

Research And Development of EHC:The .NET Questionnaire System in EHC

Shuai Sun, YongJian Zhang, JiaHui Yao, iSSS of PKU

1. Introduction

China Family Panel Studies (CFPS) aims to reflect the transition of Chinese society, economy, population, education and health by tracking and collecting data of three levels—individuals, families, communities, providing the data base for academic research and public policy analysis.

The Event History Calendar(EHC) is kindly used as another effective way to solve the complex information collection. Interviewee need to recall the related information for the questions, and the longer the recall period, the lower the accuracy of the information. EHC is an effective tool to help interviewee to better recall time information. Because the frequency of follow-up interview changed from originally once one year to two years, since 2014, CFPS has been used the EHC collecting information in three modules—the change of residence, work, marriage, in order to improve the accuracy of the information.

EHC was developed by the mainstream object-oriented programming language, applying to adult questionnaire of CFPS(China Family Panel Studies),to help interviewee to recall questionnaire answers and interviewers to input information conveniently through visualization and time table. In addition, it could also checkout the reasonability of the data by software check, hardware check, and overall logic checks for increasing the efficiency of data collection.

CFPS EHC has its own independent questionnaire system, which used dynamic configuration form in its question and jump logic, and its data will be written in the Blaise bdb synchronously to ensure the consistency of Blaise data and EHC questionnaire data.

2. Background

CFPS PI team began planning to add EHC module into the questionnaire of CFPS in 2012, and determine its needs in 2013. iSSS execution team then begin the work of EHC design and development, during the process, we learned the EHC program of PSID, read some articles about EHC finding from Blaise forum, which gave us a lot of inspiration but also let us take a lot of detours less.

After nearly a year of design and development work, we first used the EHC module during the execution of survey in 2014, EHC program runs relatively stable,and we carried out large-scale version of the update only once during the survey in 2014, but it's just one aspect of the evaluation for EHC program,it needs the assessment of data analysis team and PI team for EHC program

effectiveness in improving the quality of data collection.

