I. INTRODUCTION

1. The introduction of electronic questionnaires for long-established surveys poses challenging questions for project planners.

2. On the one hand, designing an electronic questionnaire for such a situation is an exacting task. The design decisions taken in the run-up to the project itself, concerning such vaguely defined and/or all-encompassing subjects as “lessening the burden on respondents” or “security”, lead to measures which a future user may perceive and judge in widely varying ways, depending on his or her individual situation. The results of such decisions thus become crucial factors for the questionnaire’s range of application and user acceptance, influencing the degree to which the new reporting medium will be utilized. The users will not be motivated to use an electronic questionnaire if they see no concrete personal advantage in so doing.

3. Great attention must be paid to the respondents’ security concerns and to safeguarding their privacy. For example, enterprise management often feels strongly about internet security, establishing internal security rules and restrictions on browser use (e.g., no scripting, no cookies) in order to ensure that sensitive enterprise data cannot be spied on. Such clients will not accept internet questionnaires unless they not only trust the institution carrying out the survey but also recognize that the system design of the web questionnaire itself does not counteract their security measures.

4. Secondly, the financial backers of an electronic questionnaire project attach certain expectations to their use. As a result of their investment, they anticipate savings on personnel expenses and/or improvements in processing time and quality. In general, however, the electronic questionnaire will not succeed in wholly replacing the paper version, but will form an additional data collection track that must be serviced.

5. One aspect, which at first sight offers considerable potential for economies, is the shifting of some responsibility for data editing towards the respondent. This only makes sense to the degree that it does not violate the above-mentioned principle of lessening the respondent’s burden. Moreover, such a relocation also calls for changes to the national statistical institute’s internal data editing procedures. Only...
a strategy which succeeds in making optimal use of the respective reporting tracks’ advantages on the one hand, but on the other also consolidates the data flows from the various reporting media for combined processing at the earliest possible moment, will result in actual savings. This necessitates additional investment in software and alterations in workflows and administration.

6. Thirdly, not least as a consequence of such an intervention in established processing and production systems, the employees entrusted with them – be they subject matter or IT specialists – will have to alter procedures which they have developed over time and which have hitherto proved of value. They must carry out new tasks for which they will also need new skills. In this area, aside from concrete training and other measures, top management’s explicit commitment will be required in order to overcome possible resistance.

7. This paper reports on our experiences while introducing electronic data reporting (EDR) in Statistics Austria, using the monthly Short Term Survey (Production) as an example. For this survey, two parallel electronic reporting tracks are used, in addition to the original paper questionnaire. We describe the initial situation in 1998/1999 and give an overview of the fundamental design decisions taken at that time. These have served as the basis for a general strategy in the areas of data collection and editing. Our most important products, and the design principles we used in their implementation, are presented in brief. Furthermore, we will demonstrate to what extent and in which ways the monthly procedures for processing the collected data have changed in consequence, as compared to the “paper-only” era.

8. In 1996, 50 employees were entrusted with processing the Short Term Survey from the subject matter point of view. Nowadays they number only 30, have been equipped with PCs and trained in the use of the new products. All these employees process both paper and electronic reports.

II. INITIAL SITUATION / THREE PRIMARY GOALS

9. In 1998, under heavy pressure from the Chamber of Commerce and Industry, the media and the political establishment, the then so-called Austrian Statistics Central Office decided to accommodate the wishes of business sector respondents and to develop an integrated application for statistical questionnaires. Specifically, the monthly Short Term Survey, the annual Structural Business Statistics Survey and the annual Production Input Survey were to be supported. The primary goal was to alleviate the burden on the respondents.

10. At that time, the Austrian Statistics Central Office found itself in a phase of radical change. It was to be separated from Government Services, reorganized and reconstituted as an independent body – the current Statistics Austria, a federal institution under public law. The questionnaire project appeared a suitable contribution to this reorientation, offering benefits for both respondents and the institution. From the start, however, it was clear that an electronic data reporting alternative for the aforementioned economic surveys could never fully replace the paper track, which made early integration of the various reporting tracks indispensable. The second goal of the project was thus defined as streamlining and accelerating the processing of survey data, hand in hand with enhancement of data quality.

