K.J. Poppe
G. Beers

Mededeling 533

PACIOLI 1

On data management in farm accountancy data networks

Reflection paper

August 1995

PACIOLI

Agricultural Economics Research Institute (LEI-DLO)

REFERAAT

PACIOLI 1; ON DATA MANAGEMENT IN FARM ACCOUNTANCY DATA NETWORKS; REFLECTION PAPER
Poppe, K.J. en G. Beers
The Hague, Agricultural Economics Research Institute (LEI-DLO), 1995
Mededeling 533
ISBN 90-5242-298-2
36 p., tab., fig.

This is the first of four reflection papers to provide suggestions for decision making on the further development of the RICA. The reflection papers are submitted to the management committee of the RICA by the concerted action PACIOLI. The concerted action aims to improve the quality of agricultural accountancy and FADNs. The topic of this paper is information management as a method for innovation.

This reflection paper provides suggestions for decision making on the further development of the European FADN, based on papers of the first PACIOLI workshop. Although information modelling has some risks, it seems to be a promising tool. By creating a data model or data dictionary it could improve the flexibility of the RICA system. This flexibility is necessary now that policies are changing and the EU is widening.

Farm Accountancy Data Networks/Farm return/Farm accounting

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Poppe, K.J.

PACIOLI 1 : on data management in farm accountancy data networks : reflection paper / K.J. Poppe and G. Beers. -
The Hague : Agricultural Economics Research Institute LEI-DLO). - Fig., tab. - (Mededeling / Landbouw-Economisch Instituut (LEI-DLO) ; no. 533)
ISBN 90-5242-298-2
NUGI 835
Subject headings: farm accountancy / data networks / information management.

CONTENTS

'On ne connait que les choses que l'on apprivoise'

Antoine de Saint-Exupéry: Le petit prince

SUMMARY

This is the first of four reflection papers to provide suggestions for decision making on the further development of the RICA. The reflection papers are submitted to the management committee of the RICA by the concerted action PACIOLI. The concerted action aims to improve the quality of agricultural accountancy and FADNs. The topic of this paper is information management as a method for innovation.

The Farm Return describes the data that are exchanged in the RICA. It is structured as a set of tables and origins from a time in which punch forms were used. It is not optimally geared to the accountant supplying the data, nor to the user. Since the Farm Return offers no room for voluntary supplements of data, the RICA cannot play a role in gradual innovation by expanding the data set for some of the surveyed farms.

Since the introduction of the current Farm Return in 1977 RICA's environment has changed, to put it mildly. The first PACIOLI-workshop identified some important trends that will influence farm accounting and the RICA in the coming years. Technological trends suggest that more and more farmers (will) own computers. Electronic links between computers are being established. This provides technical possibilities to access local databases and has huge consequences for the steering of information flows. One of the political issues is the budget constraints that demand cost effectiveness more than ever. The enlargement of the EU and changing rural policies (agricultural as well as environmental ones) lead to an increasing heterogeneity in farm systems. This makes it less and less likely that data gathered with the current Farm Return describes the real world efficiently and correctly. Where some farmers have a low interest in information and a correspondingly low interest in joining an FADN, others might be willing to provide a lot of data as long as they are rewarded by a feed back of reference information for bench marking. This asks for a flexible response, where access to data of the farm held by third parties (banks, suppliers etc.) could be helpful.

Where the current data management in RICA is based on an old, inflexible punch form approach, and current trends demand more flexible instruments, information models can come to the rescue. Information modelling is a cohesive aggregate of methods, techniques and tools which can be used to describe, analyse and eventually change information systems for an organisation, technically as well as functionally. The word 'model' should be interpreted in this context as 'map'. Information modelling provides, among others, a data model that describes on a logical level all the things of interest to an organisation of which data should be stored. A data model can be used by informatics experts to define data stores and files to be used

in the transfer of data. Section 6 of this paper gives an example related to table C of the current Farm Return.

Main advantages of a data model approach are:

* increased flexibility: for each attribute (data item) it can easily be agreed upon on which farms it should be recorded.
* all definitions (including derived statistics and translations) can be included in the model.
* the decisions on harmonisation and transfer of data are separated.
* meta-information can be supplied to users
* the data model and its physical representation (the data dictionary) can be stored in a work bench, which supports maintenance of the model and helps to generate software in a cost effective way.

The first PACIOLI-workshop identified potential problems when information modelling would be introduced on EU level, like the diversity in Europe, the lack of funding and the human factor. However, potential solutions were also identified.

Balancing the advantages, problems and potential solutions, and taking into account the need for more flexibility in the Farm Return and the existing tools to create this, it looks worthwhile to launch a short feasibility study to investigate if more flexibility can be created in the Farm Return. This flexibility could be based on a data-dictionary containing a data-model with the entity-types and attributes of the current Farm Return and that can be widened to incorporate additional data of interest for e.g. pluri-activity, cost of production, environmental issues, forestry etc. In the longer run one could imagine a situation where a huge virtual logical data base exists in which all the relevant data that are available in the Member States is described. Retrieving the data would become a technical and juridical issue.

The proposed study should incorporate the following topics:

+ the interest in member states and with the Commission to improve flexibility in the Farm Return in order to introduce statistics on new policy fields.
+ the interest to use a more up to date data modelling technique to improve data exchange and software development.
+ the selection of a work bench (or comparable tool) to support the development of a data model / data dictionary
+ the conversion of the current Farm Return to the format of a data model.
+ a test of the flexibility with data on gross margins (cost of production) or pluri-activity by incorporating data definitions used in member states in the data model.

An appropriate name for such a study would be RICASTINGS: RICA's Study To Install a New Generation of Statistics.

1. INTRODUCTION OF PACIOLI

This paper 1) is one of the deliverables of a concerted action in the EU's AIR-Programme, called PACIOLI (Panel in ACcounting for Innovation, Offering a Lead-up to the use of Information modelling). PACIOLI brings together scientists from several member states, who are interested in farm accountancy, farm information systems and agricultural policy. The objectives of the concerted action are:

* improvement of the quality of accountancy and FADN data;
* stimulation of the use of accountancy and FADN data;
* improvement of information management in FADNs;
* improvement of cost effectiveness;
* asses the need for and feasibility of projects for innovation of accountancy and Farm Accountancy Data Networks (FADN).