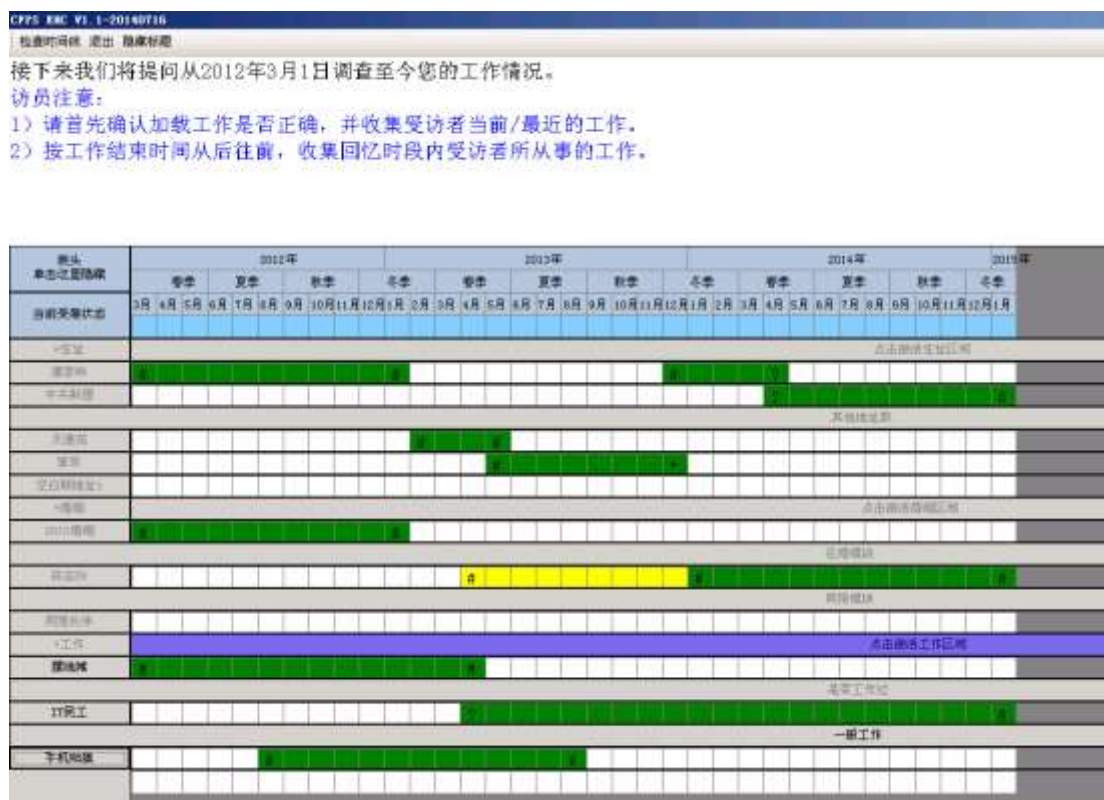


Figure 1: EHC main interface

3. Difficulties encountered

Invoking COM program written in C# for Blaise is the most difficult problems encountered in the development of early stage, we have tried to solve this problem by develop EHC using VB language which is more compatible with Blaise, but still does not solve the problem, finally we solve the problem by learning from Blaise Technical staff and finding Blaise API document. We use Blaise Alien Router function to invoke COM components, in addition we solve the problem of data transfer between Blaise DataBase and C# COM components.

It was proposed by EHC design team that the questions asked in EHC modules were too complex, which has the similar complexity with the questions in Blaise questionnaire, which took the great challenge for EHC program. To solve this problem, small questionnaire system has been developed by C# language in EHC with the character that questions and logic configured in XML files, C# language read XML files for showing the questions, jump logic in XML files will decide the next question, which has the similar appearance with Blaise. Using the tree structure displayed all answered questions and double clicking one question to the related question are the different design from Blaise.

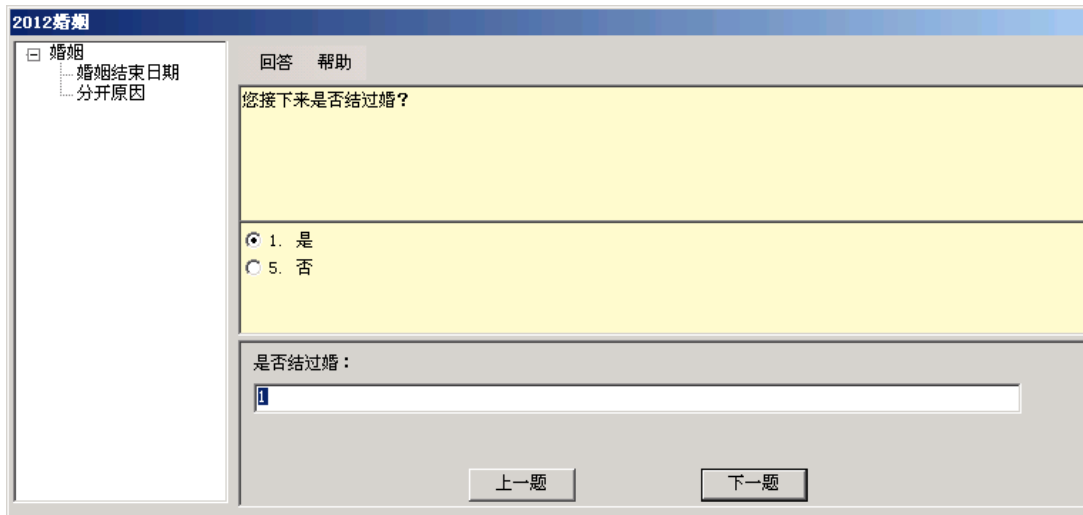


Figure 2: EHC questionnaire module

```

<Question>
  <QID>AR1</QID>
  <QName>AR1</QName>
  <BQID>EAR1</BQID>
  <BQFullName>EHCR.EAR1</BQFullName>
  <BQIfArr>False</BQIfArr>
  <QType>SingleSelection</QType>
  <QDK>False</QDK>
  <QRF>False</QRF>
  <QShowCard>False</QShowCard>
  <QHelp>False</QHelp>
  <QShorStem>2012地址确认</QShorStem>
  <QStem>根据我们的记录，您_CFPS2012_time调查时是在_residence2012居住，是吗？</QStem>
  <DicQOptions>
    <Option name="1" value="是"/>
    <Option name="5" value="否"/>
  </DicQOptions>
  <QAttention value=""/>
</Question>

