

Virusfinder Paper Questionnaire

Manual for the online validation system

Konrad Herbst, v0.1, 2020/12/03

Problem statement

For some study arms of our virusfinder-study promising participants eligible for further testing are identified via questionnaires. These can be returned to us by the participants in two ways: Either they fill in the forms online in which case the data is automatically tracked in a database. In case a participant returns the paper questionnaire via post we need to scan and digitize it's answers to feed the data into the database. We automatically analyze the scans and this workflow is continuously optimized. Nevertheless, we still need manual curation of the automatically determined answers. This is mainly because the paper questionnaires are often filled out ambiguously and some questions can be answered via free text.

Solution

A simple online application was implemented allowing to “crowdsource” the validation task of the scanned paper questionnaires (i.e. data curation). The application is tightly integrated with the questionnaire database and the workflow is as follows:

- A scanned questionnaire is computationally processed and the most likely answers are determined.
- These answers and the scans are uploaded to the server. The questionnaire then has an extra field “Needs validation” with the value “yes”.
- An interactive web application allows somebody (i.e. the curator) to view these automatically determined answers for one particular scan. In case the curator identifies a miss-classification of an answer she can easily change its value to the correct one.
- The curator submits the changed values to the server once she is satisfied with the result. For this the “Needs validation” field is explicitly set to “no”.
- The online application takes care that no more than one curator works on one particular questionnaire.

At this point I would like to thank Svetlana Ovchinnikova and Simon Anders very much for their super fast and competent help with setting this application up. Without them this wouldn't have been possible in such a short time-frame. Thanks a lot!

How can you help?

Now it's your turn! You can help by curating paper questionnaires using the web application. Access to this application is restricted. The questionnaires are in principle anonymous, but sometimes people write person-specific information on them. We urge you to keep such observations confidential. Please, just ignore and forget about them. If you would like to discuss

anything you see on the questionnaires you can of course always write me a mail:
k.herbst@zmbh.uni-heidelberg.de.

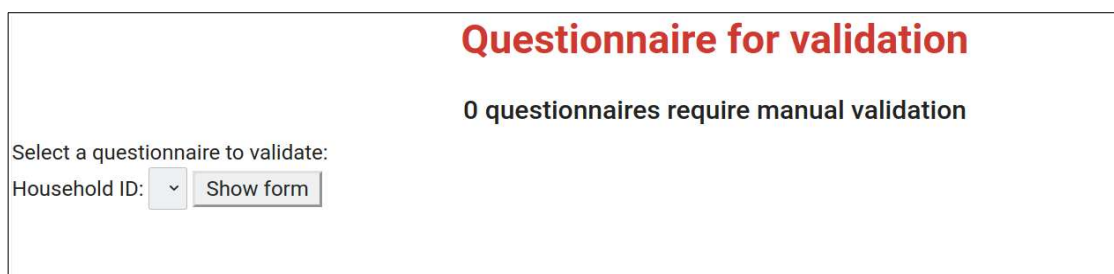
User account

Since access to the web application is restricted, you need an user account on the Django instance on which the virusfinder.de website runs. You might already have such an account. For example as a Hiwi helping in the hotline you already have such an account for filling in comments in the online tracking system. If you don't have such an account, send me a mail with your first and last name at k.herbst@zmbh.uni-heidelberg.de. I can then create an account for you if needed.

Webapplication access

The application can be accessed via https://virusfinder.de/en/paper_questionnaire:

This is how the app looks like, when no scans need to be validated. Time for a coffee break...



Questionnaire for validation

0 questionnaires require manual validation

Select a questionnaire to validate:

Household ID:

This is when there is (quite some) work to do. One can start by selecting a certain “Household ID” from the selection list or just by clicking on “Show form”.



Questionnaire for validation

679 questionnaires require manual validation

Select a questionnaire to validate:

Household ID:

Questionnaire view – What needs to be done?

This is what you see when a particular questionnaire was selected.