In the concerted action four workshops will be organised, respectively on

a] information analysis;
b] accounting and managing innovation;
c] need for change;
d] suggestions for continuation.

The papers presented in the workshops are published (see Beers et al., 1995a) as they contain interesting information for scientists, accountancy organisations and software developers in the member states. The papers are also summarized in a paper that contains all the conclusions and the highlights of the extended report (Beers et al., 1995b).

In addition to these papers the results of each workshop in the concerted action are used to provide the RICA-community with a so-called 'reflection paper' that deals with a special issue. The purpose of these papers is to provide suggestions for decision making on the further development of the FADN, based on sufficient background from the workshop papers. The reflection papers are submitted to the management committee of the RICA. The issues of the four reflection papers are determined by the coordinator of PACIOLI and the head of the RICA-unit DG VI A/3.

More information on PACIOLI can be found in Beers, 1995.

---

1) The paper is written by Krijn J. Poppe and George Beers. The authors work with the Agricultural Economics Research Institute in the Hague. George Beers is project leader of the concerted action PACIOLI. Krijn J. Poppe heads the Dutch delegation in this concerted action and represents the Netherlands in the management committee of the EU's FADN (RICA). The paper benefited from discussions in the first PACIOLI-workshop and from suggestions by Luis Florez Robles, Nigel Robson, Thierry Vard and Tim Verwaart.

2. TOPIC OF THIS PAPER

This paper is the first reflection paper of PACIOLI. It follows the workshop on information analysis, held in March 1995 on the island of Ameland, the Netherlands. That workshop mainly focused on information analysis to explore the domain, also with an eye to the need for change. The purpose of this reflection paper is to describe information management in RICA, for the current situation and with some proposals for the future. This means that the paper deals mainly with methods and tools.

Future reflection papers will probably deal with the organisation and objectives of RICA ('managing RICA and innovation') and with new data requirements. Of course in practice these topics are linked, but in this paper the focus is on data management.

The next section describes the current farm return used in RICA. The description is mainly based on one of the contributions to the workshop by Poppe (1995). It is followed by a broad overview of trends that are important for data gathering by RICA in the future. Some of these trends (especially on future data needs) will be discussed in more detail in forthcoming reflection papers. Here we take the point that future adaption of the farm return will be needed. For the conclusion of this reflection paper it is not important which topics will be covered by these adaption.

Section 5 explains methods of data modelling as explained in the first workshop of PACIOLI in papers by Graumans (1995), Verwaart & Spiering (1995) and Poppe (1991). As these methods are not always easily understood, section 6 provides an example on the well known domain of the labour data in the Farm Return of RICA.

Section 7 discusses the possibilities to exploit the technique of data modelling in RICA. In section 8 recommendations are made for further action within RICA.

3. CURRENT FARM RETURN

The Farm Return is a format used by RICA to describe the data of an individual farm in the form in which it is exchanged between the member states. The current Farm Return was introduced in 1977 (published as Regulation (EEC) 2237/77 of the Commission dated 23.9.1977 in the Official Journal L 263, dated 17.10.1977) and replaced the first one, that lasted for a decade. The Farm Return is used to gather data on nearly 60,000 'commercial' farms in the EU (the figure will be revised upwards with the data from farms in Austria, Finland en Sweden). The RICA is a network of networks: accounting offices keep records of the 60,000 individual farms and submit the data to national liaison offices, who convert them to the Farm Return and send them to Brussels.

The Farm Return is structured as a set of tables, with numbered data items, fields and code-schemes. The tables do not correspond to the work methods of accountants (who use charts of accounts), nor to the data structure which is logical from a users point of view. This makes the Farm Return no more or less than a tool for conversion of data. The tables of the Farm Return describe a specific view of the farm which includes:

A. General information on the farm.
B. Type of occupation (tenure).
C. Labour input.
D. Number and value of livestock.
E. Livestock purchases and sales.
F. Costs.
G. Land and buildings, deadstock and circulating capital.
H. Debts.
I. Value added tax.
J. Grants and subsidies.
K. Production.

Recently the Farm Return has been modified to cope with the effects of the CAP-reform. This update (published as Regulation EEC nr. 2940 of the European Commission, dated 25 October 1993 in the Official Journal L 265 dd. 26.10.1993) changes:

Table A including a code for the type of the region regarding the Structural Funds
Table G officially including the value of quota
Table J adapting codes for subsidies to include subsidies for the environment and forestry
Table K giving rules to code the set aside areas

And introduces:

Table L data on quota (buying and leasing)

Table M data on compensations in arable farming ('Mac Sharry-payments').

The tables in the Farm Return contain the details of the data items under so-called 'headings' and each heading has one or more descriptions, with a serial number for each (sub)heading described. Many of the serial numbers will not be used by some farms (e.g. headings for livestock on arable farms). Two tables demand the use of additional codes to specify the data entries. Table K uses product codes to specify the output. Table J demands the use of additional codes to specify grants and subsidies.

After defining the information that should be gathered and transmitted in the tables mentioned, the Farm Return provides additional definitions and instructions. These definitions comes in two types, depending on their juridical status.

It should be noted that nearly all the data items are obligatory. That makes it nearly impossible to exchange (on a voluntary basis) data that are available in some member states on a majority of farms on e.g. non-farm income, the environment, forestry or gross margins and production costs, and which is in demand by users (Williams et al., 1995). Much information is aggregated several times on its way from the farm to the users of the European RICA, which sometimes mean that valuable details are lost.

It is especially problematic that the costs in the profit and loss account are not allocated (with the exception of feedingstuffs) to the enterprises on those farms where that is feasible. So the costs are given by category and not by category and profit centre. This makes the calculation of gross margins or cost prices very difficult. If costs were given by category and profit centre, the Farm Return would provide more adequate information tot monitor the CAP which is based on policies by commodity (product).