11. In view of the substantial investment and convinced that electronic questionnaires will be the future standard for primary data collection, and bearing in mind that National Statistical Institutes (NSIs) must carry out a multitude of surveys, the third goal was to provide a general solution. This should allow future electronic questionnaires to be created quickly, flexibly and at minimal cost. Furthermore, it should include all software systems needed to process the electronically collected data within the framework of a comprehensive general strategy, in a standardized and efficient manner.

12. The first deployment of the electronic questionnaire was the Short Term Survey “Production”. Due to its complexity and its monthly reporting cycle, this survey was then and remains today the one that places the greatest burden on its respondents.
III. THE SHORT TERM SURVEY “PRODUCTION” IN STATISTICS AUSTRIA

13. The Short Term Survey “Production” delivers the foundation for complying with the EU regulation on short-term economic statistics. In addition, it allows for the calculation of the national production of goods. It is one of the essential bases for observing the economic cycle, thus providing the data for the economic decisions taken by government and administration on the national and European level and the European Central Bank.

14. The Short Term Survey “Production” includes all branches of industry in the NACE sections 10-45 (Mining and Quarrying, Manufacturing, Electricity, Gas and Water Supply and Construction). This survey has existed for a very long time and since 1996 has been based on a concentrated sample, the data being reported directly by the respondents. Observation units are enterprises with one or more establishments and the establishments themselves. To fulfil their reporting duties, the respondents may engage the services of external third parties such as tax consultants. The classifications NACE and PRODCOM are utilized in a version adapted to Austrian requirements.

15. Approximately 12,000 questionnaires are sent out monthly. Of these, 10,500 are directed to enterprises with one establishment, 400 to multi-establishment enterprises and about 1,200 to the latter’s establishments. The respondent’s address and some characteristics are obtained from Statistics Austria’s enterprise register and pre-filled in.

16. In total, ten different variants of the questionnaire are used, depending on the enterprise’s size and type. The questionnaire for single-unit enterprises contains about six pages with on average 90 questions, grouped into blocks by topic (employees, hours worked, salaries, orders, turnover and production). Multi-establishment enterprises must fill in three pages for the enterprise as a whole (employees, salaries, turnover) and six pages per establishment. About 25 pages of explanations and clarifications accompany the paper questionnaires. Many logical dependencies exist between questions, both within the same block and between blocks (e.g., if some employees only work part-time, the hours worked must be stated separately for full- and part-time employees).

Survey data processing up to the year 2000

17. **Incoming check and register:** Incoming (paper) questionnaires were first registered and submitted to a rough check in order to verify data completeness. Information relevant to the enterprise register – change of name, address etc. – was verified and corrections applied to the register (not shown in Figure 1) manually.

18. **Data validation:** The registered questionnaires were then validated manually, including comparing them to the reported data of the previous month and/or year. The checks included:

   a. Visual examination of the question groups in the questionnaire by the subject matter expert.

   b. Comparison of the current reported data with the previous month or year: visual check by laying the relevant paper forms side by side.

   c. Quotas (e.g. paid hours per employee) were calculated by the subject matter expert and written on the paper form; comparison with the previous month.

   d. In case of implausible answers the respondent was contacted and the results noted on the paper form.

---

2 The following description is intended to clarify the tasks involved. For easier understanding, many of the survey’s aspects have been simplified.
e. Missing data was added and corrections applied in red pencil; „old” values struck out „legibly” in order to preserve the original data and the modification history.

f. For control purposes, some question blocks were augmented by tallies of entered data fields. This enabled cross-checks on errors occurring during the data entry phase.

19. **Classification codes** were inserted in the question group Production (PRODCOM codes and units of measure) in those cases where the respondents had not done that already. The respondents received printed code lists in which they could search for their product and determine the correct code.

20. **Input:** The data were then recorded by input personnel and imported into the mainframe DB2 database.

21. **Corrections and additions** to the data were written on the paper form in red pencil by the subject matter expert, as was further information on data fields, blocks or the whole questionnaire. Corrections of the DB2 data were carried out in a mainframe terminal application by using the printout of the error list and the original paper questionnaires.

