A Framework to Integrate and Analyse Industry-wide Information for On-farm Decision Making in Dairy Cattle Breeding

by

Alfred Ainsley Archer

Department of Animal Science

McGill University, Montreal

July, 2000

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Doctor of Philosophy

0-612-69849-1

Abstract

Alfred Ainsley Archer
Department of Animal Science

The goal of this thesis was to develop a framework that could integrate and analyse industry-wide information for the support of on-farm decision-making in dairy-cattle breeding. Specific objectives included i) describing a dairy breeding information system (DBIS); ii) examining how the Internet could be exploited to improve the DBIS and its functioning; iii) describing a process for implementing a unified data model to facilitate integrated user access to information in the DBIS; and iv) developing software to support decision-making by facilitating access to a unified data model when implemented as a database management software.

The first objective was achieved by following a systems approach – defining a goal, boundary, functions, structure and performance – to describe multi-organisational information systems and, specifically, a DBIS in the Canadian dairy industry. Using this framework, the subsequent analysis of the DBIS looked at its overall effectiveness. The DBIS was also compared with other known systems, where the number of participants (as well as their roles) differs from the Canadian situation. Improvements were suggested for the Canadian DBIS by focussing on the decision-maker's ability to retrieve, integrate and consider required information through information technologies. The second objective involved using the systems approach to investigate the kinds of information (if any) provided on Web sites of the DBIS participants, and to see if the Internet could be exploited to improve this process, either in terms of improved transfer speed or data transformation. It was established that the Internet is being used for rapid, flexible access to support information by DBIS participants, but that it is being under-utilised, particularly where herd output information is concerned. Herd output information could be filtered, integrated and transformed to support specific user needs at appropriate levels of intelligence density. It was further postulated that these data could be exploited more effectively through the use of such information technologies as common data exchange mechanisms and decision-support systems. The third objective was achieved through applying information engineering methods to develop a data model to represent the DBIS. This unified model was described in conceptual, logical and physical terms, and facilitated transparent access for on-farm users to information from more than one source organisation. It was demonstrated that such a model could maintain the autonomy of participating organisations while simultaneously creating an amalgamated databank for decision support. The final objective lead to the development of a prototype user interface called DAIRIE: DAiry Information Retrieval and Integration Expert which could interact with the physical schema of the unified data model previously developed. The interface consisted of data selection, aggregation and display forms, and allowed dynamic SQL query generation for transparent information retrieval to support decision-making. Knowledge was employed in facilitating user access to information as well as its presentation and interpretation. The approach is modular and, therefore, flexible in terms of future additions and improvements. The prototype shows that there is potential for creating data driven systems that can satisfy individual uses and preferences for herd output information.

Résumé

Alfred Ainsley Archer

Département des sciences animales

Afin d'étudier l'utilisation de l'information qui est disponible aux gestionnaires de fermes laitières, pour des activités de reproduction, une structure a été développée dans le but de comprendre, organiser et accéder à l'information de façon intégrée.

Les objectifs de cette recherche étaient; de développer une méthode qui décrirait un système informatique de reproduction laitière (SIRL), à dimension industrielle; d'examiner l'exploitation de l'Internet quant à l'information sur la reproduction, en utilisant l'analyse produite par le système; de décrire un procédé pour l'implantation d'un modèle de données unifiées, qui faciliterait l'accès aux usagers du SIRL; et de développer une *logiciel-prototype-médiateur* (prototype software interface) au système implanté SIRL, pour faciliter la prise de décisions des gestionnaires sur les fermes.

L'approche d'un système informatique a été utilisée pour déterminer un SIRL afin de soutenir les gestionnaires sur la ferme-même et d'identifier comment la technologie informatique telles que EDI, systèmes de gestion de données, et un soutien basé sur la connaissance, peut améliorer le système en question. L'Internet est exploité par le SIRL et peut avoir un impact sur l'efficacité de par la réalisation d'une "jungle informatique", mitigée à travers les services axés sur le client et à un accès informatique intégré; les fonctions peuvent être plus globalisées et soutenir des niches plus spécifiques; et les secteurs devenir plus intégrés. L'impact à long-terme sera la transformation de l'Internet, d'un simple médium véhiculant l'information, à une assistance lors du traitement informatique en utilisant les technologies, telles que applets, des systèmes de soutien décisionnel et d'échange de données électroniques. Le procédé du développement des modèles de données mène au développement de modèle conceptuel, logique et physique.

L'outil prototype, soit l'Expert en extraction et intégration de l'information laitière (EEIIL) a été développé et possède des avantages permettant d'améliorer, dans la boucle décisionnelle, la composante de l'intervention humaine. Les résultats des objectifs finals ont

démontré les composantes de la médiation *(interface)* des usagers du SIRL (formule de sélection, formule d'agencement de la dimension des données et formule d'affichage de données et les sous-formules à l'intérieur de ses principales formules) et les modules de développement des SQL dynamiques (module SQL, module Multiview).

There are those who seek knowledge for the sake of knowledge, that is curiosity.

There are others who desire to know in order that they themselves may be known, that is vanity.

Others seek knowledge in order to sell it, that is dishonourable.

But there are some who seek knowledge in order to edify others, and that is love.

Bernard of Clairvaux (1090 -1153)

Table of Contents

Preface

This thesis has been prepare in accordance with the October, 1999 revision of the *Guidelines for Thesis Preparation* (Faculty of Graduate Studies and Research, McGill University). As stated therein:

> "As an alternative to the traditional thesis format, the dissertation can consist of a collection of papers of which the student is an author or co-author. These papers must have a cohesive, unitary character making them a report of a single program of research...The thesis must be more than a collection of manuscripts. All components must be integrated into a cohesive unit with a logical progression from one chapter to the next. In order to ensure that the thesis has continuity, connecting texts that provide logical bridges between the different papers are mandatory.
>
> "In general, when co-authored papers are included in a thesis the candidate must have made a substantial contribution to all papers included in the thesis. In addition, THE CANDIDATE IS REQUIRED TO MAKE AN EXPLICIT STATEMENT IN THE THESIS AS TO WHO CONTRIBUTED TO SUCH WORK AND TO WHAT EXTENT...The supervisor must attest to the accuracy of this statement at the doctoral oral defence. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to clearly specify the responsibilities of all the authors of the co-authored papers."

