

Advantages and disadvantages of migrating Blaise 2.5-based survey systems to Blaise III

Leif Bochis, Statistics Denmark

1. The use of Blaise at Statistics Denmark

At Statistics Denmark Blaise has been used since 1990. During these years at least 70 applications have been developed, including instruments for CADI, CAPI, CATI and CASI and ranging from input systems for small paper forms to complex questionnaires, such as the Labour Force Survey comprehending instruments for telephone interviewing, data input and data editing.

Most of the applications are CADI or CATI - one CAPI (conducted by interviewers from a market research institute) and two CASI-instruments used as supplement for two surveys where most of the data are collected from registers.

A central data input section was abandoned a couple of years ago, leaving the data input tasks either to be performed by out-of-house bureaus or to be put in and edited via systems developed at Statistics Denmark. A large number of these tasks are now put in and edited via Blaise instruments.

At Statistics Denmark all data input/editing and interviewing are carried out on diskless workstations with all programs, data and external files located on file servers. The main software configuration is based on Windows 3.1 upon which almost all of the basic software - e-mail, wordprocessing, spreadsheets and archives - are based. Thus, Blaise instruments are always run in Dos-sessions under Windows - except for the telephone interviewing.

2. Case : The Forms Input Systems for Hotel and Camping Site Bed-Nights

For the Hotel and Camping Site Bed-Night Surveys paper forms are collected each month from medium-sized or larger hotels and camping sites. These forms were typed in via a data input bureau and later validated in batch - like traditional forms processing.

The planning started in the late autumn of 1995 in order to modernize the data input systems into combined data input and validation. The goals were :

Faster error correction and publication (the old process was relatively efficient though - but a couple of days ..)

Cheaper data input. Since the closing of the central data input section at Statistics Denmark many data input tasks were performed by a data input service company. This traffic was - though fairly efficient - rather expensive, so it was quite a demand that we should make an in-house solution.

The data input system should not replace the tabulation programs, but on the other hand, be open to a later replacement, i.e. the solution should not exclude certain ways of replacing the tabulation programs when resources to redevelop showed up. A demand that is easily met by the fine data and metadata export facilities in either versions of Blaise.

3. The Blaise 2.5 Version

The data input system was developed in January 1996 using Blaise 2.5. The necessary administrative tasks (dunning letter draw outs etc.) were programmed in Manipula and packed into a Menu system programmed in Visual Basic. The data input system was ready for use when the first 1996-forms arrived in the beginning of February and the administrative programs were added in the period from February to April. In April the system was adapted for the Camp Site Forms.

The administrative tasks and utility programs included :

- Dunning letter draw outs,
- Lists for telephone dunning,
- Data output to the tabulation systems,
- Names and Adresses draw out for arbitrary searching purpose,
- File Browsing,
- Miscellaneous utilities such as lock removal and backup.

The first three were developed as quite simple Manipula setups, while choice of parameters was done using Visual Basic forms and passed over to the Manipula setups through BAT files.

The names and adresses were put into a file in order to search arbitrarily without knowing the key of the hotel or camping site - in lack of the more flexible lookup-facilities of Blaise III.

File browsing and lock removal utilities were developed entirely with few lines of Visual Basic using the Windows Notepad program as file browser for smaller files.

4. The Blaise III Version

Blaise III was already under consideration when development started, but due to some experiences of instability of the then existing version, lack of some promised - and promising - facilities and the developer's lack of experience with Blaise III we did not dare to base the system upon it.

Later, with the release of version 1.1j in the summer of 1996 which proved stable and comprised the promised facilities, the Blaise support started to reconsider the data input system - though mainly as a study project.

The experiences of this project were later used as a basis for developing a Blaise III instrument for Maritime Statistics Data Input and Validation. In the winter of 1996/97. This instrument is not quite finished when writing these lines, but as this project is only the second major Blaise III project at Statistics Denmark the experiences are closely connected to the former and have influenced also the conclusions of this paper.

The study project on Hotels and Camping Sites consisted of a number of steps :

Automatic conversion of the Blaise-instrument into Blaise III and the Manipula setups into Manipula 2.1 using 22III.EXE and MAN2III.EXE. Followed by the necessary changes in the code due to the new way certain things are implemented in Blaise III. This conversion took a couple of days and exposed a lot of the benefits of Blaise III - among them the ease of manipulating dates and the valuable secondary key concept introducing lookups through trigram search.

The Visual Basic program and the BAT files were then transformed into a Maniplus setup. This process lasted another four or five days - while still inexperienced - and served as a solid introduction into the way these things are done using Maniplus.

The result of this process was a program with the same functions as the version built in Visual Basic and Blaise 2.5 except for some minor differences - as, for example, the built - in possibility of printing files from the Notepad program.

The result was also, however, a program that performed considerably slower due to certain changes in the way rules are processed in Blaise III.

Conclusions

The case once again proved the good properties of Blaise 2.5, Manipula and Visual Basic as tools for rapid application development.

The study project showed that although a rather large investment should be made before starting real life Blaise III projects, Blaise III and Manipula/Maniplus have the same potential - except for the ease of visual development in VB.

The greater advantages of using Blaise III comprehend :

- one integrated tool for all: administrative tasks, data input and menu systems,
- advanced lookup facilities which made alternative search programs obsolete.

We also realized, however, that it takes quite an effort to implement Blaise III as a new standard software in the house. Many things are carried out in Blaise III in a quite different way than in Blaise 2.5, which makes the necessary education of experienced Blaise 2.5 developers into Blaise III developers quite a job. And besides, that a large amount of work is needed to get to the starting point of use - among the trivial but nevertheless time-consuming tasks are translation (when you are using a 'small' language as mothertongue) and integration with other tools such as making the necessary CIF-files for exporting data and metadata (unfortunately a STP2CIF program hasn't been delivered).

Among the disadvantages are the remarkable use of memory - 8 MB is still widespread for our workstations, and when the demands rise to 4-6 MB for the application - including DEP and Maniplus - plus the eventual demands of the external files, these workstations had to be put in front of the queue of upgradings (which the users approved, of course).

We experienced the problem of very slow performance of lookups due to the inability of Blaise III to use lists in memory. As it is stated in the README.DOC file of version 1.1j :

"The MEMORY attribute for external files will not be implemented. The need for this option has become very limited with the usage of disk caching systems."

When using diskless workstations it is - with the current versions of our operating systems - not possible to use disk caching. We have until now solved the problems by using RAM-disks for external data instead. This works and the performance is satisfactory, but we are not happy with the inflexibility the use of RAM-disks imposes on our configuration.

In contrast, the Visual Basic + Blaise 2.5 solution provides the necessary performance and lower usage of memory due to the ability of Visual Basic to share memory (common DLL's etc.) with other Windows applications, and the ability of Manipula 1.6 to perform tasks without the need for extended memory.

Another disadvantage is that users may be biased towards the old dos interface of Maniplus, if they are used to Windows shells built on top of the data input instruments. This will be slightly harder to 'sell', if it is not followed by remarkable gains in the integration of the Blaise III tools. These gains may not be as obvious for users as well as for developers. So for this reason, too, we are looking forward to the Windows version.

At Statistics Denmark we have started development in Blaise III at a low level. There is no doubt, however, we will make good use of the large benefits of this version in the coming years.

On the other hand, quite a number of Blaise 2.5 applications - mostly applications using (larger sized) lists - should be evaluated carefully before moving them to Blaise III.