By defining the Farm Return as the highest common factor of the national networks, the Farm Return is a limited instrument to exchange data on EU-level. Since the Farm Return offers no room for voluntary supplements of data, the RICA-network cannot play a role in gradual innovation by expanding the data set for some of the surveyed farms. In stead of providing leadership in the innovation of agricultural accounting it is forced to

follow and cannot easily transfer know-how on accountancy between member states 1).

As a tool for conversion the methodology is inflexible and seems outdated compared to newer methods of data management. Some examples:

* one can easily become confused by the numbers used for the headings and the serial numbers: number 90 stands for the interest paid on loans for land and buildings (heading in table F), but is also the serial number used on the magnetic tapes for the average number of equine (horses, heading 22 in table D).
* due to the record structure chosen for the magnetic tape, the current Farm Return can not be expanded any more. Nearly all available serial numbers in the range 1 - 1377 have been allocated. It is also clear that much physical record space on magnetic tapes is wasted because zero's are transmitted (e.g. table D and E) or global headings as well as subheadings are transmitted.
* in table M the reference yield of the farm (on which the Mac Sharry compensations are based) is transmitted per farm. This is however a data-item which is fixed per (Mac Sharry-)region, not per farm.
* one of the headings in the Farm Return describes the location of the holding (e.g. district) with a code. The meaning of these codes has to be provided (not necessary in electronic form) by the member state. This is however not used to connect the results of RICA to a geographical information system.

The methodology of the Farm Return origins from a time in which punch forms were used and data was described in two-dimensional tables. Therefore the tables look like punch forms, directly taken from field recording books. But in reality the data are gathered by an accounting process, using a chart of accounts to record the farm transactions. However a standard chart of accounts for European agriculture was (and most likely is) a utopia.

Representing information requirements in the form of tables, one would - from a user point of view - expect tables like: balance sheet, profit and loss account (or inputs and outputs), cash flow statement, cropping plan, general information. These are the standard statements used in providing information by agricultural accountants and agricultural (finance) management.

---

1) Of course the mission of the RICA-unit in DG VI is to provide (often confidential) policy information to DG VI and not to improve agricultural accounting or to make statistics. The RICA is a tool for that mission, not an end. The point however is that the RICA-unit needs control over the instrument to fulfill its function and that harmonized changes in the instrument cost a lot of resources (time) or are nearly impossible, both with the risk that it threatens the mission of the RICA-unit (due to too much time dedicated to data-management or due to outdated data). The next reflection paper will deal more explicitly with this dilemma.

But in the Farm Return this appears to be somewhat different, e.g. the cropping plan is integrated with production (table K), and livestock has been taken out of the balance sheet (table G) to table D and out of the costs (table F) and production (table K) to table E. Grants and subsidies are separated from production and costs (in a separate table J) but not from the table for investments (table G).

Another aspect of the Farm Return is that additional information is needed to calculate the results which are published by RICA. Examples are statistics like Livestock Units and all the income statistics like Family Farm Income and Net Value Added. The calculation of these income statistics demands rules on the valuation of the output of animals (especially the increase in value due to growth which has not yet been realised by sales and the treatment of price-developments) and the treatment of (investment) subsidies. Other data needed in order to analyse or publish results, are data on exchange rates, the inflation and data on the weighting of the farm. These types of data are not defined in the Farm Return but pop up in internal documents and publications of the Commission. This can lead to misinterpretations (e.g. two sets of livestock unit-coefficients are used, one of them based on the structural policy), inconsistency and errors.

Other additional flows of data within the RICA network are the control-program and updating coefficients for the RFS-model. The tests in the software programme are described in a technical document. These tests relate serial numbers to each other, e.g. if there is milk production there should be dairy cows, or signal unlikely high or low values (in relation to pre-defined limits per region).

In conclusion: the lay-out of the Farm Return is not optimally geared to the accountant supplying the data, nor to the user, nor is it efficient in terms of data transmission. The Farm Return has been developed as a tool for conversion of national data, and did not intend to harmonize national accounting methodology. In stead it creates a new set of data to compare results from different member states. For most of the original national farm accountancy data networks (with the exception of new member states like Greece, Spain and Portugal), the RICA is an 'add-on application' which did not influence the development of the national accountancy methods and definitions. Probably part of the lack of harmonisation in definitions, as noted by Power et al. (1989), can be attributed to the fact that RICA's Farm Return tends to follow in stead of setting developments in agricultural accountancy.

4. IMPORTANT TRENDS FOR TOMORROW'S DATA GATHERING

Looking to the future, the first PACIOLI workshop identified some important trends that will influence farm accounting and the RICA. In random order these trends can be grouped as technological trends, political issues, farmers' information needs, farmers' cooperation and costs of data. Some of these issues will be dealt with more explicitly in the next workshops (and reflection papers) of PACIOLI. They are discussed here briefly as they influence and support the suggestions made on data management.

Technological trends (Bonati, 1995) in informatics have important consequences. More and more farmers have computers themselves, and electronic data interchange (EDI) of information and documents is speeding up. This is due to new software (including the Internet and World Wide Web) and declining telecommunication costs.

As a result information and the processing of data is rapidly decentralising. This is happening on the user-side of RICA (e.g. in France where researchers throughout the country can use the individual data in the central national RICA database through a system called ARISTIDE that makes retrieval of individual data itself impossible) as well as on the accounting side. In many member states the accounts are made on personal computers, sometimes with electronic import of data. This leads to a decentralisation of software (including control programs) and integration of manuals in software. In some cases (like e.g. France or Italy) it even led to a decentralisation of authority, not to say a collapse of central power: the FADN-network evolved from a top/down approach to an alliance of local or regional organisations. This has huge consequences for the steering and control of information flows.

Electronic links between computer systems not only speed up communication, but also provide technical possibilities to access local databases: having access to information is now more depended on knowing where the data is and being allowed to use it, than on 'physically having / owning' the data. By applications like the World Wide Web the physical location of data in not of importance any more.