```

Figure 3: XML questionnaire

```

4   <Question name="AR1" jump="AR101">
5     <Answer name="1" jump="AR102"/>
6     <Answer name="5" jump="AR101"/>
7   </Question>

```

Figure 4: XML logic jump

4. Unresolved Issues

There may be data redundancy problems in the EHC questionnaire system, all the questions interviewee answered will be stored in the current version of the EHC program, but there may be some questions ineffective. The causes of the problem are as follows:

1. EHC questionnaire system is too complex , different in the questionnaire system preload data and different questions answers can cause different question Jump.

2. Interviewer can return to modify the answer, this will lead to questions already filled out becomes invalid.

3. Interviewers can jump to a question by double-click a topic in the navigation bar on the left of the EHC questionnaire system. (The navigation bar will show the navigation information of question completed only)

4. EHC system did not consider saving mechanism for data storage of invalid questions, when save data,all of the question answers will by saved.

Considering that the issue is complicated to modify, and worry about the safety of data ,the problem is not resolved in the current version, we will consider in the next version to solve this problem.

5. Data output and format

There are three EHC data format, the first type is Blaise data(BDB), all of the EHC questions will be stored in the questionnaire BDB, so to ensure the consistency of EHC data and Blaise questionnaire data.

The second type is the SQL Server data, this section of data applies to EHC migration module.In relation to five addresses (provincial, city, county, township, Murai) selection, EHC directly read the SQL Server data of IM system, This ensures consistency of data interviewer see.

The third type is XML data, The trace data which EHC recorded will be stored as XML files. An XML file will be stored locally in the EHC at the completion of each sub-block problem, The file records the interviewer trace data during operation EHC, but the current record of the trace data is limited to the answers of the question and the submit time of the question, No Blaise ADT documented less detailed.

```

<EHCDATA>
  <DataSet ModuleName="JobData" EnterTime="9-07-26" SaveTime="9-08-08">
    <JobData>
      <GC1051 BQID="EGC1051" QType="UserInput">在家务农种地</GC1051>
      <MAINJOB_2014 BQID="MAINJOB_2014" QType="UserInput">在家务农种地</MAINJOB_2014>
      <GC1052 BQID="EGC1052" QType="StartTime">
        <QRStartYear>1900</QRStartYear>
        <QRStartMonth>1</QRStartMonth>
      </GC1052>
      <GC1053 BQID="EGC1053" QType="EndTime">
        <QREndCurrent>True</QREndCurrent>
        <QREndYear />
        <QREndMonth />
      </GC1053>
    </JobData>
  </DataSet>
  <DataSet ModuleName="JobData" EnterTime="9-07-26" SaveTime="9-08-15">
    <JobData>
      <GC201 BQID="EGC201" QType="UserInput">0</GC201>
    </JobData>
  </DataSet>
</EHCDATA>

```

Figure 5: Trace data format

6. User-friendliness

6.1 Special symbols of the beginning month and the end month

The special symbols will appear at the beginning and the end of each period in the EHC calendar. "#" means the year and the month the interviewer selected; "*" represents the uncertain year the interviewer chosen; "?" represents the uncertain month the interviewer chosen. Figure 1: EHC's home page.

6.2 EHC marriage module distinguishes the marriage from the cohabitation by color

In the EHC marriage module, there will be questions about pre-marital cohabitation and marriage. The yellow color represents the period of cohabitation, if pre-marital cohabitation emerged before marriage, and the green color means the duration of marriage.

Similarly, in a separate marriage question block, the time period for the green color, in a separate cohabitation question module, the time period for the color yellow. Figure 1: EHC main interface.

6.3 Five address selection

Five- address selection module has been developed in the EHC migration module for recording the interviewee's migration address supposed to be the same with IM system, which shares the address data with IM system, in order to ensure the consistency of the data and improve the accuracy of address data entry.

