FINAL ASSESSMENT REPORT FOR

THE EUREDIT PROJECT

May 2003

Svein Nordbotten, EUREDIT Internal Evaluator

Date: 30. May 2003

Contents

0. Summary

1. Conditions of contract NT-1767/00

2. Content and sources for this report

3. Assessment of project work

4. Assessment of project deliverables

5. General conclusion and final remarks

6. Appendices

0. Summary

The main objective of the Euredit project was to investigate and evaluate methods for automatic editing and imputation. The project had 12 participants and was funded as a shared cost project supported by EU for the period 1. March 2000 to 28.February 2003, and estimated to include 499 person-months. In retrospection, an objective including also some non-automatic approaches may have provided even more useful results from the project.

The project was organized in 9 Workpackages with 41 deliverables of which 10 were public. The deliverables were submitted as planned with a few exceptions.

The project met for 7 project meetings well prepared by the project coordinator. Most preliminary deliverable draft reports were scheduled for the 24th and 36th months. If some of these deliverables had been scheduled earlier, deadline pressures in the last project months might have been avoided. Another lesson learned was that deadlines immediately before project meetings did not give partners required reading time for returning feedbacks to their colleagues.

The academic quality of the deliverables was in general high and the project must be regarded as very successful. Some end users interested in editing and imputation may not find the deliverable D6.2 (planned as 'best practice guide for edit/imputation') as ready for practical application as expected.

The general conclusion is that the EUREDIT project was a very productive and successful project. The leadership and management has been stimulating, professional cooperation good, the scientific quality of the results high, the presentation to the statistical community comprehensive and interesting, and the prospects for dissemination and exploitation seem good in spite of uncertain funding.

1. Conditions of contract NT-1767/00

According to contract NT-1767/00 between ONS and Svein Nordbotten, I shall deliver a final assessment report on the quality of the entire EUREDIT project by 31 May 2003.

This report was delivered to the Project Coordinator 30 May 2003.

2. Content and sources for this report

This final assessment is presented in 2 parts. In the first part, section 3, the work performance during the project period is assessed, while in the second part, section 4, and the project deliverables are discussed and assessed.

The evaluator has expressed opinions denoted *Comments*. These are not meant to be criticisms, but rather points to bear in mind for similar future projects.

The report is based on:

- Participation in 6 project meetings and 1 editorial meeting,
- EUREDIT project description as in *Annex 1 – Description of Work*,
- EUREDIT Website, papers, email, Work Plans, internal and external Deliverables by project partners,
- Editorial Committee meeting in Neuchatel January 2003,
- 5 previous Interim Assessment Reports,
- Other relevant materials made available to me.

3. Assessment of project work

General

The EUREDIT project structure comprised 9 Work Packages some of which were subdivided into subpackages described in detail in Description of Work in ***Appendix 2***:

- WP1: Project management
- WP2: Selection and compilation of datasets for evaluation methods
- WP3: Determining objective quality criteria for evaluating methods
- WP4: Develop and test selected new methods for error localization
- WP5: Develop and test new methods for imputation
- WP6: Evaluation and validation of results of WP4 and WP5 to determine the 'best' methods
- WP7: Integration of the individual edit and imputation methods into a single package for wider dissemination
- WP8: Dissemination and exploitation
- WP9: Project evaluation

Eleven partners participated in the project with Office for National Statistics, UK, as coordinator. The participants are listed in ***Appendix 1***.

Funding for 500 person-months was provided, of which 84, 60 and 51 were allocated to the 3 partners ONS, University of Southampton and Statistics Denmark. The workpackages WP4, WP5 and WP6 were assigned 371 person-months or about 75% of the resources.

Seven Project and Management Committee meetings have been proficiently arranged and run without problems in spite of some changes in the staff of the Project Coordinator.

Comment: The distinction between WP4 and WP5 has been difficult to maintain because some methods were used both in editing and in imputation. Subdivision reflecting the different methods might have resulted in a more 'clean' project structure.

WP1: Project management

The main activities specified for this work package were to organize cooperation among the partners by establishing standards, organizing meetings, monitoring progress and preparing project progress reports to the Commission. The central work has been done by the Project Coordinator and his management staff, which has carried out its tasks in a very professional manner.

Experience from this project, shared with many other project managements, indicates that it is difficult to make certain that deliverables are produced in time. Unfortunately, this problem had

not consequences only for the management and the delaying partner himself, but also for the project as a whole.

Deliverables were in general late for project meetings. The impression was that few partners had time for reviewing in detail papers from other partners. The peer reviews expected at the project meetings were frequently reduced to rather arbitrary comments across the table.

Comment: The benefits from peer reviews by other partners are lost when papers are delivered as late as the day before they should be discussed.

A satisfactory basis for recording activities within different workpackages and a standard for progress reporting, were established as well as procedures for recording the time and resources spent on different tasks.

The Consortium contract was discussed at several meeting, partly because of changes in the organizational status of one partner.

A group of peer reviewers for commenting on research papers was established with members outside the project. The group should contribute with constructive comments to drafts, and add prestige to the deliverables of EUREDIT. Because of the late deliveries of drafts close to deadlines, it became difficult to take advantage of the group.

The Project Coordinator and his staff prepared detailed and useful Minutes for each project and PCC meeting, Management reports and SPPRs for EUROSTAT, copies of which kept the partners well informed about the general progress of the project.

The number of documents prepared, submitted or relevant for discussions at each meeting was large and created also technical problems. Since project documents in general were distributed electronically, matching file names and document names became difficult. The lack of proper document dating made identification of the latest document version another problem.

Comment: With as many deliverables as specified for EUREDIT, and partners trying to compare their own work with others, more strict reporting instructions are needed to enforce timely and a standard form of report presentation.