The main political issues that the RICA system has to deal with in the future are budget constraints and the enlargement of the EU. To reduce budget deficits (also with an eye to the Maastricht's EMU criteria) all EU Member States are trying to reduce costs, including the costs of their government information systems and FADN. More often governments only provide basic statistics, leaving other services to market forces (Persson, 1995). This means that national FADN managers are often under pressure to cut

costs by applying cost reducing technology or by cutting the amount of information gathered. At the same time new or reformed policies demand for additional and sometimes completely different information (e.g. environmental or veterinarian data).

On the European level the situation is characterised by an increasing heterogenity in farm systems and a multiplication of data suppliers (accounting offices, liaison organisations). Both are due to the enlargement of the EU from 6 to 15 member states, and with a pressure to facilitate comparisons of data with Switzerland, Norway and CEE-countries. This heterogeneity in data suppliers and farm systems makes it less and less likely that data gathered with the current Farm Return describes the farm systems and decisions on all farms correctly. And it also makes it much harder to incorporate new data items that can be gathered on all farms.

Farmers needs for information seems to increase. This is not only due to more paperwork (e.g. as a result of the CAP reform or environmental regulations) but also a result of farm enlargement, lower margins and higher management levels. More farmers integrate accounting with planning (Enroth, 1995). This application might also be interesting for policy makers; compare RICA's Forecasting System - RFS.

Although this topic will be discussed in more detail in the second PACIOLI workshop, it suggests that there are two types of farmers and from this two methods to run an FADN (or better: an FADN should be able to cope with both type of farmers). One category of farmers has a low interest in information, is (therefore) not willing to provide much data to an FADN and as a result the FADN should not be too ambitious in gathering a lot of data. An alternative for data gathering on this type of farms is to use other sources of data than the farm itself. It is not unthinkable (and in some countries already common practice) that farm data are supplied by organisations that interact with the farm, like banks, suppliers, auctions, dairy factories etc. Availability of new information and communication technology will give a lot of opportunities for relatively low cost data gathering without too much 'bothering' of the farmer (who of course has to grant his permission for this data retrieval).

The other category of farmers is well informed (often with its own PC), is interested in feed back from the FADN (especially in reference information for bench marking) and willing to provide data. The FADN can gather a lot of data on these farms without much costs, or even has to do so in order to guarantee the farmer's cooperation.

As the RICA aims to be a representative sample, it should be able to cope with both types of farmers, and in cooperating with the second category of farmers an FADN can reach a win/win situation where additional information for policy makers and research is gathered without much extra costs. However, this heterogeneity in farmers' styles asks for flexibility in accounting methods and ultimately the software that supports that process. Technological trends in informatics and data management makes this flexibility possible.

5. DATA MANAGEMENT

Previous sections argued that the current data management of the RICA is based on an old, inflexible punch form approach (section 3) and that current trends demand more flexible instruments (section 4). Information models can provide an answer to this demand.

Information modelling is a cohesive aggregate of methods, techniques and tools which can be used to describe, analyse and eventually change information systems for an organisation, technically as well as functionally (Martin et al., 1989). The word 'model' should not be interpreted here as an econometric model (like regression analysis or linear programming) but as a map: an information model gives a picture of the information that is used in an organisation, comparable with a road map that models the (potential) traffic flows in a country. An information model contains on a meta-level a description of data, not the values of the data items.

Several papers in the first PACIOLI workshop provide more details on information modelling. Graumans (1995a and 1995b) describes the method as well as the Dutch experiences in using the technique for standardisation purposes. Poppe (1991) provides an example used in farm accounting and Verwaart & Spiering (1995) describe the application of information modelling to renew the Dutch FADN. The last paper is used here to explain two important elements of an information model: the process model and the data model.

A process model describes all activities in the organisation that are related to information and decision making. Appendix 1 provides an example. The names of processes typically contain a verb. Processes are grouped in functions like strategic, tactical and operational management. Processes can be defined including incoming and outgoing data flows.

A data model describes all the things of an organisation of which data should be kept. The objective is to define all data and the relationships between data. The data are described with a unique definition, even if it is used in several processes.

In the end the data model is often more important than the process model, which is mainly used as a check on the completeness of the data model. While processes may change, data often stay the same and should not change in definition.

Important elements of a data model are the entity types with their mutual relationships. An entity type is a fundamental thing which is of relevance to the organisation, about which data could be stored. These things can be tangible (e.g. crops, machines) or intangible (e.g. region, result of a checking procedure). Appendix 2 gives an example of a data model from the Dutch FADN with the entity types in its accounting package, like farm (hold-

ing), general ledger, livestock etc. The lines between the entity types represent the relationships and indicate that the organisation is not only interested in the entity types as such, but also in their relationship (e.g. which farm family is using what amount of farm products - in a holding with two entrepreneurs and two families this provides more information than the total consumption of all the families involved at farm level).

Each entity type is described by a definition and its attributes. The definition creates the uniqueness and thus is the focus in harmonisation activities. Attributes in fact are describing the entity type: the attributes show which information related to the object (= entity type) the organisation wants to keep. Attributes for a cow (= entity type) could be its birth date, value, milk yield etc. All these attributes are also uniquely defined, providing additional harmonisation between the different processes that use the data.

Data models are made at a logical level by the users of the information, and can afterwards be used by information experts to design an optimal file structure (to exchange data electronically) and data base structure (to store data). Once the logical data model is defined by users, this can be done by computer experts.

The first PACIOLI workshop learned that not only the Netherlands but also e.g. the Italian INEA is using this approach for data management. It has been suggested that the Dutch and Italian experiences will be further discussed in one of the future workshops of PACIOLI.

6. AN EXAMPLE: DATA MODEL OF LABOUR

The data model approach already has proven to be fruitful to support data exchange and harmonisation. In section 4 it was argued that the exchange of data in RICA should be more flexible, to incorporate new data demands and to support harmonisation when data gathering decentralises. It seems therefore important to fully understand the methodology and to see how it could replace the current, outdated, punch form oriented Farm Return. Therefore this section provides an example with table C of the Farm Return. Readers not interested in this example can turn directly to section 7.