In keeping with the guidelines, in particular the above excerpts, the contributions made by various authors to chapters in this thesis are detailed in the following paragraphs. Chapters 1, 2, and seven, as well as the prefacing, connecting texts and bibliography were prepared entirely by A. A. Archer. Dr KM Wade provided editorial input and guidance for the entire thesis and general guidance. Through extensive consultation Dr Wade also contributed to the systems analysis in chapters three and four from a domain expert's viewpoint. Dr Wade assisted in Chapter 6 by debugging, and advising on suitability of output from the software.

Dr. R. Lacroix contributed to chapters three and four to the development of the information systems approach through extensive consultation. In the preparation of chapter 5, Dr. Lacroix provided guidance with the development of the data model particularly with the metafile, entity-relationship model, and precursor code adapted for the software (particularly error trapping and SCC module and form).

Data obtained and used in the work presented in chapter 5 and 6 were pre-treated by Mr Diederik Pietersma.

All other work in chapters 3, 4, 5 and 6 was completed by A. A. Archer.

Acknowledgements

I would like to express my appreciation for all those who played a part in the accomplishment of this work. Firstly, I would like to thank Drs K M Wade, R Lacroix and R I Cue, who, with much longsuffering, guided (and at times prodded) me through every stage of my studies.

This work was made possible through the financial support of National Sciences and Engineering Research Council of Canada, les Fonds de l'Autoroute de l'Information du Gouvernement du Québec, and the Cattle Breeding Research Council of Canada. Data used in this research were provided by the Programme d'analyse des troupeaux laitiers du Québec, Holstein Canada and the Canadian Dairy Network.

I would like to acknowledge the staff and students of the Department of Animal Science: the support staff who assisted in so many ways, the faculty and postdoctoral students who were ready to answer questions and point me in the right directions, and the students who made my time in the department more than just work.

Special thanks to Diederik Pietersma, Madhu Nagaraja, Mark Strasser, Xingzhu Yang, my colleagues in the Animal Science Information Systems Group and my constant companions in the "dungeon," for all the good times and your valued support. I would especially like to honour the memories of Ken Foran and Mark McBratney; your friendship and sincerity were encouragements to me; you have been and will be missed.

I would like to express my thanks to the Lord who gave me the ability to pursue these studies and the grace to arrive at the goal. Thanks to my family, in particular my mother and father Nerissa Archer and the late Desmond Archer, upon whose shoulders I stand: their love and provision have shaped and equipped me. Finalement merci à mon épouse bien-aimée, France: ton amour a effacé toutes les épreuves.

List of Tables

List of Figures

Chapter 1. Introduction

Profitability gains for dairy farms can be improved by making the best use of information available to on-farm managers for breeding decisions. These decisions include culling, sire selection, mating decisions and replacement selections (Biron, 1998). Because these decision activities are not exclusive from each other and may use overlapping information from several sources, the decisions made tend to be complex.

It has been theorised that dairy farmers can be assisted in the decision-making process through the development of decision-support system (DSS) information technologies. These have typically taken the form of knowledge-based information technology requiring clearly structured specific information which was known before hand and expected by the knowledge-based system and a specific planned approach to the use of that knowledge, based on the expertise of one or more experts. Some prototypes have been developed to support decision making in dairy cattle breeding (Strasser, 1997, Strasser et al., 1997). The development of such an integrated decision-support system for the breeding domain was a major objective of this research effort.

Developing successful decision-support solutions for dairy managers in this domain has proved difficult because of the nature of the decisions being made and the nature of dairy breeding information systems. Managers draw on knowledge from various sources (including their own experience), some of which can conflict (Strasser, 1997), making the decision process on each farm potentially quite unique. They may include information from many organisations upon which the user is dependent. While going through the decision-making process a manager may cover a broad spectrum of time ranging from an individual test (which could be considered the smallest resolvable unit of measure where data are collected) to life of an animal or of a whole herd. The information system includes several different types of organisations that provide information to dairy farm managers. Also, the managers are independent entities in the system: they receive information from many organisations and may conceptualise various means of employing it. Profitability can be achieved in the dairy business through different means and individual farm managers may select different strategies to achieve success (for example low input - low output, high input - high output, etc.). These can all be considered successful and profitable strategies by various farm managers. Thus, while a DSS may be robust in its use of information, and represent the reasoning of a particular group of (successful) managers, it could conceivably not be adopted through a failure to represent the problem or the advice in terms that managers will accept, or because it includes or excludes information that many managers may or may not identify as being important (Newman et al., 1999). The problem is one of trying to include all sources that provide relevant information to the area being addressed

Efforts to automate the utilisation of information resources to support on-farm decision making have met with mixed success. Newman et al. (1999) highlight several aspects of DSS development which contribute to this failure, including unclear definition of beneficiaries, the inability of end users to relate to the output of the system and difficulty using the system. Assumptions about a "typical" user and the "typical" information they should select and what conclusions they should draw, based on expert knowledge, often excludes groups of users and are rarely applicable (i.e., this strategy of development may exclude many users by presuming that there is *one* typical sort of user with *one* way of employing information for predictable patterns of decision making). Such approaches may lead to successful modelling and prototyping but may fail because they are too narrow to support a large enough proportion of users. On the other hand, a DSS, broadly defined to model the expertise of most users and taking account all possible information will be unable to model the domain successfully. Even if it were possible to provide for every eventuality, each manager weights or sorts the information selected differently. A single expert may model situations on the frontier of his experience poorly. As an example, using multiple experts to recommend culling has been tried (Strasser, 1997); however experts may rank culling candidates differently which may create uncertainty or mistrust in the end user.