22. **Programmed detailed checks and corrections:** The data entered in the DB2 database was validated by a program containing about 100 separate consistency checks on single data values and their mutual relationships. This application also drew on the previous month’s data and calculated quotas. Print-outs of the error lists were provided to the subject matter experts who then corrected the values on the micro data level in parallel on the filed paper questionnaires and in the database.

![Figure 1: Traditional survey processing](image-url)
IV. FUNDAMENTAL DECISIONS AND THEIR IMPLEMENTATION

A single technology will not suffice

23. The discussions on the burdens carried by respondents, undertaken at the project’s commencement with the active input of business representatives who articulated their ideas and wishes, soon led to the realization that one technology alone could not satisfy all requirements. The various forms of electronic data collection exhibit specific advantages and disadvantages that must influence the choice of medium for a survey. Which track will be most suitable from the point of view of unburdening the respondents depends on the survey’s volume and complexity.

24. In economic statistics, for example, the effort needed to respond adequately varies greatly even within the same survey. A large enterprise with multiple establishments has to answer a much larger volume of questions – and in greater depth – than does a small one. The respondents’ demands with regard to electronic support and ease of use thus diverge enormously. A web form may completely fill all the requirements of small firms while medium-size and large enterprises demand enhanced support functions and – crucially – satisfactory performance even with large amounts of data, if the electronic track is to be perceived as useful and not an additional burden.

25. For reasons of cost and available resources, only one solution could be implemented initially. Thus, taking into account that medium-size and large enterprises carry the heaviest burden and articulate their wishes in greatest detail, we developed “e-Quest, the electronic questionnaire manager” first. This is a Windows 32 bit application implemented in Visual Basic 6. In the year 2003, we initiated the web form system “e-Quest/Web” as an additional option for smaller businesses.

26. What do both systems have in common?

- High security of input data;
- No protracted searches in voluminous production code lists and other classifications, but rather quick retrieval of the correct code, aided by display of the codes in hierarchies, full text and synonym search, and automatic insertion of codes, texts and units of measure in the questionnaire;
- An extensive multi-layered help system offers context-sensitive help only when and where it is needed;
- Automatic input checks mark incorrect or implausible data with warnings and/or error texts;
- Automatic calculations (sums, comparisons, etc.);
- Questions can be displayed or hidden dynamically, depending on their relevance in the specific case;
- Form completion can be interrupted at any time and the data saved for completion at a later time or date;
- Third-party respondents have access to their clients’ forms and functions for managing them.

27. The PC solution e-Quest is especially suitable for multi-establishment enterprises and includes further functions, e.g.:

- a standardized data interface (XML) for connection to enterprise IT systems;
- an “auto-import” function with which questions for which the answers are constant over protracted periods can be pre-filled in;
- a copying function by which data in a designated questionnaire can be copied to the current one;
- stand-alone or client-server architecture (within the enterprise);
extensive options for individual user management (password-protected level of access to specific questionnaires and/or system functions), especially suitable for client-server installations or third-party respondents.

“IT experts should only be used when they are really needed”

28. Many IT solutions not only incur one-off development expenses but also imply considerable further effort for continuous adaptations to new or modified requirements. In many cases, such changes are part of the work environment and are by no means unexpected – e.g., recurrent statistical surveys are subject to periodic revisions of their content, so that some questions may vanish while others are added. Even so, system design often neglects such foreseeable alterations, necessitating the services of expensive IT personnel.

29. In the IT field, many steps in the process only require specialist programming skills because suitable tools for use by “normal” users do not (yet) exist.

30. In e-Quest and e-Quest/Web, we placed great emphasis on these considerations and designed systems controlled to a considerable extent by metadata that can be actively supplied by subject matter experts.

31. To this end, we developed a suitable tool, the e-Quest Metadata Manager, to support the process of survey preparation. Using this application, subject matter experts can design survey questionnaires simply by dragging and dropping visual objects, and can specify field and group definitions, control flow and validations. This information is then converted automatically to a standardised XML format that controls e-Quest and from which e-Quest/Web forms are generated.

From piecework towards automatization

32. Most NSIs carry out a great number of primary surveys for which electronic forms are eminently suitable. To lessen the burden on respondents, the electronic reporting option should be offered in all such cases.