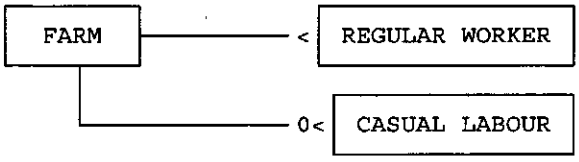
The current Farm Return specifies the following table C.

Table C Labour

Heading number and description	Code (1)	Year of birth (2)	Number of annual units (3)	Annual time worked (hours) (4)
A. Regular unpaid labour				
13. Holder / manager	51	52	53	54
	55	56	57	58
14. Holder / not manager	59	60	61	62
	63	64	65	66
15. Manager / not holder	67	68	69	70
	number of persons			
16. Spouse(s) of holders	71	-	72	73
17. Others	74	-	75	76
B. 18. Casual unpaid labour	-	-	-	77
C. Regular paid labour				
19. Manager	-	78	79	80
20. Others	-	-	81	82
D. 21. Paid casual labour	-	-	-	83

In addition to this table, the Farm Return specifies a definition for labour ('not including agricultural contract work'), holder ('person who assumes economic and legal responsibility, including share croppers'), regular ('at least a whole day of each week and in special cases for a limited period'), work ('including bookkeeping'), and annual units. Also the method of notation (time worked in hours, year of birth using the last two figures of the year) and predefined values (e.g. code 9 = unpaid managers in holdings operated as a company) are given somewhere in the instructions.

A data model on this information area would look like this:



The relationships read as follows: on each farm in the FADN there are one or more regular workers; and on each farm zero, one or more types of inputs from casual workers could exist. The description of the entity types would be as follows:

ENTITY TYPE REGULAR WORKER
Definition: person who at least works a whole day of each week

Comments: contract work is not viewed as labour input.
in special cases working during a limited period might be considered as regular.

Attributes: 1001 Code company managers
1002 Year of birth
1003 Number of annual units
1004 Annual time worked
1005 Holder
1006 Manager
1007 Spouse of holder
1008 Unpaid

ENTITY TYPE CASUAL LABOUR
Definition: all labour input not being regular labour

Comments: -

Attributes: 1008 Unpaid
1004 Annual time worked

The attributes of these entity types could be defined as follows:

number and name:	1001 CODE COMPANY MANAGERS
definition:	code that is used to identify unpaid managers in holdings operated as a company
comments:	-
format:	1 digit
possible values:	9
obligatory/optional:	optional
checks:	1005 and 1008 must be zero

number and name:	1002 YEAR OF BIRTH
definition:	year in which the person was born
comments:	-
format:	year in 2 digits, being the last two digits of the year
possible values:	01-95
obligatory/optional:	obligatory for managers or holders; optional for spouses and others
checks:	-

number and name:	1003 NUMBER OF ANNUAL UNITS
definition:	the annual work time of the person(s) involved in relation to the normal annual working time of a full-time worker in the region, with a maximum of 1 unit per person.
comments:	(text on disabled, as in footnote and additional instructions in the Farm Return)
format:	units in 3 digits, including one behind the decimal point
possible values:	0.1 - 99.9
obligatory/optional:	obligatory
checks:	-

number and name:	1004 ANNUAL TIME WORKED
definition:	Number of hours actually devoted to the work on the holding
comments:	work on the agricultural holding includes.. (see text in current Farm Return)
format:	hours, 5 digits, no decimal point
possible values:	1-99,999
obligatory/optional:	obligatory
checks:	-

number and name:	1005 HOLDER
definition:	person who assumes economic and legal responsibility
comments:	includes share croppers
format:	1 digit; values: 1 = holder, 0 = not a holder
possible values:	0-1
obligatory/optional:	obligatory
checks:	if value = 1 then 1007 and 1008 must be zero

number and name:	1006 MANAGER
definition:	person who undertakes day-to-day management of the holding
comments:	persons who undertake the management of only one activity should not be classified as manager but as 'other labour'
format:	1 digit; values: 1 = manager, 0 = not a manager
possible values:	0-1
obligatory/optional:	obligatory
checks:	if value = 1 then 1007 must be zero

number and name:	1007 SPOUSE OF HOLDER
definition:	number of person(s) who is/are married with the holder and work on the farm.
comments:	the number of persons can be higher than 1 if there are more than one holder. In that case also more records might be provided.
format:	1 digit
possible values:	0-9
obligatory/optional:	obligatory
checks:	if value > 0 then 1005, 1007 and 1008 must be zero

number and name:	1008 UNPAID
definition:	indication if the person(s) receive no remuneration for their work.
comments:	includes persons that receive a remuneration in cash or in kind that is far below the amount normally paid for the services rendered and which remuneration is not included as a cost in the Farm Return.
format:	1 digit; values: 1 = unpaid, 0 = paid
possible values:	0-1
obligatory/optional:	obligatory
checks:	if value > 0 then 1005 and 1007 must be zero

For reasons of clarity this example is a little bit simplified (e.g. the same attributes are used for regular and casual labour and no attribute is incorporated to include some 'hidden' information in table C that results from the prescription that holders should be entered in declining order of influence). However, it shows that the information in the current Farm Return can be very well described by a data model.

Such an exercise leads to a structured approach on a logical level. This implies that all software (forms, databases, transport protocols, report writers, spreadsheets, SAS- or SPSS-software etc.) in which this datamodel is implemented is always compatible. The technical translation of the logical model can be a choice made by informatics experts for a special application, as long as they can guarantee that the information can always be represented as described in the demands as stated in the datamodel. The information model itself can be stored in a workbench (a kind of database with computer added design facilities) and can directly be applied in writing soft-

ware: the attributes are often fields on a screen with the format as its lay out and the checks as an audit of the input.

The data model can be used by informatics experts to define data stores which are used in the exchange of data. This would lead to an agreement on technical aspects, comparable to those in annex III of the 1977 Farm Return. This would e.g. include that attributes with a value zero would not be transmitted.