When it became obvious that the project would not get any extra funds from EU for organizing conferences about editing and imputation, efforts were made to find other alternatives. The efforts were successful in Finland and the Finish partner in Jyväskylä did an excellent job to arrange a conference, DataClean 2002, in May 2002.

WP2: Selection and compilation of datasets for evaluation methods

Workpackage 2 included the very important task to compile representative data sets to be used in the editing and imputation experiments.

The WP 2 objective was to provide the data sets needed for the experiments and comparisons of methods. The provided datasets selected and obtained with a 'clean' version were:

- ABI (ONS),
- SARS (ONS),
- LFS (DSt),
- EPE (SFSO),
- Time Series Data (Quantaris)
- GSOEP (GSB).

Full names and sources are given in ***Appendix 2***.

From each provider, 2 'specific' data sets were obtained, one for method development and testing by the partners, and one for the final testing and subsequent comparison by the Project Coordinator. These selected sets were supposed to reflect problems with different types of data processed in a NSI.

For each set, several versions were required. The target versions were data sets, which had been edited by the providers' experts and considered 'true/clean'. An error algorithm was developed, and 2 additional 'raw' versions were generated, one generated version with missing values only, and a second version with both missing values and errors. For those working with methods requiring training/estimation of parameters, access was also supplied to small samples from the evaluation sets including both target and raw versions.

It should be noted that the target versions of the data sets are not more 'true/clean' than the providers' editors made them. Errors still remaining in the target sets may have had a significant impact on the ranking of the methods. From this point of view, synthetic target data sets with associated versions might have served the purpose of standard data sets as well, and would have reduced the formatting and distribution problems created by real data sets from various providers.

Plans for the EUREDIT project description state that "*The finally selected data sets will be stored on a CD-ROM, … ultimately for wider dissemination (subject to confidentiality constraints).*" The EUREDIT standard data sets should be an important contribution from the project to future research in editing and imputation as a benchmark. Unfortunately, because of confidentiality restrictions, the distribution of the data sets used will be limited to only some of the sets used.

Comment: Since some of the datasets were subjected to confidentiality constraints, the publicly available data sets will be limited. Referring to the alternative synthetic data sets, these would not have created any distributional restrictions. A further discussion of real and synthetic 'target' data for experimentation might have been useful. Synthetic data sets with distributional characteristics corresponding to real population data could have been an important supplement to the data sets used, and would not have caused any confidentiality problems.

In general, selecting datasets for use in editing/imputation experiments raises a number of more general questions within the scope of the EUREDIT project, which time did not permit being discussed in detail.

Comment: Examples of questions associated with this aspect are:

- ***Which are the specific criteria for selecting representative data sets?***
- ***How correct (true) are the target data sets compiled?***
- ***Which error types should be introduced by error generation and why?***
- ***How well do the error algorithms used reflect the real error mechanisms?***
- ***Should error types be introduced in separated data sets or mixed?***
- ***How important is the availability of the data sets for future researchers?***

WP3: Determining objective quality criteria for evaluating methods

Drafts of a technical report for describing the statistical quality in edit and imputation methods were circulated and discussed in several versions. The deliverable *Evaluation Criteria for Statistical Editing and Imputation* has served as a basic tool for evaluating and comparing performances of different methods developed/investigated.

The proposal for the EUREDIT project aimed at methods for automatic editing and imputation. Focus concentrated on the 'accuracy' dimension of statistical quality contributed by the editing/imputation process even though other dimensions such as processing time can be of importance.

Comment: A wider investigation not limited to automatic editing and imputation methods only and with extended attention to the impact of different methods on error perception, timeliness, resource consumption, and co-functioning with other processes would have been both interesting and useful.

This workpackage should also provide *training guidelines* for improving and evaluating the performance of editing and imputation processes. Even though the term 'training guidelines' can be difficult to interpret in the present context, some guidelines were drafted and discussed.

WP4: Develop and test selected new methods for error localization

WP4 was subdivided into 5 workpackages. For the complete specification of the structure, see ***Appendix 3***.

WP4 had 2 general goals:

1. Select, adapt, apply and evaluate editing methods currently in use,
2. Develop, apply and evaluate new editing methods.

Application and evaluation of all methods are supposed to be performed on the data sets selected.

The work of WP4 divides in 2 categories according to the objectives. In the *first category*, work with methods currently used includes adapting, experimenting and evaluating editing methods in use or related to methods in use. The main work in this category has been on AGGIES, CANCEIS (NIM), and GEIS applied on SARS and ABI data sets, respectively.

Comment: Even though some methods were developed outside EUREDIT, the work done by the responsible partners to adapt and implement the methods for use in the EUREDIT has been significant and should not be underestimated.

The *second category* of methods for editing included work on new methods for error localization. A number of methods not yet used for statistical data editing were developed or adapted, applied and tested. They included a range of methods for error detection/localization based on outlier detection, multi-layer neural networks (MLP), correlation matrix memory networks (CMM), and self-organizing maps (SOM), all documented in a number of algorithmic descriptions and reports on application. The new methods have been applied to the selected datasets and to a limited selection of variables.

Some of the new methods investigated have been developed outside EUREDIT, or by partners in connection with other tasks. The EUREDIT task in connection with these methods has been adaptation and application for statistical error detection.

Comment: It is still much work needed before some of the new methods can be included among the practical tools for NSIs.

Currently used editing methods usually rely on edits pre-specified explicitly by experts. Some of the experiments in the first category relied on edits specified previously by the experts who were responsible for the original editing of the target data sets.

In EUREDIT, several partners have done interesting studies investigating the possibility to compute implicit edits (e.g. neural network training and WAID) derived from work already done by experts.

In some experiments, the testing of methods was for example simplified by limiting the variables to subsets of variables from the original datasets, and, in other cases, to subsets of the original edits.

