Sustainable Questionnaire Development with Colectica and Blaise

Jeremy Iverson, Dan Smith, Colectica

1. Abstract

Blaise Colectica Questionnaires is a survey specification creator, and now integrates with the new Colectica Question Bank. This new tool allows users to author questions and sequences of questions with a simple, web-based interface. Users can tag questions with categories, add custom metadata, indicate how response data appears in resulting datasets, and see surveys in which questions have been used. All changes and usages of questions are tracked in the metadata repository for a full audit trail. Documentation about the resulting data can reference the source questions to provide a full lifecycle view of the data lineage. Blaise Colectica Questionnaires can connect to Colectica Question Bank or Colectica Repository to use single questions or full sequences of questions within a survey instrument.

Using Blaise Colectica Questionnaires, surveys are designed once and can be published as PDF specifications, paper forms, Blaise 5 data capture instruments, and standardized XML. The tool allows rapid, iterative survey development and testing. Survey authors can collaborate within the tool, making comments and edits that are tracked in a metadata repository, instead of exchanging Word documents or text files by email.

In the past year, Blaise Colectica Questionnaires has also added support for in-survey search, survey validation, intelligent drag-and-drop survey organization, visual calculation editing, external data lookup, and custom Blaise code in expressions and rules.

2. Introduction

Blaise Colectica Questionnaires is a desktop application to author standards-based survey specifications. The tool provides a sustainable, auditable alternative for iterative survey design, compared to ad hoc approaches that may involve word processing documents or spreadsheets shared by email or network drives. Blaise Colectica Questionnaires can connect to a centralized repository to use and share survey components throughout an organization.

Many organizations require additional control over the authoring, categorization, and approval of questions during the survey design process. Colectica Question Bank is a new server based application with an associated web front end that focuses on development and approval workflows around reusable survey components. The tool allows groups of end users to author questions, review and approve questions, and view the full history and use of such questions. It provides an organization the ability to categorize survey components into groups and hierarchies to aid in finding existing reusable components. The focus on reusable survey components allows organizations to properly distribute and track the work of specifying, reviewing, and approving surveys.

The application server uses a distributed actor model to allow scaling to any size of deployment. It can be deployed on a single server or in a highly-available distributed cluster running on Linux or Windows. It integrates with Colectica Repository to provide centralized storage of the reusable survey components. Concurrent usage by multiple end users, changes and synchronization of survey component state, and authentication is handled by the distributed design of the application.

3. Authoring Questions with Colectica Question Bank

3.1 Question Types

Colectica Question Bank supports defining the following item types.

3.1.1 Static Content

- Descriptive Text
- Graphic
- Note

3.1.2 Standard Questions

- Multiple Choice
- Text Entry
- Numeric Entry
- Date
- Time

3.1.3 Composite Questions

- Choice Grids
- Question Grids

3.2 Question Metadata

All question types support adding metadata such as names, labels, and descriptions. Custom metadata fields can also be added to all items.

3.3 Labels

A question can be tagged with a label to make it easier to find. Labels can be created in place, or hierarchical sets of labels can be managed and assigned.

3.4 Define Output Data

When specifying a question, the author can also specify the structure of the data that is created by the question. This information includes the name and datatype of the variable or variables that are output from the survey.

If all questions in a survey specification define their output data, this means a full, variable-level description of the output dataset can automatically be created, with links back to the source questions.

3.5 Organization

Questions can be specified individually, or question sets can be created that contain ordered lists of questions. Questions in a set can be used in a survey individually, or all together.

4. Review, Approval, and Publication

Colectica Question Bank provides an approval workflow for questions to be created, reviewed, approved, and published.

4.1 Statuses

Items in Colectica Question Bank have an approval status.

4.1.1 Draft

When an item is initially created, it has the **Draft** status.

4.1.2 In Review

When the author decides a question is ready to be reviewed, they can change its status to **In Review**.

4.1.3 Approved

A reviewer can review a question and move its status to **Approved**.

4.1.4 Rejected

If a reviewer decides not to approve an item, they can set its status to **Rejected**. Rejected items can be resubmitted for review by the author.

4.1.5 User Roles

Colectica Question Bank users have one or more roles.

4.1.6 Reader

A reader can view items of any status.

4.1.7 Author

An author can create new questions. Unless an author also has the **Reader** role, they can only view their own items.

4.1.8 Approver

An approver can change the status from **In Review** to **Approved** or **Rejected**.

4.1.9 Publisher

A publisher can change the status from **Approved** to **Published**.

4.2 Comments

Users can comment on items to provide feedback as they are developed. The author can make changes in response to these comments.

5. Find Questions

Colectica Question Bank allows users to list questions, and to filter them by different facets.

5.1 Approval Status

Users can filter questions

5.2 Question Sets

Users can search all questions in the question bank to which they have access, or they can limit their search to a selected question set.

5.3 Labels

Questions can be filtered by the labels.

5.4 Full Text Search

Full text search allows users to quickly find questions based on name, question text, or text in any metadata field.