To avoid such failures a fundamentally different way of supporting users should be developed. In order to permit farmers to employ familiar methods of decision-making it should direct the information they choose as little as possible, but should be able to support their needs as much as possible. This method should facilitate the manipulation and integration of the information farmers require based on an understanding of the information system in which they operate. It should encompass a broad scope of the breeding domain to support as many users as possible and employ knowledge to support their choices of information for specific breeding activities, not by constraining the information considered, but by elucidating relevant characteristics of the information selected.

Decision support should start with an appropriate comprehension of the information systems in which managers operate as a critical step to building decision-support structures which can effectively support their information needs. On-farm dairy information systems are complex and can vary greatly among farms (Lacroix et al., 1994). While information can flow in both directions between the farm and off-farm information sources, not all farms employ information from every source. A model of the information system underpins decision-support automation since it spans the biological system and the model of a decision-making environment. Information also flows among sources as well. Given an understanding of the information system, the decision-support systems that are developed may be able to be more consistent with the goals of a system.

The advent of computer and telecommunications technology makes possible network solutions for distributed information systems. Electronic technologies have become commonplace for supporting on-farm management information systems. Dairy-breeding decision making could be supported over a network like the Internet: it represents a potential means of exchanging the information in the system and improving some systems attributes. For dairy managers the interface to the networks available to them is critically important since it should facilitate their efforts rather than deter them.

Users may be hindered rather than helped by the proliferation of electronic technology to support them. Accessing information through networks may require expertise in computer hardware and software applications, which can challenge, frustrate or discourage some people. Information represented differently by each organisation in the system may force farm managers to develop multiple means of storing and accessing for the information they require. Incompatible formats may create confusion where similar terms are used to mean completely different things or different terms mean the same thing. Even high quality information may not bring clarity if the technology to deliver it is more visible than the information it is trying to support.

A method of decision support that is driven by the data selected is atypical of most knowledge-based DSS methods. Incorporating the data to maximise the freedom of dairy farm managers to choose the information they consider important would require that the DSS respond dynamically to the information selected by managers and also control the information selected. However, knowledge-based methods (including blackboard systems and Executive Information System – EIS) have been developed and can be applied. A blackboard system is a knowledge-based system that uses opportunistic reasoning of independent knowledge sources and can view a shared data space (the blackboard) while contributing towards a complete solution (Englemore et al., 1988). The information is structured within the blackboard and an overall control mechanism determines how the knowledge sources will act in the system to arrive at a solution. An EIS is a single database or system of databases that supports a user interface where knowledge is provided at the front-end to support key enquiries into the functioning of that enterprise (Dhar and Stein, 1997; Young and Watson, 1995; and Nord and Nord, 1995). Executive information systems presuppose that the user of the system is a manager in a system requiring large quantities of internal and external information that can be easily manipulated to address important issues in achieving the goals of the organisation (Nord and Nord, 1995). While preliminary work was performed to investigate how a blackboard approach could be used, EIS would seem to conform quite well to important factors in the goal of this thesis – namely the distinction of the independent decision-making role of on-farm managers, the use of database structures (which can be used to support information systems flexibly) and the data-driven approach to decision-making.

The objective of this project was to develop a way to support on-farm breeding decisions using information and knowledge available from the industry irrespective of origin (on- or off-farm). The expectation was that this research would serve as a methodological framework for developing decision-support systems for on-farm dairy managers, requiring information and knowledge from various, distributed industry-wide participants, to assist in on-farm breeding activities based on user selected information.

The study was conceived of as requiring three stages. The first stage should explore and understand the nature of information systems in dairy breeding. A suitable method for accomplishing this was required. Thus the first part of this stage was to develop and use a method for analysing industry-wide dairy-breeding information systems. The specific objectives were i) to extend the systems approach to analyse a multi-organisational information system – specifically to describe its goals, boundaries, functions and components; ii) to use this approach to analyse a dairy cattle breeding industry system; iii) to discuss its performance from a specific point of view using important attributes; and iv) to recommend how it might be improved.

The second study was concerned with investigating the use of the Internet in dairy breeding activities. Any attempt to provide information to on-farm users from many off-farm organisations must rely upon some common means of communication. The Internet can facilitate information exchange in dairy breeding and many participants in the dairy industry are experimenting with it. The specific objectives were to examine the deployment of a DBIS over the Internet; to study the impact of the Internet on the DBIS; and to investigate means to improve the DBIS through better use of the Internet.

Once an understanding of information systems in dairy breeding was achieved, the next project was to develop a way to facilitate information availability, improve access, clarity and utility, and to link information products across organisations in the system. The objective was to describe a unified data model as a representation of an industry-wide dairy breeding information system, which could be implemented as a database management system.

The goal of the final project was to develop a prototype software system, capable of integrating user requests for information from potentially many sources to support on-farm decision making through dynamic information retrieval from the unified data model of the DBIS. A front-end graphical user-interface that interprets user information needs to support on-farm breeding decisions was required to employ domain knowledge and create views of information from a database that represents the Canadian DBIS.

Chapter 2. Literature Review

Introduction

This chapter reviews the literature that concerns the central investigations of this research: namely decision-making for dairy-farm management, decision-support and data-management technologies, and information exchange technologies (including the Internet). There is a large body of literature that pertains to these areas and this provides a context for understanding several important issues that relate to this general research area. How information can affect decision making is described. It specifically discusses the types of decisions that are important in the area of dairy cattle breeding, organisations that support them and some of the information they provide. How information systems have been thought of in the agricultural domain is examined especially in terms of information technology developed to support them. The general use and development of decision-support systems in agriculture is explored. A conceptual view of networks is considered, which is followed by a consideration of information exchange among dairy participants, including how the Internet has been employed to accomplish this and what distributed systems should be able to accomplish.