33. In the face of shortening production cycles and shrinking budgets, this demand can no longer be fulfilled in traditional fashion, namely by manual software production. Not only do the expenses of developing individual software add to production costs for the paper questionnaires for each individual survey, but providing a timely electronic alternative for new or modified surveys is an almost impossible task. Even when software development is done in-house, thus avoiding delays due to public calls for tenders and subsequent evaluation periods, developing a user-friendly, secure and thoroughly tested electronic alternative takes much longer than the usual production process for paper forms.

34. We therefore first conceived e-Quest as a generic application and subsequently built an “industrial production line” for the e-Quest/Web forms which allows a complete internet web form application to be created by simply “pressing the button”.

Benefits for the data collector

35. For survey data processing, only then may one expect significant savings through electronic questionnaires when the relationship of electronic to paper-based responses is relatively high and/or the new data track triggers a climate of change in the traditional methods of processing. An infrastructure encompassing all phases of data collection is needed, and the new processes must be integrated into existing processing systems. As a rule, the latter have developed over time and cannot be simply thrown overboard.

---

36. In parallel with the aforementioned e-Quest Metadata Manager for survey preparation and the questionnaire systems e-Quest and e-Quest/Web on the respondents’ side, Statistics Austria also developed standardized software systems to accept and process the raw data. Incoming e-Quest/Web questionnaires can be directly viewed and edited in the intranet by subject matter experts using the same browser application developed for the respondents – albeit enhanced by specific internal functions. Alternatively, the data can also be transmitted to the e-Quest Package Manager (internally nicknamed the “Pot Application”) created for e-Quest responses. This is what is done for the Short Term Survey.

37. The “Pot Application” is a standardized tool for processing electronic responses. Incoming questionnaires (in the form of XML packages) are automatically distributed into various processing “pots” according to their content, using criteria defined in advance. In the “pots” they can be viewed, edited, corrected and passed on to other subject matter experts for further processing. For viewing the data, one can use an enhanced version of e-Quest that keeps track of any changes to the data, maintaining a history of modifications and comments attributable to the (internal) user who made them. Questionnaires that have been checked are moved to the “OK pot” and are then transferred via an interface component\(^4\) to other systems.

\[\text{Figure 2: Integration}\]

38. An efficient integration of electronic questionnaires into the existing processing flows was achieved with the PC correction application „KjeKorr“. This is a program that allows data from the monthly Short Term Survey „Production“ already residing in the host database to be viewed, augmented, validated and corrected. It too fulfils all Statistics Austria’s requirements for extensive documentation and modification history, which earlier was contained in the manual notes on the paper forms. Without having to take recourse to the paper questionnaires, the subject matter expert can now process the survey data quickly and efficiently and can rely on automatic data editing support:

- automatic calculations (sums, quotas, entry counters, units of measure);
- online search in the classification hierarchies (PRODCOM, NACE);
- modification history, comments and validation checks on the currently active field are highlighted;
- questionnaires that have been „set aside“ or deleted can be reactivated;

\(^4\) The interface component is programmed for the specific survey.

- automatic data validation when the data are stored.

**Respondent-side editing**

39. One aspect, which at first glance should improve the quality of data and thus lead to higher efficiency, is the shifting of responsibility for data editing to the respondent. It must however be noted that this goal does to some extent contradict the expressly formulated principle of “lightening the burden”. Errors the respondent is alerted to while entering data result in an immediate increase in time spent reading the error, determining what is wrong and correcting it (and the time spent earlier in reacting to enquiries from Statistics Austria about erroneous data is not visible at that moment). This area was therefore considered carefully during the design phase. The subsequent decisions and their results are described in detail in the following.

40. **Single field checks**: The validation of single data values is controlled wholly by the data field attributes specified during questionnaire definition, such as data type, length restrictions and arithmetic limits. The checks can be classified according to their urgency and importance.

   **Urgency:**
   - check immediately after the data value was entered,
   - check when the user says so,
   - check when the whole questionnaire is being validated

   **Importance:**
   - must be corrected in all cases,
   - important warning which may however in some cases be ignored,
   - simple warning,
   - comment for the subject matter expert

41. In the case of e-Quest/Web “immediately to be checked” validations are only really checked right away if the user has enabled Javascript in the internet browser. If not, they are carried out on the Web server at the time of the next server interaction and then displayed complete with error messages and markings to the user for correction.