Comment: The use of pre-specified edits creates a special problem because the edits were selected in a situation, which was different from the experimental, and selection of edits more appropriate for the experimental situation might have produced other results.

WP5: Develop and test selected new methods for imputation

WP5 was divided into 7 smaller workpackages. The overall general objectives were:

1. Select, adapt, apply and evaluate imputation methods currently in use,
2. Develop, apply and evaluate new imputation methods.

Application and evaluation of all methods were also in this workpackage supposed to be on as many as possible of the selected data sets. In the project description, it was anticipated that *"EUREDIT will develop and evaluate indices for the purpose of indicating the extent to which an imputed value can be trusted."*.

The evaluation procedure developed as part of the task of WP3 measured the accuracy performance of the different methods when the target values are known. This is important for comparing methods in a research situation. Computing indices of the reliability of imputed values in current surveys is different because the computation usually must be done on a small fraction of the observations, but is important in practical applications of methods.

Comment: Unfortunately, the problems of predicting accuracy of imputed values and testing the reliability of alternative indices were not pursued in much detail in EUREDIT.

The methods studied for imputation included CANCEIS, Cherry Pie, DIS, E-M algorithm, GEIS, SOLAS and logistic regression. New methods included POEM and BEM for outlier processing, MLP, CMM, SOM and SVM for more general imputation. As for the new methods in WP4, important academic work has been carried out in WP 5. A number of the reports made within WP5 contain original development and will probably end as published articles in scientific journals**.**

Comment: For the new imputation methods, the contribution to the practical imputation in NSIs will probably be limited in the short run.

No deliverables were scheduled from WP4 and WP5 in the first 24 months. Partners working with these workpackages organized, however, special meetings and presented oral progress reports to the Project meetings. During most of the Project meetings, three breakout group meetings were also organized running for a couple of hours in parallel. Each group had members from different workpackages. Experience and knowledge were exchanged, and further interactive contact by means of the web site was expressed as desirable. The web master reacted fast to this desire and set up discussion pages to be used by different groups.

Comment: The possibilities for exchanging ideas and experiences were not utilized very much. It is a paradox that researchers working in close contact with the information and communication technology still have problems to use this technology in their communication.

Requests for written plans and reports ahead of meetings were frequently repeated, and when approaching month 24, the importance of presenting reports in time for other partners to read before the meetings was emphasized. Three-fourth of the listed deliverables were due from month 24 to month 36.

Comment: As pointed out in other connections, the partners delivered close up to, on or after deadline dates leaving little or no time for reading plans and reports before project meetings. It was unfortunate that deadlines for substantial deliverables were set to the last year of a project.

During the project period, feed backs from EUROSTAT on the EUREDIT periodic reports also emphasized the importance of obtaining and presenting results, which would be considered useful by the NSIs.

Comment: More attention should have been paid to introduction of the methods in practical editing work. This task was postponed to WP 6 at the end of the project with little possibility and time for interaction among the partners with respect to this important aspect

EUREDIT aimed at documenting comparative performances of investigated methods. During one meeting among the partners working with neural network methods, it was proposed and agreed that relationships among different approaches, differences and similarities, should be outlined for non-expert users. This could be, it was said, a useful document for the deliverable D6.2 for a wider public. However, no resources were available for developing this idea further.

Comment: A project like EUREDIT, comprising a number of methods and partners with very different approaches, would have benefited from more interaction among partners representing different approaches between the meetings.

The topic 'selected editing' is important from a practical point of view. However, selective editing is not an editing method per se, but an approach combining different, frequently manual and automatic, E&I methods effectively. Selected editing is an example demonstrating the drawbacks of the limited scope of EUREDIT representing a useful topic, but not within the EUREDIT scope of automatic editing.

The methods had different requirements to software development. For some of the methods in current use, software developed at advanced NSIs already existed, and the planned experiments required adjustments to the experimental situation. For MLP methods commercial software such as SAS Insight and SPSS Clementine were used, while for other methods extensive development and coding of the methods were required.

Comment: What has been evaluated is in fact the combined effect of methods and their implementation. For example, the commercial software used for experimenting with MLP did not permit the fine-tuning of basic parameters, as did several of the other software tools tailor-made for the particular experimental situation. There can obviously be excellent methods badly implemented, and vice versa.

In a brainstorming meeting on neural network, it was proposed that the training sets should be randomly partitioned into 25%, 50% and 75% of the already distributed sets. This would make it possible to evaluate the effects of the size of the training sets on the trained neural networks, and the partitioning was done.

Comment: Unfortunately, time did not permit an analysis of the effects of the training set size on the performance of the MLP.

WP6: Evaluation and validation of results of WP4 and WP5 to determine the 'best' methods

The importance of WP 6 was discussed repetitively on several occasions and the importance of presenting the results both for methodologists and for staff with operative responsibilities was emphasized.

For some time, NSIs have discussed the concept 'quality of statistics' and the related 'best' current methods. There is a consensus that the quality concept is complex and multi-dimensional. In practical survey work with restricted resources, it is for example necessary to find a balance between timeliness and accuracy. Even though the scope of EUREDIT was limited to evaluating the statistical accuracy of editing and imputation methods, it is important to emphasize that also many aspects related to evaluation of edit and imputation methods had to be left unconsidered. For example, editing and imputation methods should not only be evaluated separately, but also as pairs of methods when possible and relevant.

The aim of the editing and imputation processes is to contribute to improvement of statistical quality both for the current survey results as well as for future survey results. In addition to detecting and correcting errors, the editing and imputation processes are therefore important sources of knowledge for improvement of other processes in future statistical surveys. This aspect should be brought to attention when discussing evaluation and validation.

