?highlight=codelist>

Condition (IfThenElse)

The routing of a questionnaire which directs respondents to the next applicable question. Conditions can apply to any construct such as entire sequences or specific questions. DDI-L: <https://docs.ddialliance.org/DDI-Lifecycle/3.2/model/item-types/IfThenElse/?highlight=ifthenelse>

Loop

Loops are used for repeated constructs such as a group of questions. DDI-L:
<https://docs.ddialliance.org/DDI-Lifecycle/3.2/model/item-types/Loop/?highlight=loop>

StatementItem

Text within a questionnaire that does not fulfil the criteria of any other DDI questionnaire element. Common examples include introductory or explanatory information (e.g. "The next questions are about your health"). DDI-L:

["https://docs.ddialliance.org/DDI-Lifecycle/3.2/model/item-types/StatementItem/?highlight=statement"](https://docs.ddialliance.org/DDI-Lifecycle/3.2/model/item-types/StatementItem/?highlight=statement).

Benefits

Why is it important to document your questionnaire?

Data collection is a core part of your research process and it is a research output in itself. If you don't document your questionnaire, it's not clear how you carried out your research or where your data comes from. For external people engaging with your data, this makes your research less transparent and harder to trust, reuse or replicate. It also makes it harder to understand your own research and track how your variables link to questions and data collection processes.

Questionnaire metadata provides the provenance of your data and transparency around your data collection process. People can gain an in depth insight into the questions in your questionnaire, how people could answer those questions (e.g. free text, categories etc.) and the overall structure. This makes your research easier to understand and more trustworthy, meaning it's more likely to be cited and reused. You can also deposit questionnaire metadata on data catalogues [give examples], so your research can be widely discovered.

If you are working on longitudinal study, having strong documentation from the first wave will make it easier to create and document the following waves. It also strengthens your overall documentation as you can make links across the surveys and track relationships between questions and how they evolve.

- *Provenance - can link variables to question - understand how the variable came about, how the data was produced*
 - *Makes your research more transparent and trustworthy*
 - *This information is increasingly required by funders, institutions etc.*
 - *People are more likely to reuse your data*
- *Can fully interpret the questionnaire*
-
- *Link to question banks (on catalogues) that have questionnaire documentation*
- *Makes your research and its data more FAIR*

When should you document your questionnaire?

It is best practice to document your questionnaire as you develop it in the research design phase. Documenting it from the beginning allows you to track changes and versions of the questionnaires as you move through the research lifecycle. You can also clearly map and manage the relationships between the variables and the questions.

Using DDI-L documentation can help inform the design of your questionnaire so it is structured and consistent. This will help you later in the research lifecycle when you analyse your data and share your research.

It is possible to document your questionnaire retrospectively. However, this may take more time and you may run into difficulties identifying the latest version of your questionnaire or which variables map to which questions. You may also miss the opportunity to document how the questionnaire evolves, reducing the transparency and provenance of your research.

[[You can also build on and reuse existing questionnaire documentation, for example when conducting the next wave of a longitudinal study]] - does this belong here?

Why use a standard to document your questionnaire?

Standards save you time by providing a ready-made framework of how to document your questionnaire. This includes what to document, standardised terms to describe elements of questionnaire and established relationships between different elements. They also make your questionnaire documentation more interoperable with other studies using the same standard.

If you're working on a longitudinal study, using a standard helps you be consistent with how you document your questionnaire on each wave. You can also reuse this documentation, meaning you don't have to design and document your questionnaire from scratch every time.

Some catalogues and repositories will require you to use a standard so you can deposit your (meta)data on their sites. If catalogues use the same standards, your documentation can be harvested and appear on their sites, making your data more findable.

- *Benefits of using a standard*
 - *Interoperable*
 - *Machine actionable*
 - *Links question banks*
 - *Reuse items (saves you time)*

Why DDI Lifecycle?

Lifecycle

- ((Complexity and relationships are mapped between items - it's not a flat description))
- Navigate inside the questionnaire flow