



北京大学 中国社会科学调查中心
Institute of Social Science Survey, Peking University

CFPS EHC

iSSS
PKU



Outline

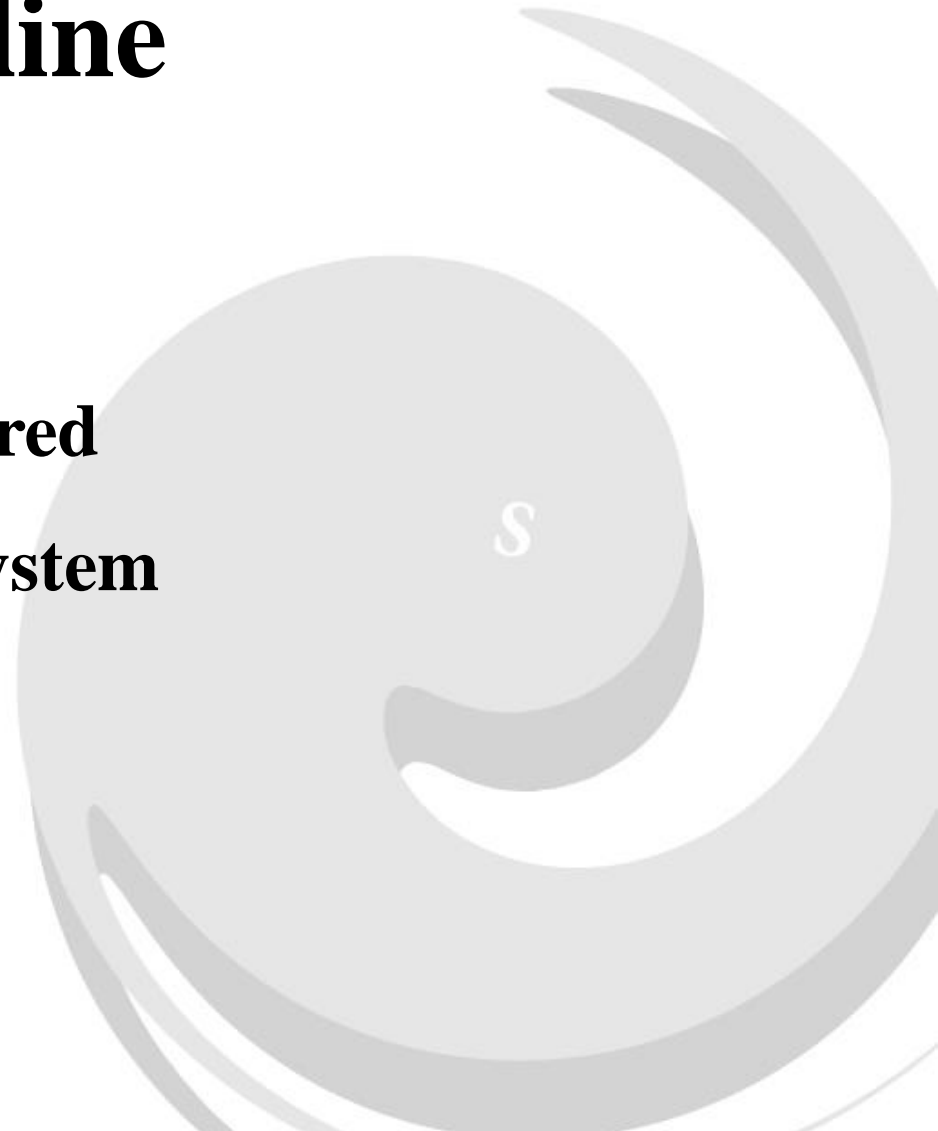
I. Introduction

II. Background

III. Difficulties Encountered

IV. EHC Questionnaire System

V. Summary & Thanks



Introduction

➤ CFPS

<http://www.iss.edu.cn/cfps/EN/>



The screenshot shows the homepage of the China Family Panel Studies (CFPS). At the top, there is a navigation bar with the following links: Home, About CFPS, News, Documentation, Data Center, For Interviewees, and For Interviewers. Below the navigation bar, the main heading reads "CFPS China Family Panel Studies 中国家庭追踪调查". The main content area features a paragraph describing the survey:

China Family Panel Studies (CFPS) is a nationally representative, annual longitudinal survey of Chinese communities, families, and individuals launched in 2010 by the Institute of Social Science Survey (ISSS) of Peking University, China. The CFPS is designed to collect individual-, family-, and community-level longitudinal data in contemporary China. The studies focus on the economic, as well as the non-economic, wellbeing of the Chinese population, with a wealth of information covering such topics as economic activities, education outcomes, family dynamics and relationships, migration, and health. The CFPS is funded by the Chinese government through Peking University. The CFPS promises to provide to the academic community the most comprehensive and highest-quality survey data on contemporary China. Three key features of the CFPS are worth noting here:



Background

- Planning : 2012
- Determine the requirements :2013
- Starting development :2013
- First Apply :CFPS2014





Difficulties Encountered

- Invoking COM program written in C# for Blaise

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.