Comment: Several ideas to make the evaluation process more general were discussed and documented in reports. One was the operational aspects of comparative evaluation, in which factors for practical comparison of the methods where identified as a supplement to the statistical measures. Another was a presentation of qualitative and operational criteria, which provoked useful discussion. Both presentations contained ideas, which had useful impacts on the work in WP6.

To meet the wider requirements to evaluation, a highly needed document*, The Way Forward - The EUREDIT evaluation Handbook,* was prepared.

To support the final evaluation process, a set of programs for evaluation of editing and imputation experimental results was developed for use by the Program Coordinator. Several participants also indicated interest for and used the evaluation software to test the results of alternative versions of their editing and imputation methods before their final experiments.

In addition to discussions in last couple of project meetings, an editorial committee was established to support the work on the final deliverables of WP6. This committee met a couple of times and contributed to the preparations of these publications.

Comment: In the last part of the project period, it became obvious for most partners that successful completion for EUREDIT depended to a large extent on the presentations from WP6, which was responsible for 2 of the most important deliverables from the project. However, the work in WP6 started too late.

WP7: Integration of the individual edit and imputation methods into a single package for wider dissemination

The objectives of WP7 were to develop portable, prototype software of selected techniques investigated in EUREDIT. The prototype was planned to have a limited distribution. Because the partner responsible for the implementation of the software depended on the results from other partners experimenting with the selected methods, the work on this task had to start late in the project.

Comment: The work in this work package was delayed because the dependence on the work of the other partners, and little time was left for discussing the prototype. Again, results from the WP4 and WP5 at an earlier time than month 24 would have eased the time pressure, and more time for discussing the form and content of the prototype.

As mentioned above, some of the software used in the experiments were developed outside EUREDIT, and create problems for developing a comprehensive prototype.

Comment: Principles for how to solve these problems should have been discussed more thoroughly in an early stage of the project. In general, is even a prototype implementation of methods which are being developed up to the last few months before the end of a project like Euredit, be a realistic goal?

WP8: Dissemination and exploitation

The WP8 objectives were to present the results of EUREDIT to the public and prepare exploitation of the results.

The project established its own website with a public (open) and private (restricted to project partners) section. The website was started early, and has function well through the project period.

The private pages have been continually developing. At the end of the first 6 project months, a number of relevant reports had already been submitted and partners were encouraged to submit their papers to the web site.

New ideas for the website were brought up during the project meetings such as the possibility to have links to bibliographies for editing and imputation, introduction of news groups, monitoring the use of the different pages to identified those most frequently used, a list of participating staff members with their e-mail addresses and URL's in the private section for easy reference, etc.

In addition to the website, a mail server has also been implemented and functioning. I hope that both the website and the mail server will be used more in the future for informing participants about the progress of the project and its work packages.

Comment: While the web site functioned well as a means to distribute documents for meeting, the use of the site by the different partners for exchanging ideas was disappointing.

Already in the proposal, the need for funding of conferences to present project results and receive feedback was pointed out. No additional funding from EU was approved in spite of proposals and applications.

The Finish partners managed, however, to obtain local funding for a conference, DataClean, hosted by the University of Jyväskylä. An organizing committee was established, a large response was received to the call for papers and a successful conference was arranged.
.

Comment: It is important for the dissemination of results from EUREDIT that proceedings from this conference are completed and published.

The partner responsible for the dissemination of the software carried out a market and concluded there was a need for a commercial product and that partner was prepared to satisfy this need. A final TIP was developed on this basis.

WP9: Project evaluation

To observe the activities and review the deliverables, an internal evaluator participated in project and management meetings, and prepared biannual reports for the Project Coordinator in which the project progress was assessed. These reports were also made available to EUROSTAT in connection with the annual project reviews. The present report is the final project assessment by the internal reviewer.

4. Assessment of project deliverables

In total, 41 deliverables were planned. A list with details is reproduced in ***Appendix 4***.

The results from the Euredit project can be divided into the following categories:

- Internal deliverables
- Public deliverables
- Published papers
- Conference reports
- Other

Internal deliverables

Most of the deliverables, 31, belonged to the category of Internal deliverables. As mentioned above, the deadlines for the first deliverables from WP4 and WP5 were month 24. The deadlines were respected, but quite a number of the reports were available only short time before the meetings at which they should be discussed.

The deliverables in this category were either guidelines/standards to be used in the partners' reporting issued by the project management, reports on progress from the different partners or draft reports from research. The internal deliverables were of limited interest for the public. Many of the internal deliverables from the WP4 and WP5 can be considered as outlines or drafts to parts of the public deliverables of WP6 discussed below.

Some of the deliverables were, as already indicated, of significant scientific value, and to the extent they are not included in the public deliverables, it should be hoped that their authors would complete and submit these deliverables to scientific journals.

Public deliverables

The 10 public deliverables were:

- D3.2 Prescriptive guidelines for evaluation
- D3.3 Final report on measuring quality of edit/imputation procedures
- D6.1 Report on evaluation of all methods
- D6.2 Best practice guide for edit/imputation for given data structure
- D7.1 CD-ROM containing prototype software and documentation
- D8.1 Project presentation
- D8.2 Initial Project Web site
- D8.4 Report of exploitation potential (TIP)
- D8.5 End of project conference (if funded)
- D9.2 Final assessment report on project quality

D3.2: Prescriptive guidelines for evaluation

The purpose of this deliverable was to give guidelines for the evaluation of the methods from a general and general point of view. The purpose was to supplement the deliverable D3.3.

D3.3 Final report on measuring quality of edit/imputation procedures.

This deliverable was a core product of the project in the sense that it specified the underlying methodological framework for measuring the statistical quality of the experimental performance of considered methods applied to available data sets. The deliverable focuses on the accuracy component of statistical quality.

A first draft was discussed in month 6. Several 'final' versions have been published on the Web site, by ONS and as a chapter of D6.1.