42. **Inter-field validations** check rules describing the relationships of field values to each other. Thus they always result from a separately specified condition. The electronic version of the Short Term Survey includes about 50 such validation rules, i.e. about half of those checked during subsequent processing. Very few of these are classified as “must be corrected” (which prohibits sending the data before such errors are corrected). In e-Quest, a further error level has been introduced in which the respondent can insist on a value marked with a severe error if he or she adds a comment as to why the value is correct in this special case.

43. **Inter-questionnaire validations** in the e-Quest system check relationships between data values in more than one logically connected questionnaires, such as questionnaires for an enterprise and its separate establishments. In practice they are only used for rough checks, as the respondent should not be forced to open and correct multiple questionnaires concurrently, although e-Quest can display them in separate windows. As a Web browser application is not really suitable for displaying and manipulating concurrent forms, e-Quest/Web forgoes such checks.

44. **Abstain from checks wherever feasible**: as correcting errors causes increased effort for the respondents, type, level and amount of validations must be weighed carefully. Some errors may be easily corrected automatically or by subject matter experts and thus could be ignored in the data input phase. From the respondent’s point of view, the questionnaire will then be easier to use.\(^5\)

---

\(^5\) We must not forget that in our case, the respondents are required by law to provide the information but are not necessarily pleased to have to do so.
Security aspects

45. **Risk avoidance**: One question discussed extensively during the project was whether one should provide the respondents with data from the previous month as initial values or for the purposes of checking and correcting the current data.

46. In the case of *e-Quest*, the data are stored locally in a relational database and can be accessed by the respondent through the application. Functions are provided for, e.g., copying the contents of a whole questionnaire into a new month and then only editing the changed values.

47. To implement such facilities in the case of Web forms, the previously reported data of all enterprises in the sample would have to be stored and kept accessible via the internet for the duration of the survey, which appeared to us to carry unacceptable risks. The potential damage a successful hacker could do in such a situation far outweighs the benefits.

48. We therefore adopted a design which aims to expose the data for as short a time as possible to a “hostile environment”. In *e-Quest*, the data are stored locally and transmitted by the user in strongly encrypted form via e-mail or FTP to Statistics Austria. Data from Web forms – transmitted via secure connection – are stored in a Statistics Austria database but transferred from there to another database inaccessible from any Web application at the earliest moment.

49. **Security Audit**: An important feature of *e-Quest/Web* is its global security architecture. Together with an external software partner, we created a design that focuses on secure and confidential communication with the user over the public and a priori untrustworthy internet. The completed design was analysed and validated by external experts with regard to mechanisms and application flows relevant to security. This method of procedure was intended to minimize the risk of design defects and also to provide confidence of Statistics Austria and the public in its Web questionnaire system. The results of the expertise were worked into the design before commencement of the implementation phase.

50. **Taking the respondents’ security concerns seriously**: Security concerns worry many enterprises, particularly when making use of the internet. Internal security rules and restrictive browser settings (e.g., no scripting, no cookies) aim to ensure that sensitive company data cannot be spied upon. We believe that these clients will only use Web forms if they firstly trust the collector of statistical information and secondly do not perceive electronic data reporting as corrupting their internal security measures.

51. For *e-Quest/Web* we therefore took the decision not to base the application on the use of cookies. Nor do the validations require Javascript to be enabled, although if it is, they will be carried out immediately upon data input instead of at the next server interaction.

52. **Gaining trust**: With respect to protection of their privacy, users often fear that internet applications will spy on their PC, the software they use, or their “surfing” behaviour, and then analyse or use that information in unknown ways. In order to strengthen respondents’ confidence in this area, we guarantee in writing that Statistics Austria only routes such information onwards for processing which they have confirmed by pressing the “Send” button, and that no other information of any kind is kept.

53. If the user has temporarily stored some data on the server in the course of filling in the form, this data is deleted automatically when the final version is confirmed. Such data is also deleted after a

---

6 This section does not treat general measures to ensure data security and confidentiality, but is only concerned with those aspects that seemed of special importance to the authors during the project’s implementation.