Dairy Farm Management and Decision Making

Dairy producers are autonomous managers and decision-makers; they ultimately control the decisions made that could affect on-farm production. Good decisions require sound information as a foundation. The on-farm management environment is complex with many things affecting the decisions taken. Each dairy manager uses information uniquely to satisfy decision making goals particular to his circumstances.

The quality of the decisions made is affected by uncertainty and equivocality of knowledge: uncertainty from a lack of information and equivocality is ambiguity from conflicting interpretations (Goodhue et al., 1992). Equivocality is defined as uncertainty (Daft and Lengel, 1986). Completeness of knowledge is considered a trait of declarative knowledge while uncertainty is a property of procedural knowledge (Hogeveen et al., 1994).

Complexity can increase with the amount of information produced (Goldsmith et al., 1996). The complexity increases the need for sharing information resources (Malone and Crowston, 1994; and Lacroix et al., 1994).

On-farm decision-making relies on many sources of information. Records have been created to evaluate animals, diagnose problematic areas and improve management strategy. Various operations within the system create information products and information flows among the operations within the system, as well as off the farm (Pietersma et al., 1998). Records collected for one purpose can be used to assist decision making in another area of production (Etgen et al., 1987).

Higher integration of information in the system may allow managers to make better decisions due to their interdependence on information from off-farm sources. Useful records to on-farm decision-making include maintenance practices and records of performance (Burnside, 1993 and Etgen et al., 1987). Most dairy producers rely to some extent on information from off-farm sources. Increased dependence on information creates a greater need for information (Goodhue et al., 1992) and this could be true in the dairy sector. The origin of much of the information that flows back to the farm is on-farm records (Christensen and Fehr, 1993; Groeneveld and Lacher, 1992), but managers often use assistance from off-farm organisations in decision making. Furthermore one should recognise that the system is tied together economically and biologically. Goodhue et. al. (1992), however, suggest that it is wrong to assume that data integration is always good or should be considered an end in itself. Developing a system for maximal data integration can create problems functionally (Hall, 1991) and may make the development of the system untenable.

Breeding activities of on-farm managers: An adequate breeding strategy is a balance of replacement and culling, sire selection, and mating to maximise the advantage of available genetic material (Biron, 1998; Burnside, 1993; Christensen and Fehr, 1993; and Lacroix et al., 1994). It requires sound information from the herd and organisations controlling genetic information on animals. Optimum cow culling should be about 25-30% of a herd (Etgen et al., 1987). Quebec has had a replacement rate of about 36 - 41% and these figures are not atypical of other replacement rates in Canada (Martin, 1995; and Pellerin et al., 1995). High replacement rates limit choices managers have among heifers and may limit genetic progress (Etgen et al., 1987; and Martin, 1995). This may influence the choices a manager makes in selecting sires with "stayability" for the herd. Biron (1998) describes a breeding strategy, aimed at consistent genetic improvement, based on replacing 25% of herd members. Cows remaining in the herd should be classified according to their potential for improving the herd genetically: superior (35%), intermediary (25%) and poor (40%). Superior cows will be bred to elite sires, the intermediary cows bred to young sires and the poor cows should be bred using low-priced sires. Sire selection towards genetic progress in the herd for economically important traits and limited impact of traits that limit production performance; however the manager ought to consider the prices of elite semen verses the value of young sires and the costs of lower ranked sires. The proportion of cows in each group depends on participation in young sire programs, the breed of dairy cattle, and replacement rate (Biron, 1998).

Off-farm dairy industry organisations involved as information sources: Breeding decisions may rely on many off-farm organisations which are involved in managing and adding value to information in order to provide various services to on-farm managers. Organisations store and control records from the farm, usually managing them on behalf of on-farm clients, providing performance measures (Appleman and Noble, 1997). Most of these organisations add value by applying their knowledge and expertise and the resulting information products are predictive, diagnostic or prescriptive in nature. The utility of the information produced has varying time-horizons. In Canada the major off-farm organisations involved in dairy cattle breeding include milk recording, breed associations, artificial insemination organisations (Biron, 1998; Burnside, 1993; Christensen and Fehr, 1993; Canadian Dairy Dictionary, 1989; and Lacroix et al., 1994).

Milk recording organisations collect milk production data and may also collect other management data such as breeding, feeding, culling reproductive and disease reporting records. Sixty percent of Holsteins in Canada are enrolled in milk recording (Biron, 1998). Most of the information is descriptive of what is happening in the herd at the time the records were made (Appleman and Noble, 1997). Configuring the information to reflect different time periods allows different decisions to be made. Managers can use test-day information for short-term decisions on groupings, feeding, breeding, drying off and selling dairy cows. Lactation-to-date, mature equivalent production and persistency information allow managers to evaluate and adjust short-term management strategy in feeding programs or breeding performance for example. Long-term information (e.g., multiple lactations, 12, 24, 36 month herd milk production comparisons) allows the analysis of overall management strategies such as sire selection, culling practices, genetic progress, and service/conception rate (Appleman and Noble, 1997).

Breed associations contribute conformation data obtained from farms and contribute to the production genetic information for these traits on sires and cows. Individual breeds establish traits and standards consistent with breed objectives for type improvement. Thirty-six percent of Holsteins animals in Canada are enrolled in this program (Biron, 1998). A breed association also disseminates genetic information related to that breed. To the extent that a producer's income is dependent on breeding true type animals (e.g., for selling breeding stock or avoiding disease), short-and long-term decisions can depend to a great extent on type information.