D6.1 Report on evaluation of all methods

The public deliverable D6.1 is presenting the methodological results and evaluations from the EUREDIT project. The publication has an introduction, 8 chapters and 6 appendices (Appendix 5) and cover more than 400 pages.

There are 3 groups of methods investigated. The first group contains automatic methods, called standard methods, which are used or ready for use in national statistical institute. The next group includes outlier methods and robust imputation methods developed partly from scratch by Euredit partners, while the last group comprises methods developed outside EUREDIT by project partners or others for more general use, but adjusted for editing and imputation by EUREDIT partners. Chapter 1-8 discuss the different methods investigated and evaluated. Most of these chapters have already been presented in several versions as internal deliverables.

The introduction is a presentation of the EUREDIT project objectives, plans and implementations prepared by the coordination partner, *Office for National Statistics*.

In Chapter 1, the project partners, *Instituto Nationale Di Statistica* and *Centraal Bureau voor de Statistiek*, present ***Evaluation of Edits and Imputation Methods Using Standard Methods***. With about 100 pages, the chapter is the longest in the volume with 10 different standard methods are described and evaluated. In the context of Euredit, a standard method is a method already used or ready for use in a national statistical institute. It is typical for these methods that they require pre-determined edits and imputation rules in some cases deduced from edits.

In the second chapter, ***Evaluation of Edits and Imputation Using Robust Methods,*** reports on experiments and evaluation of methods developed by *Swiss Federal Statistical Office* and *University of Southampton*. These methods approach the editing and imputation from as an outlier problem without any requirements to preset edits. Also in this chapter, 10 different methods are discussed, and finally discussed.

Chapter 3 is on ***Neural Network MLP*** and prepared by Danmarks Statistik. *Multiple Layer Perceptrons* were originally models developed for simulating how human sets of neurons worked.
In the context of statistical editing and imputation, the hypothesis was to investigate if a MLP can simulate the work of human editors performing statistical editing and imputation. The MLP is prepared for the specific task by an iterative training process, which requires a training set of data reflecting the human editing and imputation. Commercial programs were used for the experimentation.

University of Jyväskylä has prepared the fourth chapter, ***Evaluation of SOM based Editing and Imputation,*** with contributions from *Statistics Finland*. *Self-Organizing Maps* is another type of neural networks, but in contrast to MLP discussed in Chapter 2, *SOM* nets do not require any training set. For the EUREDIT project, a special tree-structured variant of SOM, ***NEAT-DATA***, developed at the *University of Jyväskylä* was used in a number of experiments.

Chapter 5, ***Evaluation of Edit and Imputation Using CMM***, was prepared by *University of York. Correlation Matrix Memory* is a third set of neural networks investigated in EUREDIT. This type of networks requires training, but, in contrast to *MLP*, only a single pass through the training set is enough. *CMM* methods have been developed and studied by *University of York* for many years, and were adjusted for editing and imputation within the EUREDIT project.

Chapter 6 is prepared by *Royal Holloway College Univ. London* on ***Imputation Using Support Vector Machines***. *SVM* are methods for non-linear classification and regression and have some similarity with *MLP*. The SVM methods and algorithms have been developed before the EUREDIT, but have been adjusted to handle imputation problems in statistical processes.

The next chapter, ***Evaluation of Imputation Methods for Panel and Time Series Data***, reports on the investigations and evaluations done by *Quantaris*.The methods are all for imputing missing data in financial panel/time series data. Ten methods are studied and evaluated.

Chapter 8, ***Evaluation of Edit and Imputation Performance***, is a modified version of D3.3 and adjusted to fit as a reference for the readers of this volume.

The chapters 1-8, all end by comparisons and identification of weakness of the respective methods.

This volume also contains 6 appendices. They contain a complete index to all experiments, a detailed description of the data sets and the perturbations, results from an experiment to determine editing rules automatically, papers on adaptive censoring and Bayesian networks, and finally a description of the attached CD (see section on D7.1 below).

Comment: Even though the drafts were not discussed to the extent they may have deserved, the strategy of requiring the drafts as internal deliverables was effective and contributed to better final deliverables.

Survey planers at NSIs will probably find the Chapter 1 most useful, and consider the methods in the remaining chapters in a development stage too far from practical use.

University statisticians will on the other hand probably find the approaches in the chapters 2-7 interesting, and in editing and imputation see a new and important research area.

These chapters should represent a valuable source of information for methodologists and students working with editing and imputation method development and evaluation.

A Chapter 9, named* Comparative overview of results*, was planned to contain an extensive comparative evaluation of all methods. In the last stage of the preparations this chapter was dropped and the readers referred to Chapter 4.1 of the next deliverable, D6.2.

This volume is intended to be disseminated on a CD attached to the next deliverable, D6.2. In my opinion, this valuable and unique source to experiences from the so far most extensive experiments with editing and imputation methods should be available as an independent publication.

D6.2 Best practice guide for edit/imputation for given data structure

Deliverable 6.2 was in the project plans referred to as *Best practice guide for edit/imputation for given data structure*. This is a too ambitious title since the methods investigated are limited and the criteria for identifying the *best* are not unique. A less pretentious title will probably be chosen for the final presentation.

D6.2 was intended to be the presentation of the problems and findings for users and readers not interested in methodological details or the scientific aspects of the research as presented in D6.1. The volume was planned with 5 chapters, and a list of references, an appendix of formulae and a glossary attached.

Comment: The volume has become more important than originally planned because after dropping Chapter 9 in the previous volume, the section 4.1 is now the only place a comparative review of experiments with all methods is reported, and because Chapter 5 presents the final answers from the project in the form of guidelines.

Chapter 1: ***Editing and imputation*** written by *ONS* gives an introduction to automatic editing and imputation. The next chapter, ***The EUREDIT project***, also prepared by the coordinating partner *ONS*, summarizes the project framework**.** ***Overview of the methods*** by University of reviews the methods investigated in the project, and is an excellent introduction.

