A successful match between concepts and techniques in data collection: the Teler project

Hans R. Stol  
*Stol Advice inc.*  
Noordeinde 20-22  
2611 KH Delft, Netherlands  
hstoladv@wxs.nl

1. Introduction

One of the projects in the 4th R&D framework of the EC was TELER: TELematics in Enterprise Reporting. The main objective was demonstrating how to lower the administrative burden on enterprises by using information technology and telematics¹. There are several ways to lower the administrative burden on the enterprises. One of these ways is simplifying the administrative rules by the governmental institutes that causes the burden. The Teler project concentrated on the development of information technology solving the problems raised by the existing administrative rules. The figure below shows five alternatives for the enterprises to ease the burden caused by the governmental institutes.

---

Fig 1. The alternatives

The three indirect ways for the enterprises to lower the existing administrative burden are indicated as: using the professional organisation or the accounting firm as an intermediary, using an independent organisation for (parts of) the collection and sending of the data.

The direct ways are:
- the use of existing accounting software with an integrated module for the selection and dissemination of the requested data
- the use of an independent software module as intermediary between the automated information systems at the enterprise and the information systems at the data collector.

In the NSI part the latter option has been experienced. The software module, developed by Statistics Netherlands in co-operation with Cap Gemini, is called: EDISENT. In this article the emphasis will be on the NSI-part of the project. This experience was not only remarkable because of the success in a technical sense but also because of the co-operation between 8 countries. The

---

¹ The consortium consisted of a French company (Cesia), administrative manager of the project and co-ordinating the contributions of a professional organisation (Eurofer: steel industry) and the European representative of the accounting profession (Edificas), both with demonstrations in France. On behalf of Statistics Netherlands I co-ordinated the contributions of 8 NSI’s in the following countries of Europe: Sweden, Finland, Germany, the Netherlands, Italy, Spain, Portugal and Slovenia.
emphasis was not on the differences but on the common problems. This attitude and the use the adequate modelling techniques made this part of the project to successful.

2. The concept
The overall conceptual framework of the Teler project is visualised in the figure below. On the left are the respondents, being enterprises or intermediaries like accounting firms and professional organisations acting on behalf of the enterprises. On the right the data collectors of which the NSI is an example.

![Figure 2. The Teler-concept](image)

If the data of the enterprise go via an intermediary organisation to the data collector, a permanent storage of these data is recommended. Data required by the data collector can be selected in a flexible way out of the database.

In order to collect the data automatically from the automated systems one has to overcome the problem of mapping the metadata (or the meaning of the data stored in the automated information systems) of the respondent into the metadata (or the meaning of the data required by the data collectors) of the data collectors. The easiest way is to include in the (standard) accounting software a general code structure for the chart of accounts and template(s) representing the general used metadata by the data collectors. Using standardised transformation rules the administrative burden for data items that are automatically stored would be brought close to zero! Unfortunately - in general – nor the respondents’ metadata nor the data collectors’ metadata are standardised\(^2\).

Table 1 shows the situation in the countries involved in Teler. (N= no, Y= yes, P= partly)

<table>
<thead>
<tr>
<th>Standardised chart of accounts</th>
<th>FR</th>
<th>NL</th>
<th>IT</th>
<th>DE</th>
<th>SE</th>
<th>FI</th>
<th>PT</th>
<th>ES</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>De facto standard chart of accounts</td>
<td>-</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Standardised metadata at data collectors</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>P</td>
<td>P</td>
<td>N</td>
<td>N</td>
<td>P</td>
</tr>
</tbody>
</table>

3. The process
The main functions of the software module used in the NSI-part of TELER are:

- storage of the variables from the questionnaire of the data collector
- definition of extraction rules between data items of an output file of the information system at the respondent and the variables
- after the definition of the extraction rules between the data items and the variables (once and for all) data items are transformed into variables and the variables are sent to the data collector.

\(^2\) This means that the same terms might have different meanings, the chart of accounts of the financial information systems differ, the data asked for by governmental agencies are not harmonised, etc.
So before the software can be used, preparations should be made by the data collector (in fact the distributor of the module among the respondents). One of these preparations is the definition of the variables and accompanying metadata.