Artificial insemination organisations provide information to producers on sires' value for making genetic progress of the herd and exact information on breeding records without which reliable genetic evaluations would not be possible (Biron, 1998). Eighty percent of Holstein animals are bred through AI organisations (Biron, 1998). Expertise of AI organisations technicians may be made available for long-term, decision-making on farm.

Genetic evaluations are produced through genetic evaluation centres. Managers depend on genetic information for long-term productive gains through selecting and culling cows and for sire selection and use in their herds. Selection for genetically superior traits is simplified by the use of aids such as cow indexes, total economic value index and herd inventories (for protein, protein percent, milk, fat etc.) (Biron, 1998; Burnside, 1993; and Christensen and Fehr, 1993). The organisations that participate in the production of genetic evaluation provide expert advice to managers for on-farm breeding activities. Their emphases vary and managers do rely on them.

Information Systems in Agriculture

An information system can be defined as a set of parts that is particularly concerned with the maintenance, exchange and production of information to the benefit of its parts individually or corporately towards a specific end. In situations where resources are shared in a coordinated effort, they must be managed to best achieve the over all end of the system (Malone and Crowston, 1994). The terms information system (IS), management information system (MIS) and computer information system are used interchangeably (Ahituv et al, 1994; Barrett and Konsynski, 1982; Dickson et al, 1977; Groeneveld and Gutzmann, 1990; Groeneveld and Lacher, 1992; Huber and McDaniel, 1986; and Jofre-Giraudo et al., 1990). Some justification for this exists since technologies that serve ISs do not exist entirely independently of them and do manifest them. However, an IS, as defined above, is not a software program nor an electric mechanical apparatus, but an entity that exists at the level of information transaction. The literature may describe an IS in terms of its expression rather than its essential nature, possibly because this vantage point is more comprehensible from the perspective of the user. It is reasonable that an IS should be explored and understood for its own sake in order to employ technologies to exploit it. The following sections look at ISs in agriculture and some information technologies which aid decision making.

Information systems have been described in dairy and other animal agricultural industries to assist aspects of on-farm decision making. Nelson (1994) recognises ISs in agricultural development and describes networks which enhance the potential for exchanging information products in them. Kok and Gauthier (1986) describe an expandable database management system (DBMS) prototype for on-farm cropping systems management. They deal with the issue of the organisation and storage of information required for gleaning knowledge from diffuse and interconnected areas of farm management. In their work, developing a DBMS for dairy cattle feeding and management, Lescourret et al. (1993) emphasize the underlying data model and the modelling process. Also using database technology, Lescourret et al. (1994) describe a DBMS for intermittent data supporting feeding regimes as end applications. Computer-based processes in ISs perform transactions that are routine, simple and well modelled involving large quantities of data that may be difficult for an individual manager to process consistently and easily (Dhar and Stein, 1997). The development of a structured framework to exchange agricultural information using specific protocols and formats, which can be easily interpreted by the end-user, can help the decision-making process (Pietersma et al., 1998; and Spahr, 1993). Automated record-keeping computer systems for the exchange of dairy data in management activities, integrated swine management using MIS and integrated data recording and processing facilities are examples of this (Dijkhuizen et al., 1996; Groeneveld and Lacher, 1992; Tomaszewski 1993; and Verstegen et al., 1995).

The importance of these studies is partially the IS manifested by computer ISs however the explicit recognition of the IS varies greatly from study to study. For example, Nelson (1994) explicitly recognises ISs in agricultural production in an attempt to discuss methods of implementing technology to support them. At another extreme, McKendrick et al. (1994) and Revie et al. (1996) represent ISs as computer based tools when describing hybrid information systems (a type of software system combining different information technologies) used to aid animal production in sub-Saharan Africa. In between these two positions Jones and Spahr (1991), in describing a natural language tool to simplify information retrieval for dairy producers from a database, implicitly recognise an IS by the domain semantics used by producers which is separate from the software components.

As more uses are found for computer technologies in animal agriculture, dependency upon information technology increases in ISs. Dependence on computer-based technology grows in tandem with the need to manage the complexity and interdependence of information (Greer et al., 1994). Christensen and Fehr (1993) consider computer aids essential in assisting farmers. Misztal and Lawlor (1999) and Tomaszewski (1993) similarly describe many potential uses of the information technology to improve on-farm decision making, which may serve to make management complex because of the many systems that must be simultaneously consulted.

Management information systems in the dairy industry: Management information systems can be useful for improving decision-making. A management information system (MIS) is a tool to facilitate the manipulation of the information products in an IS. A MIS provides an interface that is flexible and user friendly (Goldsmith et al., 1996) and a database management system (DBMS) which stores, accesses, retrieves and manipulates data more efficiently and powerfully than flat file computer storage or hard-copy filing systems (Groeneveld and Gutzmann, 1990). Information is organised and indexed to support retrieving information independently from how it was stored. Goldsmith et al. (1996) describes MIS as having economic value and this value contributes to the economic value inherent in the overall system. The DBMS can be implemented in distributed architecture (Codd, 1990; and McFadden and Hoffer, 1996) so the MIS can include many participants in the dairy industry that belong to the IS providing the network architecture permits it. Management information systems improve farm management practices (Jofre-Giraudo et. al., 1990). Thus users can spend less time on data manipulation tasks and more on actual decision-making. Kroeze et. al. (1996) and Lokhorst and Kroeze (1996) present a common MIS in the Netherlands (UNIFORM) which dairy farmers use to organise information for decision making relating to many aspects of herd life including production, feeding, and economics.

The DBMS is dependent on a data model (Harrington, 1998; and McFadden and Hoffer, 1994). Lescourret et al. (1993) and Lescourret et al. (1994) have described data models to implement DBMS for on-farm management activities. They describe the information available and develop conceptual and logical models using an entity-relationship approach. What Lescourret et al. (1993) call the conceptual model is also known as the conceptual schema of the data model (Harrington, 1998), and is the basis of the logical model.