A file on farm 88012 with two holders/managers (born in 1955 and 1957, working 2,500 and 2,350 hours) and one part-time paid casual labourer (1,000 hours) could lead to a file with the following records (assuming the introduction of an identifying number per entity-type as a key for the persons with the values 1, 2, 3 etc. and an attribute with the value 101/102 that identifies the entity type regular worker / casual labour):

```
1002 1 101 88012 55
1003 1 101 88012 1.0
1004 1 101 88012 2500
1005 1 101 88012 1
1006 1 101 88012 1
1008 1 101 88012 1
1002 2 101 88012 57
1003 2 101 88012 1.0
1004 2 101 88012 2350
1005 2 101 88012 1
1006 2 101 88012 1
1008 2 101 88012 1
1004 1 102 88012 1000
```

Other record structures (including the current ones) could also be agreed upon. The example showed already how checks on data-entry level can be included in the data model. In addition the data model of the FADN could include an entity type 'Check Point' with attributes like 'Formula', 'Reference value(s)' and 'Result of check point'. Income and other statistics can also be easily included in a data model. Take for instance the example of the indicator Family Work Units (FWU). This is clearly an attribute of the entity type Holding: a user is interested in the descriptor FWU at farm level. In other words: a farm is (among other attributes) described by the number of FWU. A description of this attribute would be:

number and name:	1015 NUMBER OF FAMILY WORK UNITS (FWU)
definition:	the total annual work units provided by regular working unpaid labourers.
comments:	see attribute 1003 for a definition on annual work units
format:	units in 3 digits, including one behind the decimal point
possible values:	0.1 - 99.9
obligatory/optional:	obligatory

formula:	derived by adding the values for the attributes 1003 of entity type Regular Worker on condition that attribute 1008 = 1. [a notation in mathematics would be possible too]
checks:	-

By including these derived statistics in the data model it includes all the definitions which are in use in the network. The decision to transmit this data can be separated from the logical model: one could decide in a later stage not to transmit and store the value of attribute 1015 (as it can quickly be calculated every time when needed) or to transmit it (to check calculations and transmission at different levels in the network). One could even decide to agree on a common definition of attributes like 1003, and decide to transmit only the aggregated attribute 1015, if details are not very likely to be used in analysis at higher levels in the network.

Including all the know-how in the network in an information model that can be stored in a workbench or database could have some additional attractive aspects. To support the development, dissemination and maintenance of information models a large variety of commercial workbenches are available. Some of them have facilities to generate software (databases, forms etc.) directly from the information model. If these facilities have grown mature in the coming years, the adaption of the information model generates directly a new release of the system. That will provide an enormous flexibility and cost effectiveness.

Such tools can provide users more easily with information on the meaning of the data - so-called 'Meta-Information' (Defays, 1993). That is becoming more and more important in an enlarged EU where information travels more easily. Where policy makers and researchers in the 6 founding member states were quite familiar with each other's farming systems, this is less and less the case in the EU of today and the next millennium. Another advantage is that the data model could be used as a support tool in translating activities. If the base model would be made in e.g. English or French, additional attributes could be incorporated to include names of entity types in other languages in special attributes. This leads to a thesaurus to be used in automatic translation.

The increased flexibility of the data model over the current approach becomes very clear at the moment that new data demands are dreamed up. Take for instance the proposal of a DG VI consultant (Kshatriya, 1994) to gather information on forms of pluri-activity. The information proposed to be gathered is structured once again in a complex table with about 85 fields. If this information is defined as an addition to the data model introduced above it means the following changes:

* an extra entity type 'non-involved household member' would be introduced to gather information on household members older than 16 years who don't work in the farm. Most of the attributes would be the same as those for Regular Worker, except attributes on work in the

farm (an experienced information analyst would thus introduce a so-called sub-entity-type).

* extra attributes would be created for the entity type Farm (reason for excluding data on pluri-activity, number of households per holding, number of household members over 16).
* 6 additional attributes would be created for the entity type Regular Worker (sex, education, hours worked in a self employed business not being the farm, type of that business, hours worked in an employment activity, type of that job).

This example shows how data models can be used to develop a system in modules which creates an enormous flexibility related to the current 'all or nothing' Farm Return.

7. TOWARDS A RICA - DATA-DICTIONARY / REFERENCE INFORMATION MODEL ?

Would information modelling be useful on an EU level (in general or for RICA)? That is one of the questions the first PACIOLI workshop discussed. Having listed all the advantages in previous paragraphs, this leaves the question: are there reasons why it will not work ?

In a brainstorming session the following potential problems were identified:

1. Europe has a large diversity of farming systems (which is however a problem for any information system).
2. Lack of funding: high costs are involved in creating information models.
3. Maintenance is necessary and difficult to organise.
4. There is a large variety in the quality of European information systems (but also a need for quality management).
5. Acceptance of a single model (with a need for consensus) could be problematic.
6. Theoretical issues and national attitudes towards information modelling techniques.
7. Existing systems (and investments) are a blocking factor.
8. Human factor: it is difficult to understand the information modelling technique.
9. Future is difficult to predict and the technological environment changes (although this leads also to a need for quicker software development).
10. Political resistance (including privacy laws) and unwillingness to exchange additional data.

The workshop also indicated potential solutions to overcome these potential problems:

1. Diversity is not a structural problem but a matter of costs and one could start with a more general model. The work on data dictionaries is carried out anyway in the member states. Diversity means also variation in data available, which implies that the focus of RICA could partly shift from a common data set in Brussels to access facilities to the total data set.
2. Funding is already done in member states. By better coordination duplication could be eliminated. In the case of RICA it could be aided by an EU-regulation for a new system. There is clearly an imbalance between the costs of collection and investments in analysis and making data available. Contract research and better marketing could increase

resources, although that would also imply more demonstration of data quality and representiveness.