7 Albeit only accessible through the Web application

8 According to a Cap Gemini study, large companies in Germany list computer data security as the most urgent concern in their IT strategies. 83 percent of 152 CIOs from companies with a yearly turnover of more than 250 million Euro taking part in the study “IT trends 2005” mentioned this as highest priority.

9 Except for a non-permanent session cookie that the Web server requires in order to associate incoming requests with the correct session.
specified time interval if he or she afterwards decides against transmitting the data electronically. The temporary data on the server can also be deleted explicitly by the user pressing the “Reset” button.

IV. THE SHORT TERM SURVEY “PRODUCTION” TODAY

54. At present, about 2,600 (22%) of the 12,000 monthly questionnaires are reported via e-Quest. As can be seen from the table below, e-Quest use correlates with the size of the enterprise. While in June 2004 only 12.3% of companies with up to 20 employees made use of this reporting track, the quota for the largest enterprises was 46.6%.

<table>
<thead>
<tr>
<th>Employee number class</th>
<th>Paper</th>
<th>e-Quest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...% of the companies in this size class report on paper</td>
<td>...% of the companies in (2) employ ...% of total employees</td>
<td>turnover of the companies in (2) is ...% of total turnover</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>up to 20</td>
<td>87,7%</td>
<td>87,4%</td>
</tr>
<tr>
<td>21 - 99</td>
<td>79,4%</td>
<td>77,3%</td>
</tr>
<tr>
<td>100 - 499</td>
<td>65,0%</td>
<td>62,6%</td>
</tr>
<tr>
<td>500 and above</td>
<td>53,4%</td>
<td>49,9%</td>
</tr>
<tr>
<td>together</td>
<td>80,7%</td>
<td>64,7%</td>
</tr>
</tbody>
</table>

Table 1: Reports by enterprise employee number class / June 2004

55. Starting with reporting month Jan 2005, single-unit enterprises have been offered e-Quest/Web as a new alternative, though without explicit public relations measures, except for a letter included with the paper forms. Of the questionnaires received until March 8th (altogether about 7,000), 9% made use of this new questionnaire, as against 69% on paper and 22% via e-Quest. The number of Web respondents is thus roughly equal to the number of e-Quest respondents when that application was first introduced.

56. A first reporting stream analysis, based only on the questionnaires received up to March 8th and comparing them to Oct 2004, shows that of the 1753 e-Quest users in this sample in Oct 2004, 299 have moved to the Web questionnaire. Of the 5194 paper users in Oct 2004, 292 used the Web form this time and 88 e-Quest. Of 64 users new to the survey, 8 used e-Quest and 8 the Web track. These numbers offer only a first impression, but they confirm our assumption that the respondents make deliberate decisions as to what relative importance they attach to ease of use vs. security concerns.

V. SUMMARY

57. In 1996, 50 employees were entrusted with processing the Short Term Survey from the subject matter point of view. Nowadays they number only 30, have been equipped with PCs and trained in the use of the new products. All these employees process both paper and electronic reports.

58. At the start of the project, many of our colleagues – in both enterprise statistics and IT departments – questioned the sense of this undertaking and showed considerable resistance. On the one hand, they doubted that an electronic solution for such a complex and voluminous survey would be accepted by respondents. On the other, many feared a deterioration in the work environment. It seemed inconceivable that a workflow which required several paper questionnaires to be laid side by side in order to effect the necessary comparisons and corrections, could ever be represented on a comparatively small screen.
59. Both fears proved unfounded. In the second case, the correction application “KjeKorr” for the first time in Statistics Austria shifted data editing from the mainframe to the PC level. It boasts a graphical user interface and its functions were oriented directly to the users’ needs and wishes. Consequently, this new processing tool has found wide acceptance.

60. Besides being used for the Short Term Survey “Production”, e-Quest has supported the Short Term Survey “Services”, the Structural Business Survey, the Production Input Survey and the Work Costs Survey for several years. e-Quest/Web has already been used in additional surveys such as the Livestock Survey and an ad hoc questionnaire for civil registry offices in the few weeks it has been in production. Further e-Quest/Web questionnaires are under construction. The speed and relatively minor effort with which these questionnaires can now be produced prove that the concept of e-Quest and e-Quest/Web was correct and made economic sense.