Conceptual modelling describes the domain in terms that are independent from a DBMS (Graves et al., 1996). The logical model describes the IS domain in terms that are useful for ISs experts for implementing in a database. Probably the best known model is the relational model introduced by Codd in 1970 and the entity-relationship model is similar and also very common (Codd, 1990; Harrington, 1998; Lescourret et al., 1993; Lescourret, 1994; and McFadden and Hoffer, 1994). The relational model is similar to the mathematical concept of relations as sets that share common characteristics. Other models such as the hierarchical model (Groeneveld and Gutzmann, 1990; and Harrington, 1998;) and the network model (Harrington, 1998) have been developed. However the relational model remains widely accepted and is readily implemented particularly in distributed systems (McFadden and Hoffer, 1996). Lescourret et al. (1993) used the entity-relationship model to describe a dairy production and health DBMS design. Lescourret et al. (1994), used the same logical model to formalise the data structures of the data model as a logical schema. When the system has been developed as a DBMS this is known as the physical model.

The front end of the MIS facilitates access to the data. A standard means of accessing relational DBMS is structured query language (SQL) (Harrington, 1998, McFadden and Hoffer, 1994, and Microsoft, 1997). SQL is a high level language, however the user must be familiar with its use and the data model for it to be used effectively to support decision-making in a MIS.

Knowledge based systems: A Knowledge based system (KBS) combines a computer's capacity for processing voluminous input with an expert's capacity for heuristics and reasoning with uncertainty (Barrett and Jones, 1989; Crowe and Mutch, 1990; Davis, 1986; and McKinion and Lemmon, 1985), thus increasing the user's capacity for informed decision-making. A KBS can process large and complex data requirements and can potentially manage growing volumes of information and several sources of data used in complex decisions (Lacroix et. al., 1994). Papazoglou and Zeleznikow (1991) assert that with the development of powerful computers there is now the opportunity for better tools to move from data based assistance to knowledge based assistance. Barrett and Jones (1989) suggest that an expert system or KBS is best employed when the problem is ill-defined, poorly structured and requires the help of a human expert. Comparing a KBS with a human expert, it is superior in modelling various types of relationships simultaneously thus more closely reflecting the biological complexity in dairy systems (Hogeveen et al., 1991, 1992). In KBS the knowledge is permanently stored and more easily transferrable making the decision support given more consistent and less subject to change. Its permanence can also be a major drawback since KBS cannot learn and change once created, unlike a human expert, making it less flexible.

There are several structures and functions common to KBSs. Extending from the database concept, the knowledge base is a repository of domain expertise – encoded symbolically into the software (Barrett and Jones, 1989; and Hogeveen et al., 1994). The user interface allows the person making the inquiry to interact with the computer (Davis, 1986). It interprets the human language representations and presents the required information to the KBS and vice-versa when given its response or requesting further input (Barrett and Jones, 1989). Simulation models that contain knowledge from research, can be modelled using conventional computer software such as spreadsheets (Hochman et al., 1995b; and Hogeveen et al., 1992). The inference engine of the KBS selects those parts of the KB relevant to a particular occurrence of the problem and makes conclusions about the occurrence useful to the user (Davis, 1986). According to Crowe and Mutch (1990) it "controls the execution of the expert system, links all of the operations and searches the data bases through a number of modules in response to input provided by the user." Knowledge is maintained in what is called the knowledge base. In the dairy industry, research has been conducted on the representation of knowledge in various forms including rule-based, fuzzy logic, neural network, Bayesian belief network, conditional causal model or simulation models (Hogeveen et al., 1991, 1992, 1994; Hochman et al., 1991, 1995a, 1995b; Lacroix et al., 1994; Pellerin et al., 1994; and Strasser et al., 1997). Different methods of knowledge representation have specific advantages for the type of data in the system, requirements for explanability to the end user, adaptability to problems being modelled and the expected role of the decision-support system.

Distributed decision support: Distributed decision support is the use of network computing to deliver and receive information important to decision making. There are many advantages to distribution including the ability to provide on-farm support from many organisations. However there are challenges in a distributed environment including heterogeneity, autonomy, and adaptation of legacy systems (Nicol, 1993; Bauer et al., 1994; and Attaluri et al., 1995).

Decision Support : Information Technology Support in Agriculture

Decision-support systems (DSS) are useful tools to increase the value of information imparted to the user through knowledge (Dhar and Stein, 1997). They can be used in advisory roles (Strasser et al., 1997), strategic planning (Lacroix et al., 1996; and Wade and Lacroix, 1997), and diagnostic processing functions (Hogeveen et al., 1991; Morimoto et al., 1995; and Spahr, 1988).

The use of DSS in agriculture, specifically involving knowledge based reasoning techniques, has been investigated since the mid 1980s (Barrett and Jones, 1989; Davis, 1986; and Greer et al., 1994). Researchers have developed prototypes in beef production and pasture management (Hochman et al., 1991, 1995a and 1995b), swine production (Backus et al., 1995), hydroponics and greenhouse control and management (Morimoto et al., 1995; and Kok and Lacroix, 1993), and dairy production (Allore and Jones, 1995; Allore et al., 1995; Dijkhuizen et al., 1996; Hogeveen et al., 1991, 1992, 1994; Pellerin et al.,1994; Pellerin et al., 1995; and Strasser, 1997). Gauthier and Kok (1988) described functions, technologies in cognition, data and knowledge management and information processing that would enable nonrestrictive data and knowledge bases to be developed for on-farm decision support. This is just a sample of DSS that can be developed. A potential problems end up being how to manage a plethora of decision-support systems.