Southampton Chapter 4 contains ***Results of the EUREDIT project*** starting with an ***Overall summary.*** The summary was written by *ONS* and *University of Southampton*. The summary is of particular interest because a comparative analysis across *all* methods is not found any other place in the public deliverables.

Then follows sections reporting on the performances of the investigated methods. Opposite to the presentation in D6.1, the presentation in this chapter is by *data sets*. The user can choose the

data set representing his area of interest, and read how the different methods perform on this particular data set.

The last Chapter 5 of D6.2, ***Recommendations towards edit and imputation strategies***, was intended to give the users guidelines for selecting an editing-imputation strategy, and the chapter to which practitioners will look for to find guidance and support for their plans.

The D6.2 also includes an extensive list of references to publications on editing and imputation, a summary of formulae used.

Comment: This publication is well written and the chapters can be read without a very advanced background in editing and imputation methodology.

However, the acceptance of D6.2 by the main target group, developers in national statistical institutes, may be varying. Some will welcome this volume because it gives the formalization of automatic editing and imputation they have been missing. They will also find useful method information and comparisons. Others readers may be disappointed because only parts of the volume present methods which can be applied in practical survey development without major adjustments. In spite of many interesting new ideas, these readers may find the presentation of the new methods too 'academic' for the practionners.

D7.1 CD-ROM containing prototype software and documentation

An important component planned on the CD-ROM is the collection of data sets. The number of data sets will be reduced to because of confidentiality restrictions.

Implemented prototype software is also planned to be included on the CD-ROM. The partner NAG is responsible for the implementation. During the last project meeting in February 2003, the software was explained and the interfaces demonstrated.

Comment: This deliverable was not available at the time of finishing this report.

D8.1 Project presentation

The Project presentation was prepared in the beginning of the project. It was delivered to and well accepted by different audiences. It has also served as useful project description in several reports and other deliverables.

D8.2 Initial Project Web site

The web site was established according to the time schedule in month 1. It has been conscientiously maintained and extended, and particularly the private part has served as an effective means for disseminating guidelines, standards and reports from the project management.

Comment: The partners did not take full advantage of the possibilities offered by the web site as a means for information exchange.

D8.4 Report of exploitation potential (TIP)

NAG has carried out a market analysis, and concluded that there is a need for imputation software and is prepared to satisfy the market needs.

An ***eTIP*** has been prepared and uploaded to the EU-Cordis web site (http://etips.cordis.lu).

D8.5 End of project conference (if funded)

As indicated above, EU did not grant funding for project conference. An 'outside' funding was contributed by Finnish sources for a conference in Jyväskylä in May 2002.

Comment: It must be a severe misallocation not to grant funding to promote the result of a large project like EUREDIT to the target population.

D9.2 Final assessment report on project quality

The present report.

Published papers

In addition to the scheduled deliverables, partners have, based on their EUREDIT activities, also submitted papers to journals and conferences with referee requirements. The following papers represent some indirect results of the project:

- Cédric Béguin and Beat Hulliger: *Détection de valeurs aberrantes*, French Statistical Society conference on survey methodology.

- Cédric Béguin and Beat Hulliger: *Detection of Multivariate Outliers by a Simulated Epidemic*, ETK/NTTS 2001 Conference Paper.

- Cédric Béguin and Beat Hulliger: *Multivariate Ausreisser-Entdeckung in unvollst¨andigen Stichprobendaten,* Deutsche Statistische Woche, October 2002.

- J Breckling, P Kokic and O Lubke: *A note on multivariate M-quantiles*, Statistics & Probability Letters, Volume 55(1), 39-44.

- P Kokic, O Lubke and J Breckling,: *Outlier identification using multivariate expectiles*,

- P Kokic, J Breckling, and O Lubke: *A new definition of multivariate M-quantiles*, L1-Norm conference in 2002.

- R Chambers: *Evaluation Criteria for Statistical Editing and Imputation*, UN/ECE meeting in Cardiff 18-20 October 2000

- R Chambers: *Evaluation Criteria for Statistical Editing and Imputation*, Report No. 28 in the *National Statistics Methodology Series.*

- P Piela: *Introduction to Self-Organizing Maps Modelling for Imputation – Techniques & Technology*, ETK/NTTS Conference Papers.

Conference reports

As reported above, deliverable D8.5 was planned in case additional funding was obtained. No funding was obtained and that deliverable could not be realized. However, the Finnish partners managed to obtain some local funding for a conference, DataClean 2002, in Jyväskylä, Finland, in 29-31 May 2002 on erroneous and missing data. A majority of the papers presented were based on EUREDIT research and were well accepted.

Given the restricted funding, the conference was a great success.

Comment: A report from the conference has been planned, but so far not been published.

Other

Last, but not least, the Euredit project has contributed to the partners' understanding of the statistical editing and imputation processes and their competence to solving problems in this domain.

5. General conclusion and final remarks

General conclusion: EUREDIT has been a very productive and successful project. The leadership and management has been stimulating, professional cooperation good, the scientific quality of the results high, the presentation to the statistical community interesting, and the prospects for dissemination and exploitation seem good in spite of uncertain funding.

It has been a great pleasure for me personally to serve as an internal evaluator for the EUREDIT project. I have observed both good management and research by motivated and competent professionals.