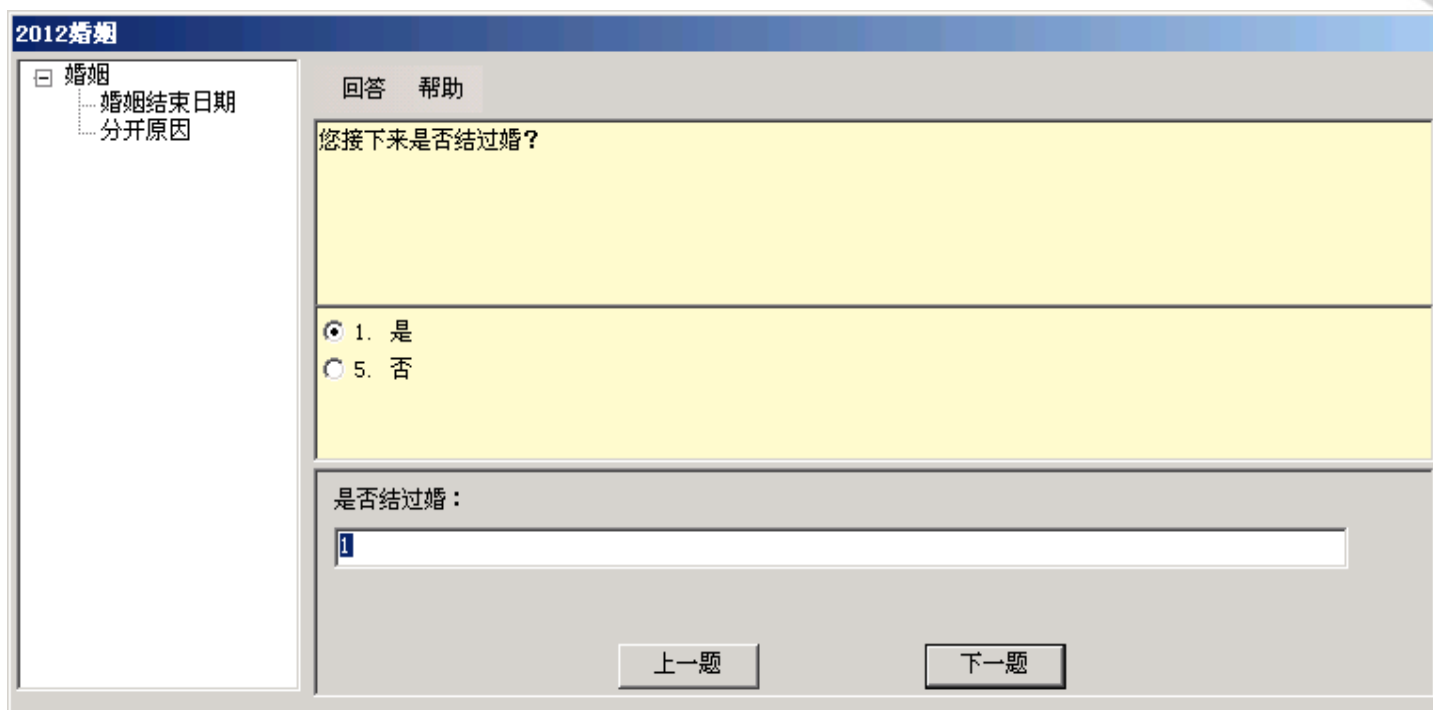
EHC Questionnaire System

➤ Development background

It was proposed by EHC design team that the questions asked in EHC modules were too complex, which took the great challenge for EHC program. To solve this problem, a questionnaire system has been developed by C# language in EHC

EHC Questionnaire System

➤ Screenshot



2012婚姻

婚姻
└─ 婚姻结束日期
└─ 分开原因

回答 帮助

您接下来是否结过婚？

1. 是
 5. 否

是否结过婚：

上一题 下一题



EHC Questionnaire System

➤ Questionnaire XML File Format

```
<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>
```

➤ Questionnaire Jump Logic XML File Format

```
1   <Question name="AR1" jump="AR101">
2     <Answer name="1" jump="AR102"/>
3     <Answer name="5" jump="AR101"/>
4   </Question>
```



EHC Questionnaire System

➤ Problem To Be Resolved

There may be data redundancy problems in the EHC questionnaire system, all the questions respondent answered will be stored in the current version of the EHC program, but there may be some questions ineffective.