The potential exists to provide knowledge and information dense advice on-farm for management decision support (Barrett and Jones, 1989). The advantage is that the average worker can be raised to the level of expert (McKinion and Lemmon, 1985). Less time would have to be spent on routine activities such as data sorting, analysis and interpretation allowing the producer more time for actual decision making and other activities. Where decisions on production are being made DSS can consider large volumes of data and support the decision making process by reducing the time necessary for analysis and remove ambiguity created by human bias or short comings. Automating decision making is also a valuable objective in some situations where the decisions that need to be taken are repetitive and can increase speed and efficiency if made by a KBS (Lacroix et al., 1994).

Multiple Knowledge Sources in Supporting Decision-Making

Knowledge based systems can support decision making in a narrow domain. To improve decision making, cooperation of several KBS in aiding agricultural problem solving accomplishes more than simply bringing all relevant sources of information together; it integrates information and knowledge across many knowledge areas and disciplines thus enhancing the value of that knowledge (Hochman et al., 1995b). In Hogeveen et al. (1994) various knowledge representation methods are described and the characteristics of the types of knowledge that are best represented by them; the most appropriate representation of each knowledge source can be included. Maximizing the information and knowledge available in the system when a decision is to be taken improves the potential for making the best decision with respect to the user by optimising the genetic resource. With independently acting knowledge sources any one or more of them may contribute to the problem as it evolves.

Hogeveen et al. (1991) considered the use of several domain experts in one KBS to be an advantage. It can give decision support from many points of view and get beyond the subjective preferences or biases of an individual expert. Integrated decision support for management on dairy farm allows the knowledge of experts from several domains to be applied to a single problem broader than any one domain, (Hogeveen et al., 1992). Hochman et al. (1995a) state that cooperating knowledge sources may enhance decision making: each domain may have weaknesses being distinct from one another but their strengths address different aspects of the problem. The accrued knowledge thus becomes more reflective of a true-to-life situation. No comment was made about how the design of the system might affect the success or failure of using multiple sources of knowledge.

Difficulties can arise in decision support offered when more than one expert is used in knowledge engineering process (Barrett and Jones, 1989). Conflicting advice may result (Malone and Crowston, 994) and the knowledge acquired may not be complimentary nor complete enough to mitigate equivocality. Schmoldt and Rauscher (1994) assert that the design of the knowledge system will to some extent contain the framework (cognitive organisation) within which the experts' own experiences were formulated suggesting that using more than one expert to develop a KBS may lead to incompatible and ambiguous results since their approaches may differ. Research into cognitive style and its effect on the design of DSS suggests that individual approaches to decision making exist and they influence the approaches to problem solving (Huber, 1983; Ramaprasad, 1987; and Robey, 1983).

Methods used for multi-knowledge sources in decision support: Two multiple knowledge source methods that support problem solving in different ways are the blackboard approach and the data driven approach. A blackboard model uses various problem solving approaches (specific to the knowledge sources) to solve complex problems (Barrett and Jones, 1989; Craig, 1995 and Englemore et al., 1988). Each knowledge source can be represented in different ways including, for example, spreadsheets, simulation models or geographic ISs (Erman et al., 1988; Hochman et al., 1995b; and Batchelor and McClendon, 1992). The value of independent knowledge sources is the incorporation of expertise specific and dedicated to a partial solution of a greater problem. Batchelor and McClendon (1992) in their use of a blackboard approach for resolving conflicting irrigation and insecticide scheduling recommendations, used a relational data base design, for the development of their blackboard, where information commonly used by knowledge sources to support the system are shared. The knowledge sources can affect only what has been placed on the blackboard.

An Executive Information System (EIS), or an on line analytical processing system, is a type of DSS designed to support managers needs for information (Young and Hugh, 1995). Thus they handle aggregations of information produced within a domain, which may originate from many sources and may represent different areas of expertise. Dhar and Stein (1997) consider data driven systems as a method of decision support that is not model based.

That is, they do not depend on a single method to implement the knowledge. The data required from the system determines what process knowledge is called thus knowledge from several sources can be employed. Supplying knowledge to the solution space, along with data, but not structured in one particular model of problem solving, driven by the inference engine. Usually EISs include a graphical user interface, integrated capabilities, *ad hoc* queries, "drill down" capabilities and access to a variety of external data sources (Nord and Nord, 1995).

In a variety of management climates EISs have been employed (Koh and Watson, 1998; Nord and Nord, 1995; and Young and Watson, 1995) and, while commercial and research EISs have been developed (Dutta et al., 1997; and Varhol, 1995), prototype applications for on-farm management have been described but have yet to be implemented (Ivanovic and Budimac 1999; and Lacroix et al., 1994). An EIS would be an appropriate model for developing on-farm decision support since it focusses on providing information to the management level without many prior assumptions about how the information should be selected. In Bauer et al. (1994) a peer-to-peer model of computing which would support a distributed independent concept of domain expertise focussed on a toolkit with sufficient breadth to allow users to assemble solutions. The overall architecture should be developed so that it is transparent enough to allow ordinary users in the environment to serve their own interests.

Networking and Distributed Systems in Agriculture

Nelson (1994) defines networking as 'a collaborative process of information exchange, around a central theme, carried out by actively interested parties'. Furthermore, he states that information exchange can be carried out in networks through audio-visual or written modes, in person, by post or electronic exchange. Electronic networks would create a greater advantage to the movement of information from the perspective of costs, time and the format of capturing and retrieving the data. Systems to support farm managers, with information from several organisational sources of information, would have to support an industry-wide network and multi-media formats. Networks currently sponsor information exchange in one-to-one (individual to individual) and one-to-many (individual to group) information transfers (Lacroix et al., 1998 and Lacroix et al., 1998b).