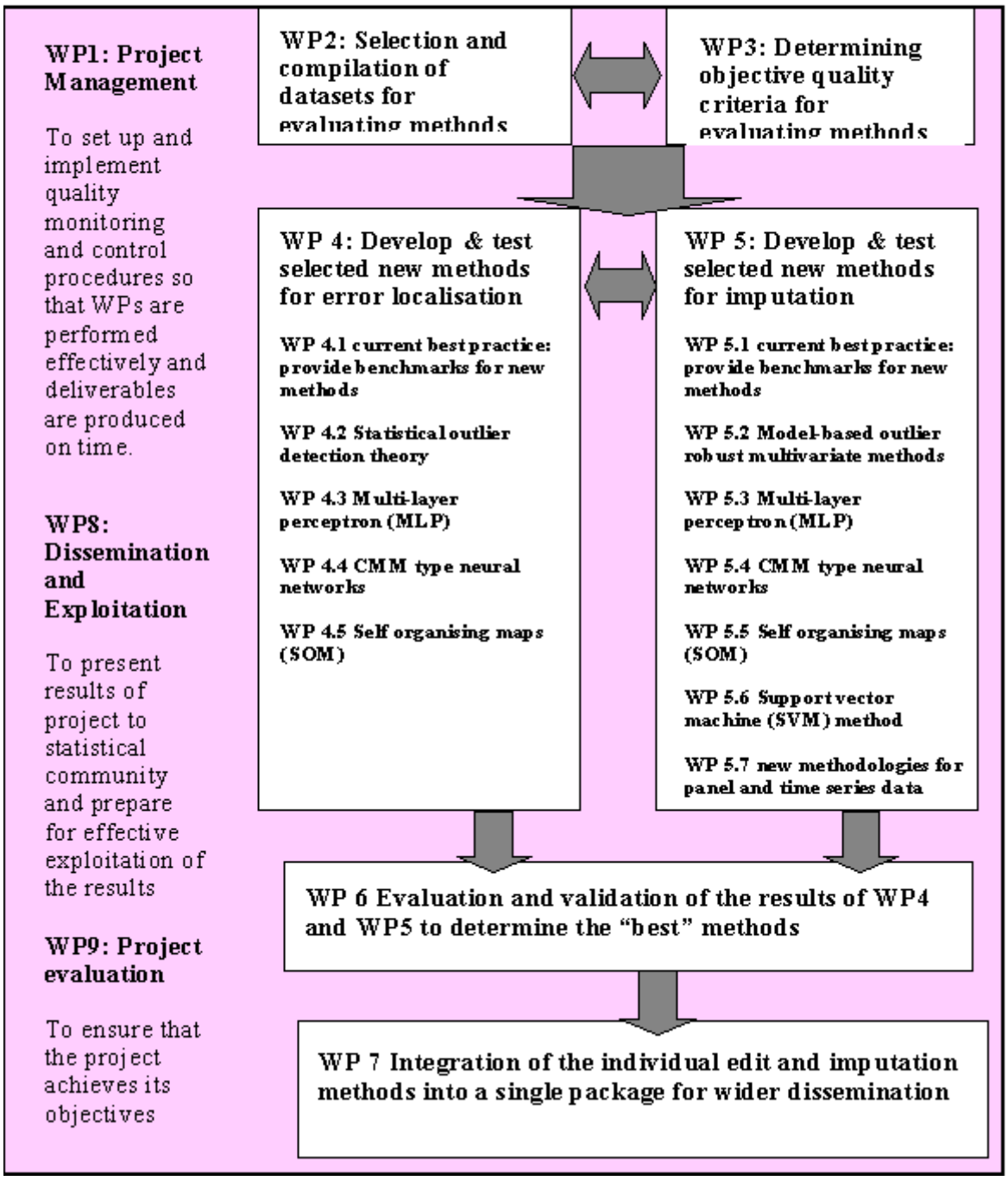
During the 3 project years, I have learned much about project work in general, and editing and imputation in particular. I would like to thank all of partners for friendly cooperation, and particularly thank John Charlton for help he always was ready to give as Project coordinator.

Appendix 1: Partners in Euredit

Participating organisation	Country	Contact-Person
Office for National Statistics (coordinator)	UK	John Charlton
Royal Holloway College Univ. London	UK	Alex Gammerman
University of Southampton	UK	Ray Chambers
University of York	UK	Jim Austin
The Numerical Algorithms Group Ltd	UK	Geoff Morgan
Centraal Bureau voor de Statistiek	Netherlands	Ton de Waal
Tilastokeskus (Statistics Finland)	Finland	Seppo Laaksonen
University of Jyvaeskylae	Finland	Pasi Koikkalainen
Swiss Federal Statistical Office	Switzerland	Beat Hulliger
Quantaris GmbH	Germany	Phil Kokic
Istituto Nazionale Di Statistica	Italy	Giulio Barcaroli
Statistics Denmark	Denmark	Peter Linde

Appendix 2.

EUREDIT: Diagram to show the relationships between the different parts of the project.



Appendix 3.

Partner	*Role*	*Country*	*Budget (K. EURO)*	*Person Months*
Office for National Statistics	**AC**	**UK**	**348,807**	**84**
Royal Holloway College	**CR**	**UK**	**88,385**	**18**
University of Southampton	**CR**	**UK**	**251,737**	**60**
University of York	**CR**	**UK**	**291,196**	**37**
NAG Ltd.	**CR**	**UK**	**118,752**	**26**
CBS	**CR**	**NL**	**183,301**	**47**
Statistics Finland	**CR**	**FIN**	**89,310**	**25**
University of Jyvaeskylae	**AC**	**FIN**	**123,671**	**33**
SFSO	**CR**	**CH**	**227,132**	**29**
Qantaris	**CR**	**D**	**131,092**	**48**
ISTAT	**CR**	**I**	**131,092**	**42**
Statistics Denmark	**CR**	**DK**	**246,617**	**51**

Appendix 4.