EHC Questionnaire System

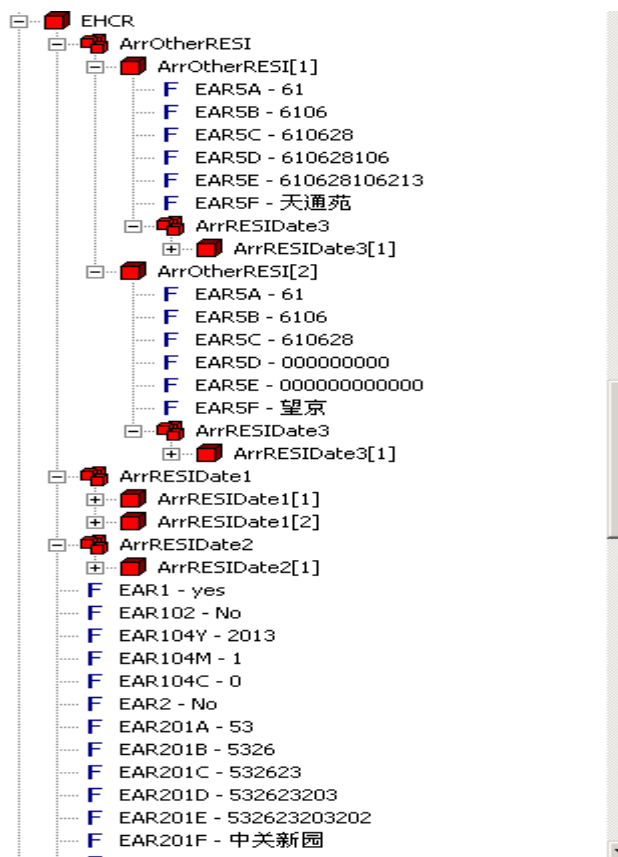
➤ Data Output

There are three EHC data format.

- First :Blaise database(BDB)
- Second : SQL Server data(MDF)
- Third :XML data

EHC Questionnaire System

➤ Data Output



	sCodeType	sCodeValue	vDescription	vLongCodeValue
1	PROV	61	陕西省	
2	CITY	6101	西安市	61
3	DIST	610102	新城区	6101
4	STRE	610102001	西一路街道办事处	610102
5	CONT	610102001001	民乐社区居委会	610102001
6	CONT	610102001003	红会社区居委会	610102001
7	CONT	610102001004	广场社区居委会	610102001
8	CONT	610102001005	尚朴社区居委会	610102001
9	CONT	610102001007	兴盛社区居委会	610102001
10	CONT	610102001008	新民社区居委会	610102001
11	CONT	610102001009	西三路社区居委会	610102001
12	CONT	610102001012	兴乐社区居委会	610102001
13	STRE	610102002	长乐中路街道办事处	610102
14	STRE	610102003	中山门街道办事处	610102
15	STRE	610102004	韩森寨街道办事处	610102
16	STRE	610102005	解放门街道办事处	610102
17	STRE	610102006	自强路街道办事处	610102
18	STRE	610102007	太华路街道办事处	610102
19	STRE	610102008	长乐西路街道办事处	610102
20	STRE	610102009	胡家庙街道办事处	610102



EHC Questionnaire System

➤ Trace Data Format

```
<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>
```



EHC Questionnaire System

➤ Five Address Selection

Five- address selection module has been developed in the EHC migration module for recording the respondent's migration address.

EHC Questionnaire System

➤ Five Address Selection Screenshot

2014住址

迁移
当前主要地址确认

回答 帮助

当前主要的居住地址是?

省/自治区/直辖市 北京市 乡/镇/街道 以上都没有

市 北京市 村居 以上都没有

县/县级市/区 海淀区 详细地址 中新园

海淀区
门头沟区
房山区
通州区
顺义区
昌平区
大兴区
怀柔区

下一题

工作



EHC Questionnaire System

➤ Death List

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.

EHC Questionnaire System

➤ Death List Screenshot

2012婚调

婚姻
婚姻结束日期
分开原因

回答 帮助

对方是因为什么原因去世的?

死亡原因：
1.0.10

上一题 下一题

死亡原因列表

- 1 A. 传染病和寄生虫病小计
 - 1.0 传染病计
 - 1.0.1 伤寒和副伤寒

1 A. 传染病和寄生虫病小计

- 1.0 传染病计
 - 1.0.1 伤寒和副伤寒
 - 1.0.10 败血症
 - 1.0.11 流行性乙型脑炎
 - 1.0.12 流行性出血热
 - 1.0.13 麻疹
 - 1.0.14 病毒性肝炎
 - 1.0.15 艾滋病
 - 1.0.2 痢疾
 - 1.0.3 肠道其他细菌性传染病
 - 1.0.4 呼吸道结核
 - 1.0.5 其他结核
 - 1.0.6 钩端螺旋体病
 - 1.0.7 破伤风
 - 1.0.8 百日咳
 - 1.0.9 脑膜炎球菌感染
- 1.1 寄生虫病计

- 2 B. 肿瘤小计
- 3 C. 血液、造血器官及免疫疾病小计
- 4 D. 内分泌、营养和代谢疾病小计
- 5 E. 精神障碍小计

Code: 1.0.1

确定 取消



Summary

After more than one year's design and development, CFPS EHC program has experienced many difficulties, which most of them 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.



Thanks

Particular thanks to GINA and the Michigan Survey Team, 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.



Thanks