3. Maintenance is closely linked to the previous point. It is essential to avoid obsolescence.
4. The variety in quality of information asks for more, not less standardisation. There is a need for pilot surveys and sample checks in case new data is collected. An interchange of data between member states will improve the quality of the data.
5. Acceptance is very important and should be organized. On the single model, it is suggested to stress the fact that it is not necessarily one model but a common framework to standardize definitions and procedures. Accounting institutes are used to increasing international standardisation.
6. Theoretical methods could perhaps partly be solved because the Commission already adopted a Euro-method. Problems could be passed to specialists when clear objectives and the time available would be framed.
7. To overcome the problem of existing systems, the information model should be a description of interfaces, not of systems themselves. With time this problem would fade away.
8. Human problems can be partly solved by training and by employing specialists as needed.
9. Change is seen as essential to avoid obsolescence and in fact this is the maintenance problem in another form. The uncertainty in future policies can partly be resolved by contracting facilities and people in stead of buying them. Needs like costs of production and estimation of future income are seen as rather stable. A frequent (2-3 year) monitoring of user needs by market research is also suggested.
10. Political resistance can partly be solved by implementing the model as an interface and not as a prescription for a national system. Variable geometry could be use here: exchange with those who are willing to do so. Institutional aspects (like privacy regulations) can partly be solved by technology. It could also be helpful to use the RICA-payments by differentiating them to the quality of the data delivery (amount of detail, being on time etc.), although some RICA liaison offices are not very motivated by the amount of money that flows into the national treasury. Discussions to find simple solutions and to develop communication should be promoted.

Balancing the advantages, problems and potential solutions and taking into account the urgent need for more flexibility in the Farm Return and the available technical tools to create this, it looks worthwhile to launch a short study to see if more flexibility can be created in the Farm Return. This flexibility could be based on a data-dictionary containing a data-model with the entity-types and attributes of the current Farm Return and that can be widened to incorporate additional data of interest for e.g. pluri-activity, cost of production, environmental issues, forestry etc.

This approach of flexibility would break the trend to make the gathering of all data items in the Farm Return obligatory. It potentially facilitates a situation where the Farm Return would make a voluntary exchange of data possible on all the data that are gathered in national farm accountancy data networks anyway, and that would (in a harmonized methodology) be useful to the users.

In fact part of the description of each attribute in the data model is whether it is obligatory or not and basically it is possible to define who records a certain attribute and who not.

For instance it is known that some RICA-partners exchange aggregated data on gross margins per arable crop through a Paris-based organisation called IAGC to improve estimates of cost of production. Another example is data on non-farm income which is available in some member states, and where the RICA-committee now works on a voluntary exchange.

In the longer run one could imagine a situation where a huge virtual logical data base exists in which all the relevant data that are available in the Member States is described (figure 7.1). Within the data dictionary of the data base the hard core would be the current Farm Return that defines the data to be provided by all member states on all farms in the network.

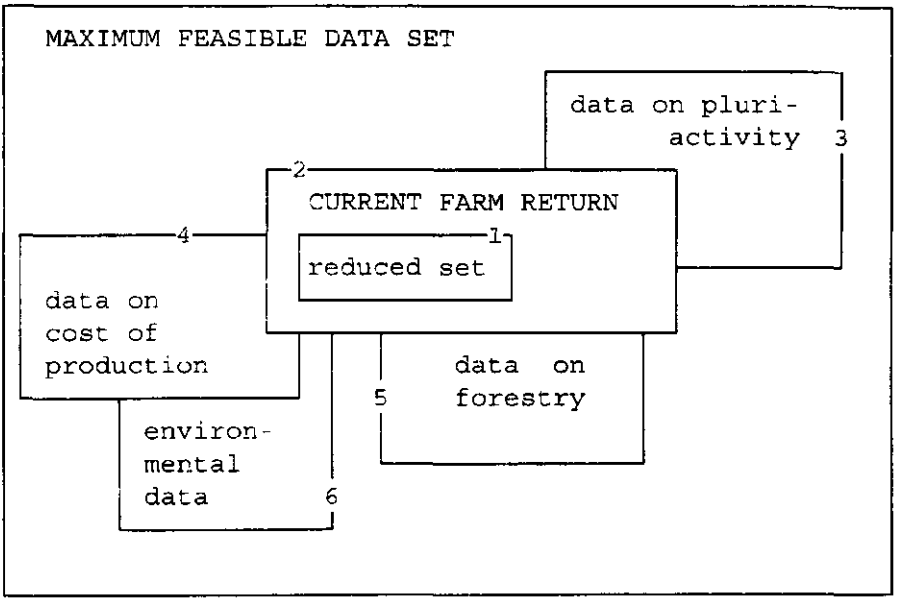


Figure 7.1 Data dictionary / logical data base oriented to a more flexible exchange of data

Perhaps that it is even possible to reduce this data set (e.g. by omitting details in table K) for application in Central-Eastern Europe and for speeding up data delivery in Western Europe. In addition to the current Farm Return additional data sets would be identified that some member states are willing to exchange (e.g. on a voluntary basis) on some farms. The fact if data are transmitted in advance (creating data bases in several member states) or can be retrieved on the moment of demand by using the data dictionary becomes a technical and juridical issue that can be settled on attribute-level and that can be recorded for each attribute in the data-model itself.

A study to investigate if such a flexibility is feasible and if a data-model, and the physical data dictionary based on it, would be helpful to create this flexibility should not only incorporate discussions with the member states, but also with the users of the RICA data in general and with those in the European Commission in particular. Information modelling is not only a user-oriented approach, but it is also important to involve the users that in the end will have to pay (part of) the cost of the exercise.

This suggests that the proposed study should incorporate the following topics:

+ the interest in member states and with the Commission to improve flexibility in the Farm Return in order to introduce statistics on new policy fields.
+ the interest to use a more up to date data modelling technique to improve data exchange and software development
+ the selection of a work bench (or comparable tool) to support the development of a data model / data dictionary
+ the conversion of the current Farm Return to the format of a data model
+ a test of the flexibility with data on gross margins (cost of production) or pluri-activity by incorporating data definitions used in member states in the data model.

In a follow-up project the methods developed (and reviewed) could then be introduced in the RICA-unit, the member states and could be used to standardise new data definitions and to exchange data. The example in section 6 and the experience of the first PACIOLI workshop show that such a study would not result in much extra work for the unit A/3 if it is carried out by experienced members of the RICA-committee.