Deliverables list

Del. no.	Del. Name	WP no.	Lead participant	Estimate person-months	Del. Type	Security*	Delivery (proj. month)
D1.1	Guide/template for common standards	1	ONS	6	Report	Int	3
D2.1	Standard datasets on CD-Rom with software to generate errors / missingness plus users' guide incl data dictionary	2	ONS	13	CD of data and software	Int	6
D2.2	Report on the rationale for including the chosen data sets	2	ONS	1	Report	Int	12
D3.2	Draft technical report describing the underlying methodological framework for measuring statistical and operational quality in edit and imputation methods. ε	3	U.Soton	11	Report	Int	5
D3.2	Prescriptive guidelines for evaluation	3	U.Soton	11	Report	Public	6
D3.3	Final technical report on measuring quality of edit/ imputation procedures ε	3	U.Soton	3	Report	Public	12
D4.1.1	Software/ documentation for methods in WP4.1	4.1	CBS	6	software	Int	24
D4.1.2	Report on evaluation of methods in WP4.1	4.1	CBS	2	Report	Int	24
D4.2.1	Report describing methodology and results of evaluation for methods in WP4.1	4.2	SFSO	6	Report	Int	24
D4.2.2	Algorithms used to	4.2	SFSO	2	software	Int	24

	implement and evaluate these methods						
D4.2.3	Production of SAS-Macros tool box, recommendations with documentation.	4.2	SFSO	3	software	Int	24
D4.3.1	***Report describing the application of MLP technique to data editing, and experimentation.***	4.3	DSt	2	Report	Int	24
D4.3.2	Prototype MLP neural networks editing system.	4.3	DSt	6	software	Int	24
D4.4.1	Report describing the application of CMM techniques to data editing and the experimental method.	4.4	U. York	2	Report	Int	24
D4.4.2	Algorithms and software used to evaluate CMM methods.	4.4	U. York	6	software	Int	24
D4.5.1	Report describing the application of SOM to error localisation.	4.5	StatFi	2	Report	Int	24
D4.5.2	A logical description of the method and algorithm.	4.5	StatFi	6	software	Int	24
D5.1.1	Software and necessary documentation implementing the selected methods.	5.1	ISTAT	6	software	Int	24
D5.1.2	Report covering an evaluation of the selected methods.	5.1	ISTAT	2	Report	Int	30
D5.2.1	A report detailing the methodology and results for outlier robust multivariate imputation.	5.2	U. Soton	2	Report	Int	24
D5.2.2	Software used to implement and evaluate these	5.2	U. Soton	6	software	Int	24

	methods.						
D5.3.1	A Report describing the methodology developed and experiments conducted.	5.3	DSt	2	Report	Int	30
D5.3.2	A prototype imputation system based on MLP technique.	5.3	DSt	6	software	Int	24
D5.4.1	A report describing the application of CMM techniques to data imputation.	5.4	U. York	2	Report	Int	30
D5.4.2	Software and algorithms used in evaluation of CMM neural networks.	5.4	U. York	6	software	Int	24
D5.5.1	A report describing the method and experimental results.	5.5	StatFi	2	Report	Int	30
D5.5.2	Software for imputation based on SOM type neural networks.	5.5	StatFi	6	software	Int	24
D5.6.1	A report describing the application of SVM-based techniques to data imputation.	5.6	RHUL	2	Report	Int	30
D5.6.2	Software based on SVM technique.	5.6	RHUL	6	software	Int	24
D5.7.1	A description of the algorithms used in data imputation.	5.7	Insiders	6	software	Int	24
D5.7.2	Final report on methodologies and experimentation developed for WP 5.7.	5.7	Insiders	2	Report	Int	30
D6.1	Report on evaluation of all methods ε	6	ONS	3	Report	Public	36
D6.2	Best practice guide for edit/ imputation for given data structure ε	6	ONS	3	Report	Public	36

D7.1	CD-Rom containing prototype software and documentation	7	NAG	15	Software	Rest/ mostly Public**	36
D8.1	Project presentation	8	ONS	0.3	Report	Pub	3
D8.2	Initial Project Web site	8	NAG	1	Web-site	Pub	1
D8.3	Dissemination and Use Plan (initial TIP)	8	NAG	1	Report	Int	6
D8.4	Report of exploitation potential (final TIP)	8	NAG	1	Report	Pub/Int	36
D8.5	End of project conference (if funded)	8	NAG	(6)	conference	Pub	33
D9.1	Assessments of work in progress	9	ONS	3	Report	Int	6-monthly
D9.2	Final assessment report on project quality	9	ONS	1	Report	Pub	38

**Int.* *Internal circulation within project (and Commission Project Officer if requested)*
Rest. *Restricted circulation list (specify in footnote) and Commission PO only*
IST *Circulation within IST Programme participants*
FP5 *Circulation within Framework Programme participants*
Pub. *Public document*
*** Most of the material will be public but some components may include licensed software*

Appendix 5.

Chapters in D6.1.

0. Introduction
1. Standard methods
2. Robust methods
3. MLP neural networks
4. SOM neural networks
5. CMM neural networks
6. Support Vector Machines
7. Methods for panel and time series data
8. Evaluation Criteria
Technical Appendices
A. Results of all key experiments. To include 'Table of Contents' for tables.
B. Detailed descriptions of datasets and their perturbations.
C. Results of using WAID on ABI to determine key edit rules.
D. Adaptive Censoring
E. Bayesian networks
F. Contents of CD-ROM

Appendix 6.

Chapters in D6.2: *Best practice guide for edit/ imputation*

Foreword
1. Editing and imputation issues
2. The EUREDIT project
3. Overview of the methods investigated by the EUREDIT project
4. Results of the Euredit project (user perspective)
4.1 Overall summary
4.2 ABI dataset
4.3 EPE dataset
4.4 DLFS dataset
4.5 SARS dataset
4.6 GSOEP dataset
4.7 Panel and Time Series dataset
5. Recommendations towards edit and imputation strategies
References
Appendix of Formulae