The screenshot displays the questionnaire interface with three main sections:

- select another form:** A dropdown menu for selecting a questionnaire to validate, with a 'Show form' button.
- scan with answer overlay:** A paper questionnaire titled 'Symptom-Fragebogen' with a QR code and instructions. It lists symptoms and their frequency, with checkboxes for 'ja' (yes) and 'nein' (no). The 'ja' boxes are highlighted in red, and the 'nein' boxes are highlighted in green.
- Questionnaire for validation:** A form titled 'Questionnaire for validation' with the text '101 questionnaires require manual validation'. It contains fields for 'SUBJECT:', 'PHONE:', 'EMAIL:', and a list of symptoms with corresponding dropdown menus for 'yes', 'no', or 'often'.

The upper left again contains the selection field for selecting a certain questionnaire. On the left we see the first page of scans for this particular questionnaire. On the right you see assigned values for particular questions, they are approximately in the same order as on the paper questionnaire. Scrolling down leads to the second page and the button for submitting the validated questionnaire answers to the database (more on this below).

Answers in scans are overlain by boxes colored depending on assigned value – more on this below. If you notice a classification mistake you can easily change the value for this answer by selecting

This detailed view shows the 'Husten mit Auswurf?' section of the questionnaire, which is divided into three rows. Each row contains a list of symptoms and their frequency, with checkboxes for 'ja' (yes) and 'nein' (no). The 'ja' boxes are highlighted in red, and the 'nein' boxes are highlighted in green. The 'Husten mit Auswurf?' section is highlighted in green. The 'EXHAUSTION' section is highlighted in red. The 'Husten mit Auswurf?' section is highlighted in green. The 'EXHAUSTION' section is highlighted in red. The 'Husten mit Auswurf?' section is highlighted in green. The 'EXHAUSTION' section is highlighted in red.

Symptom	ja	nein
Husten mit Auswurf?	<input type="checkbox"/>	<input type="checkbox"/>
Halsschmerzen	<input type="checkbox"/>	<input type="checkbox"/>
Atemnot	<input type="checkbox"/>	<input type="checkbox"/>
Muskelschmerzen/Gliederschmerzen	<input type="checkbox"/>	<input type="checkbox"/>
Müdigkeit/Abgeschlagenheit	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Kopfschmerzen	<input type="checkbox"/>	<input type="checkbox"/>

EXHAUSTION:

Symptom	yes	no	often
EXHAUSTION:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
HEADACHE:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
RUNNY NOSE:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CHEST PAIN:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

the correct value with the selection fields on the right or by clicking directly on a colored box. Clicking multiple times on a colored box cycles through the allowed values for this answer, so this is probably the fastest way to correct an answer.

Answer value options

Obviously, most answers can be answered with either “yes” or “no”. People sometimes fill in answer ambiguously, for example because they check neither “yes” or “no” or instead both. In such cases we consider the answer to be “unknown”/“-----”. Also for some questions (question 2 on old form, questions 7 and the last question “Ich lehne die Verwendung[...] ab.”) no explicit “no” checkbox is given. In this case there are only possible answers: “yes” if the checkbox is checked, “no” if unchecked.

Some questions cannot be answered simply by “yes”/“no”. People can write additional symptoms for question 2, their employment in question 5 and questions 6 and 13 are answered in form of numbers. In such cases clicking on the colored box directs your cursor directly to the respective text field so that you can type in the correct answer. Please enter the text people might have written into free text fields for questions 2 and 5. For question 6 the number of people in a household might have been not correctly identified (especially if somebody wrote it into the respective free text field). We had also already cases where more than one number was given, then just take the highest for this particular question. For question 13 answers are coded from top to bottom from 1 to 8 (highest to lowest degree) and the automatic answer algorithm (at the moment) enters all possible checked answers in the respective field. Since question actually asks only for the highest degree you might just enter the respective number coding for the highest degree (which is the lowest number value of checked boxes). Take care when this question has only four instead of eight answer options (see also below). In such cases, we still code the answers as if there were all possible options, so “5” to “8” for the bottom four checkboxes. The automatic answer identification will make a mistake here and most likely include “34”. Just remove these numbers then.

Questionnaire variations

There are three different questionnaires in circulation, one for each age group. They don't differ much, actually only the 2nd page sometimes shows some missing answer options and/or questions at the top and the bottom of the page in which case also the numbering of questions might change. Usually, answers for these questions will be anyway identified as "no" or "unknown", but you have to take care if there are four instead of eight options for question 13 "Was ist ihr höchster Bildungsabschluss?". Answers are still coded as if there were all eight options, so "5" to "8" and the answer identification algorithm will wrongly include something like "34". Please, correct the answer accordingly.

full number of value options for question 13

13) Was ist Ihr höchster Bildungsabschluss?