Administrative regions in China is divided into five level: provinces, cities, counties, towns

(street), Murai (community), One of the superior administrative unit may contain multiple lower-level administrative units, and each lower administrative unit belongs to only one superior administrative unit, such as a province may have more than one city, but a city can only belong to a province. we have developed five addresses linkage module against this background and uses Standards data from the National Bureau, Interviewer just need to choose to start from the provinces. (Select the primary address, the next level will automatically filter, leaving only the lower address data which is Included in the selected address)Figure 6: The address five drop-down box.

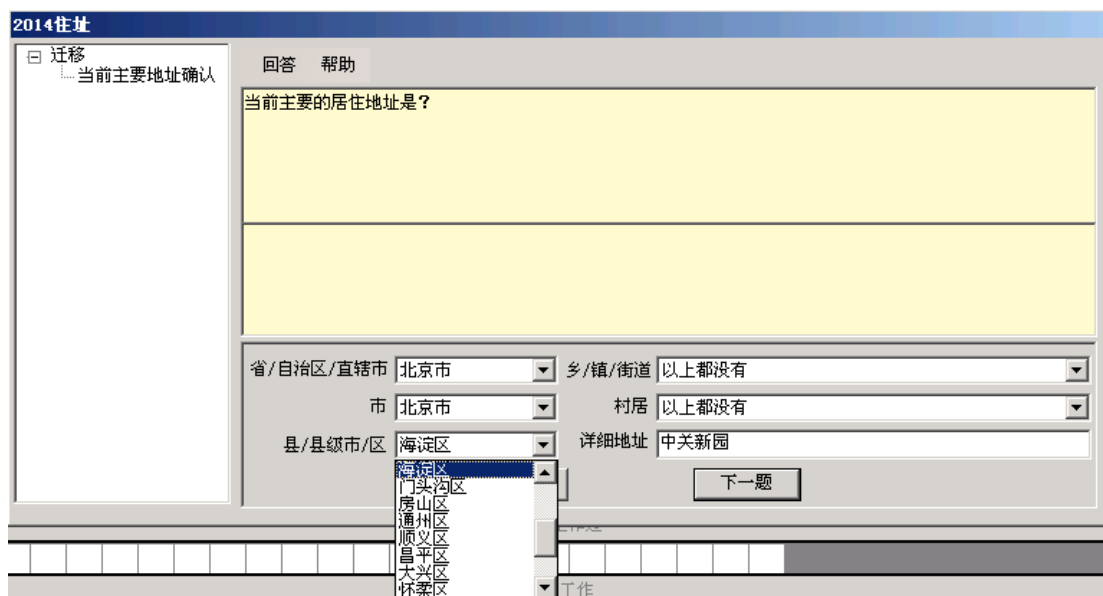


Figure 6: The address five drop-down box

6.4 Death list selection

In marriage module of EHC, if the situation interviewee spouse dies, we will ask the cause of death, where in order to maintain consistency with Blaise questionnaire, considering the data's accuracy and the convenience for the Interviewer, the death list selection module has been added, interviewers in the topic can simply double-click the left mouse button or click the keyboard spacebar to bring up the death lists. Figure 7: A list of death EHC questionnaire module.