6. Question Lineage and Audit Log

Colectica Question Bank tracks all changes to question content. Approval status and comments are also saved. All this information can be reviewed, providing a full audit trail of the development of a single question or a full survey.

This rich lineage increases the credibility of the data collection methods and provides a source of documentation for future researchers who use the resulting data. This sort of lineage is not possible when using ad hoc survey authoring methods.

7. Use Question Bank Items in Blaise Colectica Questionnaires

Blaise Colectica Questionnaires has always had the ability to use questions from the Colectica Repository. Since items from Colectica Question Bank are published to Colectica Repository, Blaise Colectica Questionnaires can also insert items from the question bank. Only approved items are available for use in surveys specified in Blaise Colectica Questionnaires.

8. Question Bank Technical Details

Colectica Question Bank uses a distributed actor model to manage all aspects of the survey components, workflows, and categorizations. Since end users simultaneously interact with the application through the web interface and desktop applications, the application must be able to manage many different concurrent actions. Additionally, in order for the application to scale to large deployments and provide high availability, this concurrent usage must also operate within a distributed environment. The Colectica Question Bank usage of distributed actors allows the required concurrency across a distributed deployment using message passing and transactional component state updates.

While Colectica Question Bank can be run as a single server, it is designed to scale elastically, and is usually deployed as a group of application instances running as a cluster. When deployed in a distributed fashion, new application instances can be joined to an already running cluster. It is also fault tolerant, and can handle a machine leaving the cluster due to a machine failure or a cluster size scale down. This is useful to allow high availability, maintenance windows for the machines in the cluster, and minimizing downtime during upgrades.

The new Colectica Question Bank is also able to be deployed cross platform. It can be deployed on Linux, on Microsoft Windows, or within various containers such as docker or podman. The application server is written in C# and runs on dotnet core, an open source runtime from Microsoft. It can be deployed on physical machines or virtual machines, on premise or using a cloud provider. In addition to the current self hosted options, Colectica also has plans to offer a hosted cloud version next year.

9. Integration with Blaise Control Centre

Colectica also released an addin for the Blaise Control Centre to integrate with its new Colectica Question Bank. The addin creates a new “Insert from Question Bank” button in the Blaise Control Centre ribbon. When clicked, the user sees the Question Bank Search dialog. They can browse or search all approved questions and question sets from the question bank. When the user selects a question, Blaise code is generated and inserted into the open Blaise source code.

10. Additional New Functionality in Blaise Colectica Questionnaires

In the past year, Blaise Colectica Questionnaires has also added support for in-survey search, survey validation, intelligent drag-and-drop survey organization, visual calculation editing, external data lookup, and custom Blaise code in expressions and rules.

10.1 In-Survey Search

Users can now search for content within a survey. All question text and other metadata are searched. This feature allows users to quickly find items in a large, nested survey.

10.2 Survey Validation

Colectica Questionnaires is capable of checking for errors that may prevent the correct publication of a survey. Colectica Questionnaires currently checks for the following issues.

- Inputs that are unfulfilled
- Multiple choice options that are not numeric
- Duplicate multiple choice option values
- Display logic with missing or incomplete expressions
- Edit Checks with missing or incomplete validation expressions
- Roster definitions with incomplete or inconsistent looping information

10.3 Intelligent Drag and Drop

To reorder items within a sequence, the user can now use the mouse to drag the item to its new location.

However, reordering questions can cause logical problems with the survey. If moving an item causes any inputs to break, the user will be prompted whether to continue. If the user chooses to continue, they will need to fulfill the broken inputs using other inputs from the new scope.

10.4 Visual Calculation Editor

The visual calculation editor allows users to specify computations without writing custom source code. Computations in survey allow the author to create derived data, calculate display logic, and compute dynamic text fills. When creating a computation, the author specifies the following information:

- Inputs: what previous values are used to perform the calculation
- Outputs: what new values are created by the computation
- Expression: the functions applied to the inputs to create the outputs

The following functions are available.

10.4.1 String

- Assignment
- Concatenation

10.4.2 Number

- Assignment
- Add
- Subtract
- Multiply

- Divide
- Absolute value
- Round
- Truncate

10.4.3 Date

- Assignment
- Current date
- Add to date
- Subtract from date

10.4.4 Time

- Assignment
- Current time
- Add to time
- Subtract from time

10.5 External Data Lookup

A data lookup item allows you to search for a record from a database. One or more values from the record can then be stored with the survey data. When generating Blaise 5 source code for a data lookup, the Blaise 5 LOOKUP command is used.

10.6 Custom Blaise Code

Computations and display logic expressions can now be specified using Blaise source code, instead of using the visual editors. Input and output tokens are displayed above the code entry box. These should be surrounded by double curly braced. For example, to refer to the fields named dogYears and dogAge, code like the following can be used.

```
{{dogYears}} := {{dogAge}} * 7
```

When generating Blaise code, Blaise Colectica Questionnaires will replace the terms with the appropriate field names. This allows custom Blaise code to be re-used across instruments, even when the fields used as input to the computation may have different names in different instruments.