- ☐ Promotion
- ☐ Masterstudium, Diplomstudium an Universitäten oder Fachhochschulen
- ☐ Bachelorstudium an Universitäten oder Fachhochschulen
- ☐ Meisterausbildung
- ☐ Gymnasium (Oberstufe), Technisches-/Wirtschafts-/Berufliches Gymnasium, Ge (Oberstufe), Berufsfachschule mit Hoch-/Fachhochschulreife
- ☐ Berufsschule (Duales System), Berufsfachschule, Beamtenanwärter*in mittlere Ausbildung Gesundheits-/Sozialberufe, Berufsgrundbildungsjahr
- ☒ Hauptschule, Realschulabschluss/mittlere Reife, berufliche Schule mit mittlerer
- ☐ Kein Abschluss/noch in Schulausbildung

reduced number of options for question 13 (the question text overlaps with checkbox 3 and 4 leading to wrongly including this answer values)

13) Was ist Dein höchster Bildungsabschluss?

- ☐ Gymnasium (Oberstufe), Technisches-/Wirtschafts-/Berufliches Gymnasium, G (Oberstufe), Berufsfachschule mit Hoch-/Fachhochschulreife
- ☐ Berufsschule (Duales System), Berufsfachschule, Beamtenanwärter*in mittlere Ausbildung Gesundheits-/Sozialberufe, Berufsgrundbildungsjahr
- ☒ Hauptschule, Realschulabschluss/mittlere Reife, berufliche Schule mit mittlere
- ☐ Kein Abschluss/noch in Schulausbildung

In addition, we have optimized answering of question 2 by explicitly giving the opportunity to answer with “no”. You can identify these forms by a third column of checkboxes instead of arrows and a small filled circle at the bottom of the page. Take care: the checkboxes are then ordered from left to right as “no”, “yes”, “often”, instead of “yes”, “often” for the older version.

older form:

2) Haben Sie eine oder mehrere der folgenden Symptome (bitte Symptome ankreuzen)?

Erhöhte Temperatur/Fieber	<input type="checkbox"/>	Wenn ja	<input type="checkbox"/> Ich habe dieses Symptom häufiger
Husten (trocken)	<input type="checkbox"/>	Wenn ja	<input type="checkbox"/>
Husten mit Auswurf?	<input type="checkbox"/>	Wenn ja	<input type="checkbox"/>

newer form:

2) Haben Sie aktuell eine oder mehrere der folgenden Symptome?

	NEIN	JA	Ich habe dieses Symptom häufiger
Erhöhte Temperatur/Fieber	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Husten (trocken)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Husten mit Auswurf?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Submitting a validated questionnaire

Once you are through a questionnaire and feel comfortable that now all answers are correctly identified, you can submit these values to the database by selecting “no” for “Needs validation” (this can only be selected on the right side of the page) and clicking the “Submit” button at the bottom of the page. If you select “yes” or leave it at “at work” for “Needs validation” the questionnaire will stay in the to-be-validated-queue and eventually be handled by somebody else. This means, that if you are unsure about the correct handling of a particular answers you can just leave the questionnaire for somebody else and work on the next one.

00163
Seite 2

vecke ab.

DATETIME: 2020-12-02 10:42:40

LASTUPDATE: 2020-12-02 10:58:30

VERSION: 1

CONSENT: ☐

NEEDS VALIDATION:

Final note

If you have any question you can of course always write me a mail at k.herbst@zmbh.uni-heidelberg.de. If it is concerned with a particular questionnaire please include the household ID, so that I might be able to check myself. Usually every day from Mo to Fr, starting from around 10:00 there will be the first questionnaires in the validation queue, so if you log onto https://virusfinder.de/en/paper_questionnaire and refresh your browser once in a while you can start validating.

This leaves me with saying thank you for your help. We very much appreciate your time and effort, because it greatly helps increasing the quality of the data for the virusfinder-study. Thanks a lot!