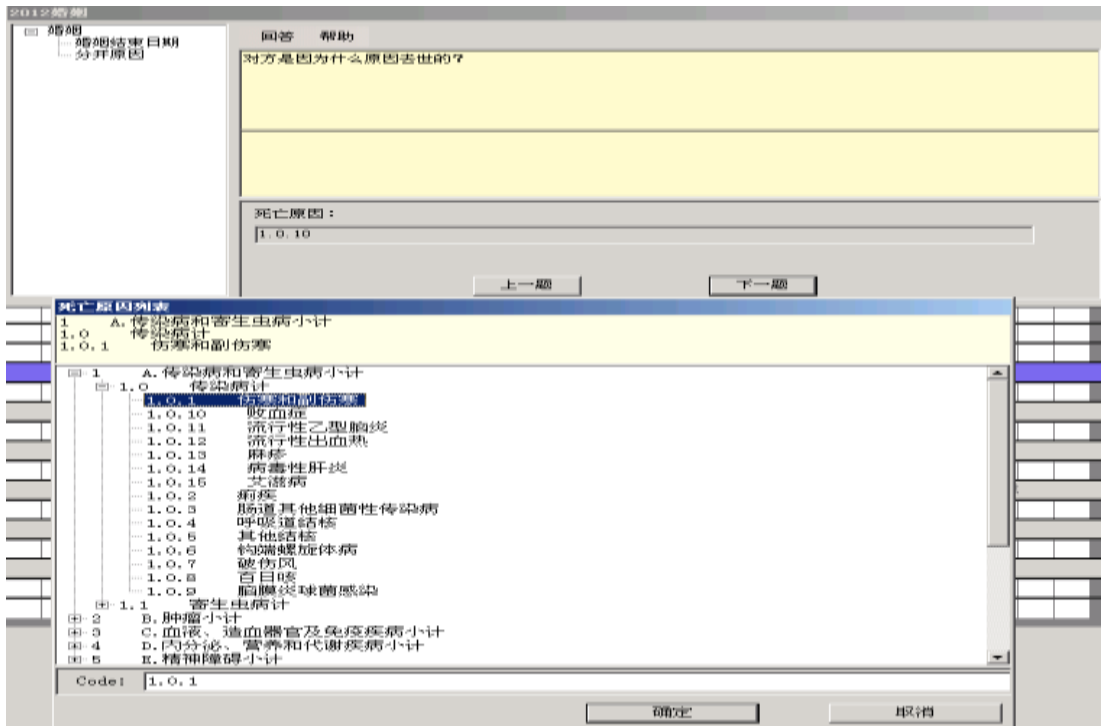


Figure 7: A list of death EHC questionnaire module

6.5 Message box

If interviewers put the mouse on any finished time region for two seconds, the system will automatically pop-up message box which displays the section title and the start and end time for the convenience of interviewers and avoiding the error due to the time period is too long or too far away from the first line of the date title.

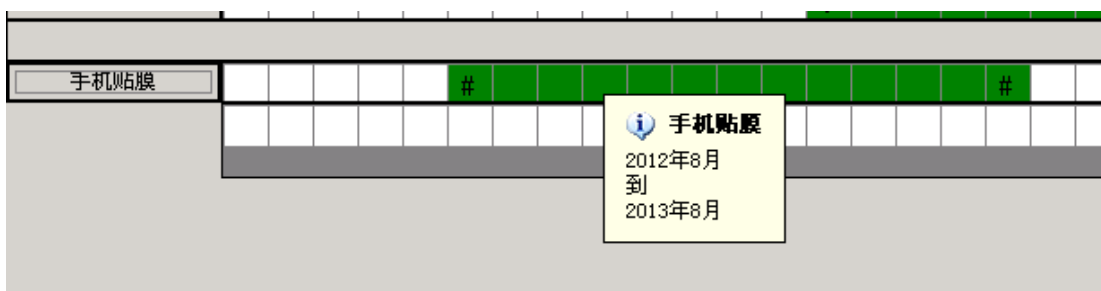


Figure 8: message box

7. Summary

After more than one year's design and development, CFPS EHC program has experienced many difficulties with expected and unexpected questions, which has already been solved one by one through the joint efforts of the iSSS'S technical group and the help of the survey team in Michigan

and Blaise technical support team. We, during that process, not only learned the new technology, but more importantly, got a lot of problem-solving approaches.

Particular thanks to the Survey Center of Michigan University, for their selfless help on technology. Thanks to other organizations and individuals who developed and used EHC before us and you are our successful guides.

References

1. Development and Programming of an Employment Event History Calendar in the Panel Study of Income Dynamics

April Beale, Mary Dascola and Youhong Liu, University of Michigan

2. Customizable .NET Event History Calendar: Looking to the Future

Roberto V Picha, Daniel Moshinsky, Mecene Desormice, Seth Benson-Flannery, U.S. Census Bureau

3. The Use of Metadata in the Design of a Customizable .NET Event History Calendar

Daniel Moshinsky, Mecene Desormice, Seth Benson-Flannery; US Census Bureau, USA