8. CONCLUDING REMARKS

This reflection paper provides suggestions for decision making on the further development of the European FADN, based on papers of the first PACIOLI workshop. Although information modelling has some risks, it seems to be a promising tool. By creating a data model or data dictionary it could improve the flexibility of the RICA-system. This flexibility is necessary now that policies are changing and the EU is widening.

It is suggested to launch a short feasibility study to check the results of this first workshop and its reflection paper in all the member states and in the Commission. This should incorporate a stock taking of opinions and show the advantages of the technique by its application on the current Farm Return and a policy topic on which new statistics are needed. An appropriate name for such a study would be RICASTINGS: RICA's Study To Install a New Generation of Statistics.

REFERENCES

Beers, G. (1995)
Introduction PACIOLI; in G. Beers et al. (1995a)

Beers, G., K.J. Poppe, D.F. Spiering and H.C. Pruis (1995)
PACIOLI 1 Farm accountancy data networks and information analysis; Workshop report; The Hague, LEI-DLO, 1995a; Mededeling 532

Beers, G., K.J. Poppe, D.F. Spiering and H.C. Pruis (1995)
PACIOLI 1 Introduction and information modelling; Summary; The Hague, LEI-DLO; 1995b

Bonati, G. (1995)
Trends in information technology and their implications for agriculture; presentation at Ameland, March 1995

Kshatriya, S. (1994)
Pluriactivity and non-farm incomes of agricultural households; Brussels, EU-DG VI-A/3

Defays, D. (1993)
Statistical Meta Information Systems Workshop; Luxemburg, Eurostat

Enroth, A. (1995)
Economic planning and monitoring on Finish farms; in: G. Beers et al., 1995a

Graumans, C. (1995a)
Information engineering: a short introduction; in: G. Beers et al., 1995a

Graumans, C. (1995b)
Information engineering: Dutch experiences; in: G. Beers et al., 1995a

Martin, M. and J. Leben (1994)
Strategic information planning methodologies; New Jersey

Persson, P. (1995)
Defining information requirements; in: G. Beers et al., 1995a

Poppe, K.J. (1991)
Determining farmers' financial information requirements; The Hague, LEI-DLO; Mededeling 444; Reprinted in G. Beers et al., 1995a

Poppe, K.J. (1995)
RICA's farm return: introduction and comments; in: G. Beers et al., 1995a

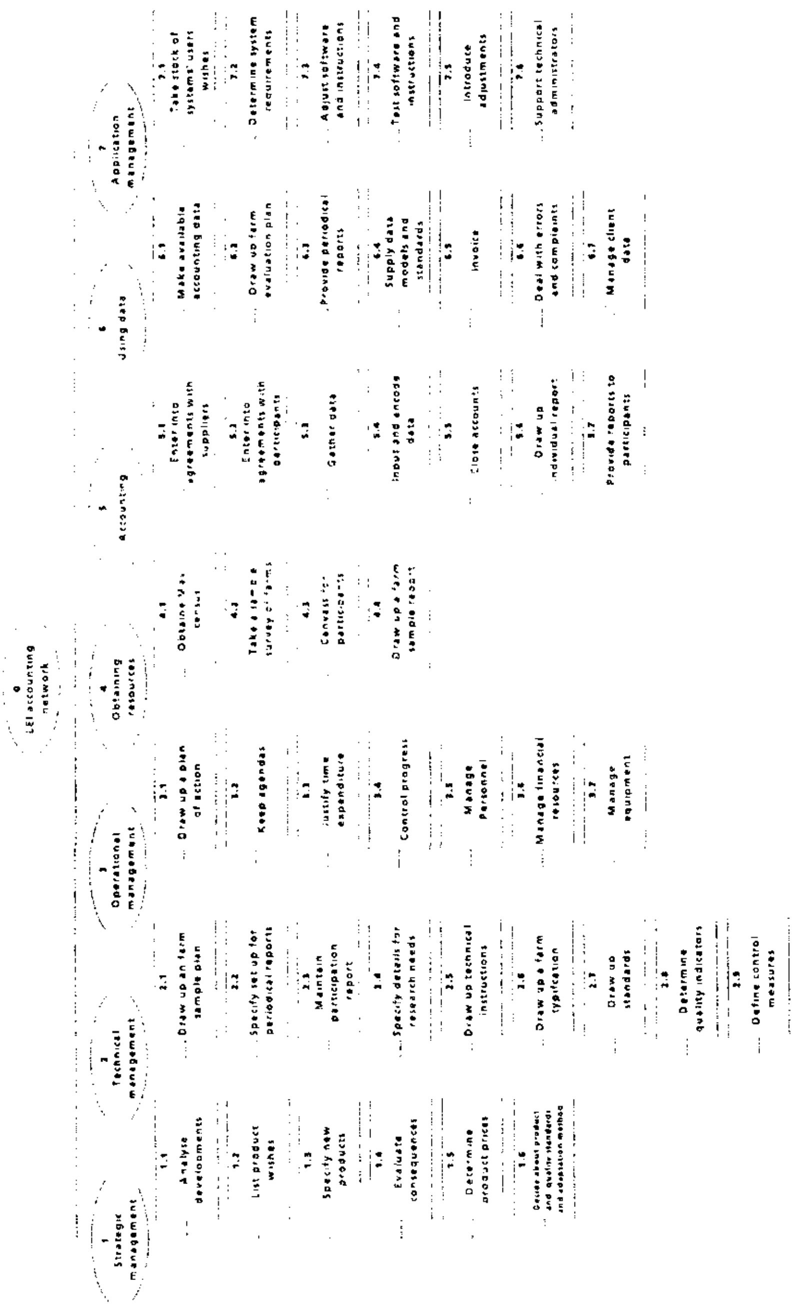
Power, R. et al. (1989)
Harmonisation of the FADN Farm Return; Dublin, Teagasc

Verwaart, T. and D. Spiering (1995)
LEI accountancy 2000; in: G. Beers et al., 1995a

Williams, N., A. Bailey and S. Dedman (1995)
Information disparities in the FADN/RICA - Causes and consequences; in: G. Beers et al., 1995a

ANNEXES

Annex 1 Process model Dutch FADN



Annex 2 Data model Dutch FADN (simplified)