Knowing that the NSI’s involved in the demonstrations of the project should prepare these issues, knowing that the software was not ready yet (the concept was there but the development of the software did not even start), also knowing that participants who are only listeners and not active players are not really involved, I was confronted with the challenge how to launch the project so that from the start the participants could/would co-operate. One important issue is to assure that the participants experience the same problem and that it needs handling. This is already difficult within an (big) organisation. It is even more difficult in a project with 8 countries.

My solution was to design a model of the metadata involved in the intended software. In the first meeting of the NSI’s this model has been discussed. The next step was the translation of the model in an access database to be filled in by the NSI’s. Figure 3 shows the first data model implemented in the access database.

---

**Figure 3. The first (meta-)datamodel**

Various adaptations on the model itself and implementation of the model in the access database followed. Before the demonstration phase the content of the data base was converted to the EDISENT software and made ready for demonstrations at enterprises and accounting firms.

The use of the common model caused an important coherence between the participants in the project. The NSI-participants were concerned about the upward compatibility of their metadata in stead of complaining about delays in the delivery of the software. As soon as the software was ready for production (although it was meant to be a prototype) the participants were able to fill in their metadata immediately.

In the deliverables of the TELER-project steps in the process are described: from the investigation of the user requirements to the demonstrations.

4. The results and conclusions

In the demonstration the electronic questionnaire and the other metadata were filled in by the NSI’s. Besides that each NSI selected enterprises for the trials. In all the countries the enterprises were willing to participate and evaluated the results as promising, most of them wanted to continue immediately. After installation of the software an export file was created from the appropriate part of the enterprises’ information system. The next step was “linking” the statistical variables (from the questionnaire in EDISENT) with the metadata of the export file. After this step the EDISENT module was able to calculate the statistical variables out of an export file of a certain period. The last step was the sending of the data to the NSI. This process turned out to be very efficient for monthly and quarterly surveys.

---

3) In an multinational project also the language of the metadata of the software module itself has to be defined by the participants!
Table 2. Result of the demonstrations

A calculation of the time needed for installation, linking the metadata of the questionnaire with those of the export file of enterprises’ information system, the creation of the export file and the sending of the data showed a cost reduction of 85% compared with the manual fill in of the questionnaire. (In this calculation one of the preconditions is that each year 10% of the questions in the questionnaire change and once in 5 years a totally renewed questionnaire should be “linked” to the information systems of the enterprise.)

The enthusiasm about the result is huge. Most of the NSI’s want to continue using EDISENT in the next three years. The table shows the targets of the NSI’s involved.

Table 3. Target for the next 3 years

Future Research and Development will be concentrated on exploring:

- The possibilities to co-operate between governmental institutes, at first within in specific country (main objective: avoid more than once questioning the respondents, use each others data as much as possible). The problem of meta data is prominently there!
- The possibilities to involve the software suppliers and intermediaries. How to get more players in the market. The more players the less burden on the enterprises.
- Definition of criteria and organisational conditions for effective and efficient data collection by governmental institutes. Among these criteria the possibilities (relevant characteristics in the EU-countries), advantages (better control and more efficiency) and disadvantages (less flexibility, loss of advantage edge, less effectiveness) of standardisation will be taken into account.

REFERENCES

RESUME
Le projet TELER a démontré une réduction considérable de la charge administrative pour les entreprises a cause de EDI. Les Institutes Statistique de huit pays (NL, IT, DE, SE, FI, PT, ES, SI), coordonné par CBS, ont construit une module de software « EDISENT ». « EDISENT » traduit les contents des systèmes informatiques d’entreprise en les variables des questionnaire. En huit pays d’Europe la module de software « EDISENT » a démontré. Dans chaque pays la software a installé à les entreprises (5-15). Apres la traduction des metadata de la système informatique d’entreprise en les metadata de la questionnaire, les données a transmis à l’Institut Statistiques. Les résultats techniques et économiques sont très positives. Le coopération de les huit pays est une exemple pour projets Européens. L’usage pratiqued’une model de metadata pour le coordination des activités etait très effective.