CATI follow-up Application

Resolving Communication links between applications using XML

Stéphane Lamarre, Statistics Canada, October 19th 2010

Outline

- Introduction
- Complexities surrounding our Business CATI Applications
- Proposed Design Solutions
- Experimenting with XML
- Conclusion

Introduction

- Designing a Business Blaise application to be used in a multimode environment has resulted in a strong tool that can be used for CAPI, CATI systems integrated with our non-response follow-up process.
- Having Blaise interface with external systems in a Multi-Mode environment has given us some interesting design challenges in data transfer between systems.

Complexities surrounding our CATI Apps

- The Regional Office Business Surveys have several complexities to them that have an impact on how we load and evaluate the data.

- These complexities involve:
  - collection infrastructure
  - design standards for the applications.

Complexities surrounding our CATI Apps Collection Infrastructure

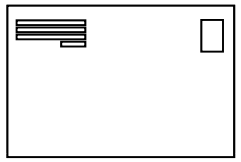
- The Business Register
  - A system that contains information on all businesses across Canada.
  - Source for identifying the samples for different business surveys.
  - The information contained in the BR includes
    - details on the businesses
    - information on the respondents
    - information involving the collection burden on each of the respondents.

Complexities surrounding our CATI Apps

- The Business Survey Regional Database (BSRD)
  - The BSRD is a MS SQL Server database with a .Net interface that manages samples during collection.
  - Inherits information about the businesses and respondents from the BR if different than collection.
  - Contains data and flags about the details on how each case is to be collected. (preferred method of collection)
  - Ability to set the Method of Collection for Internet collection
  - Placeholder

Complexities surrounding our CATI Apps

- Collection Modes
  - **Data capture mode (Blaise)**
    - involves a team that captures the surveys received by mail or fax in “non-editing” mode.



  - **Telephone capture mode (Interview)**
    - guides the interviewer to correcting the data or suppressing the edits that are confirmed by the respondent.
    - The telephone mode is primarily used for non-response follow-up.

Complexities surrounding our CATI Apps

- **The Collection Cycle**
  - The collection cycle is monthly with 18-20 business days to collect the required data for a given survey
  - The questionnaires need to be either completed or followed-up via a CATI application within a given cycle.
  - Timing between multi-mode collection systems can cause complications in balancing the work load.

Complexities surrounding our CATI Apps

Mixed collection frequencies in samples

- Businesses can choose to report
  - Annually,
  - Semi-Annually
  - Quarterly
  - Every 4 months
  - Monthly
  - Every 4 weeks

- These differing frequencies impact the size of the sample from one cycle to another.
- Requires our data repository (BSRD) to manage and maintain the collection characteristics for each unit

Complexities surrounding the CATI Apps

The Collection Roster

- The Monthly surveys are rostered with questionnaires (cycles) to collect / follow-up on.
- Currently, our Monthly data model can hold up to 15 previous cycles.
- Following-up on units by telephone from an external system can be time constraining on the whole collection process.

Design Solutions

- **Loading the Internet Data**
  - Implement two way communication between the Internet Collection System and the Regional Office Business Surveys.
  - Add controls to manage the collection responsibility.

Design Solutions

- **Two Way Communcations**
  - The sample is loaded in the CATI application and units destined for the ICS would be extracted in XML.
  - This CATI application would then receive / edit / follow-up on the processed units received from the ICS

Design Solutions

- **Place holder for business and contact info**
  - Business and contact info updates now have a place holder for review (during follow-up).
  - The interviewer needs to verify the updates done from ICS or mail capture modes.
  - This new process ensures the latest updates are sent to the BR for their review / updates.

Design Solutions

- **Controls for Collection Responsibility**
  - Exclude Mail, Fax and Internet collection units from the CATI day batch until these units are received for follow-up.
  - Set a flag for the units with the collection method “Telephone” resulting from the data capture process and from the ICS.
  - Release all Mail and Internet units in the current cycle to be touched and followed-up. This ensures that all efforts were made to collect the necessary data within a given time frame.

Receiving Data from the Internet

- There are two systems designed for collecting data over the internet
  - Electronic Data Reporting (EDR)
  - E-Questionnaire

Receiving Data from the Internet

- Electronic Data Reporting System
- A key file provided as input to the EDR system
  - Example of a key file:

    MSM

    Profile.Legal_Name

    Profile.Operating_Name

    Profile.Contact_Name

    Comments.EDR_1

    Comments.EDR_2

    CMonth.Questions.GOM

    CMonth.Questions.Total_Shipments

Receiving Data from the Internet

- Electronic Data Reporting System
- Receive a flat file from EDR to update our Blaise CATI applications
  - View of the flat file:

23234534;Profile.Legal_Name; My Technologies
23234534;Comments.EDR_1; Company Closed
23234534;CMonth.Questions.GOM;345235.23
23234534;CMonth.Questions.Total_Shipments;345235.23
12341234;Profile.Legal_Name; Your Technologies
12341234;Comments.EDR_1; Expanded locations by 2 this year
12341234;CMonth.Questions.GOM;125.23
12341234;CMonth.Questions.Total_Shipments;345.66

Receiving Data from the Internet

- E-Questionnaire System
  - Designed to make use of XML for it's simplicity, generality, and usability over the internet.
  - It improved the load time for updating the units
  - Debate over it's security of the file transfer process between the Internet and our CATI systems.

Experimenting with XML

- Example of an XML file (content):

```
<?xml version="1.0" encoding="UTF-8"?>
<Dataset><Datarecord>
<PriKey>232345345</PriKey>
<EDR1>1st_Comment</EDR1>
<EDR2>Middle_Comment</EDR2>
<Val1>201006</Val1>
<Val2>1234.45</Val2>
<Val3>356</Val3>
</Datarecord>
<Datarecord>
<PriKey>12341234</PriKey>
<Com1>Another_Comment</Com1>
<Com2>Comment_Again</Com2>
<Val1>112</Val1>
<Val2>3456</Val2>
<Val3>4</Val3>
</Datarecord></Dataset>
```

Conclusion

- Currently at Statistics Canada, we are reviewing Blaise 4.8 and it's improvements that it offers.
- Beginning stages of converting our Blaise applications.
- Make XML **the standard protocol** to communicate in a distributive environment.