Electronic data interchange is a method that automatically exchanges information between computer systems through structured standardized messages (Lokhorst and Kroeze, 1996; Silver and Silver, 1994), and could be applied over the Internet. It is useful for ISs where data are being reused; speed of transactions is important; and a high volume of data is exchanged among many participants (Lokhorst and Kroeze, 1996; and Spahr, 1993). Additional benefits from EDI may be realised when combined with other emerging technologies such as remote sensing, electronic identification of animals and automatic recording of events (Kalter et al., 1992). Tomaszewski (1992 and 1993) and Jones and Spahr (1991) discuss the establishment of electronic modes of exchanging milk recording and other types of information among farms and DHI and other off-farm sources. Presently it is possible to download test-day results or reports for use in PC-based management software from a DHI to farms (Rowe, 1997; Spahr, 1993). The Dutch dairy industry has developed an EDI system (Agricultural Data Information Standard) for their dairy sector and expect it to be useful for various functions such as daily production monitoring and official herd reports (Dindorp et al., 1996; Lokhorst and Kroeze, 1996; and Koorn, 1996).

An Information Super-Highway as a network: While the term information super-highway (ISH) is dated, the concept is useful. The basis for the ISH is the production of powerful computer technology such as supercomputers, capable of handling billions of commands per second, processing and analysing complex applications, and the development of relatively inexpensive and reliable telecommunications technologies (Dertouzos, 1995). With wireless technologies, the potential for expansion of the ISH is immense (O'Malley, 1995). Telecommunications consist mainly of telephone networks but other network components may be involved such as cable or satellite communications (Olson, 1991; and Voss, 1994). Dertouzos (1995) described a true ISH as widely available, easy to use and the foundation for many useful activities. The WWW could be considered as approaching an ISH.

The Internet, particularly through the WWW, is a communication vehicle that has added a dimension to common communication. It is a relatively highspeed facilitator of two-way communication in text and visual format (Voss, 1994). Several services have been developed to express these forms of information exchange: e.g., Electronic mail (E-mail), file transfer protocol (FTP), TELNET, USENET groups and list-servers (Ba et al., 1997; Davis et al., 1994; Egeberg and Rice, 1994; Fournier et al., 1997; Fowler et al., 1994; Innes, 1994; Lacroix et al., 1998a; Lacroix et al., 1998b; Rimmington et al., 1994; Varner and Cady, 1993). Manguerra (1997) sees the Internet as a suitable medium for the development of DSS since both technologies rely on well-structured information-management techniques for their development.

The WWW is becoming a transparent means of mass communication as the medium supporting other information technologies and supported by a critical mass of communication technologies. Intranets and Extranets, within and across organisational networks respectively, have become important for supporting users in many industries (Ba et al., 1997; Barrett and Konsynski, 1982; Covill, 1998; Lloyd and Boyle, 1998). Technologies supported by the WWW include information transactions, commercial services, knowledge-based search and analysis, and data warehouses (Brick and Henry, 1999; Hackathorn, 1999; and Kimball, 1999). The improved efficiency of workflow, simplified and streamlined processing, along with a value-added focus and potential for real-time delivery, can create true e-commerce (Trustman and Meshako, 1999).

Several farm and agricultural organisations have also taken advantage of its communication capabilities by linking extension services and by providing direct on-farm support from experts and general information sources (Davis et al., 1994; Getz, 1994; Ridson, 1994; Smith, 1994; and Varner, 1993). The Internet has been employed to support extension offices by providing data and information services such as weather reporting and futures' prices (Fowler et al., 1996). List-servers support a range of dairy production and improvement topics (Fournier et al., 1997, Varner and Cady, 1993) and E-mail can provide technical support on a one-to-one basis over the Internet.

There may be some technical and sociological limitations to the use of the Internet in some agricultural systems. Strong et al. (1997) describe several issues that might impede access to high quality information through the Internet including inconsistent definitions and formats of information among systems; difficult timely access due to information volume; changing information relevancy as users' tasks and providers' environments change; and lack of automation across independent organisation sources, making integration of large volumes of specific data difficult. In agricultural systems the geographical distance and rural location makes the establishment of cost-effective networks more challenging than in other business environments. Also the independence of individual farmers may contribute to a proliferation of definitions and formats of the information stored and expected at the farm level. Employment of the Internet is also a function of the adoption of computer technology by producers.

The demographics of users in many agricultural industries suggests that they may be resistant to adopting information technologies. In the European Union 10% to 20% of producers have access to computers (Waksman and Harkin, 1996; and Koorn, 1996) which is similar to levels in the U.S. (15% to 25%, Dukas, personal communication 1997). In New Zealand 47% of dairy producers own computers with 28% of them being connected to the Internet (Rosskopf, 1999). Thompson-Jones (1999) found that in Canada 29% of rural and small town residents had one member of the household who had used computers at least once and only 3% reporting typical use for business purposes. Those findings conflict with a survey of Quebec dairy producers that showed more than 50% owned a PC less than 5 years old and used it for on farm work (Lacroix et al., 1996b). Of this group many may not own PCs appropriate for Internet connectivity (less than 10% were using the Internet for on-farm work) or may not live in an area where inexpensive Internet access is available. Connectivity limitations will be resolved as technologies develop to provide intuitive means of connecting to the Internet, such as Web TV, Internet cellular phones (Lacroix and Wade, 1996a), and are developed to be suitable to the mobile, outside environments (Rosskopf, 1999).

Distributed systems: The primary characteristic in a distributed system is that the parts of the system are not located at the same physical location but are controlled by the participants who generate or collect the information products (Papazoglou and Zeleznikow, 1991).

Realistically this will mean differences in data formats, storage methods, database systems and applications used, operating systems, hardware and network configurations among information providers. Distributed information technology integrates geographically distant information sources into what appears to be a single logical system of data/knowledge and underlying facilities, (Papazoglou and Zeleznikow, 1991) allowing each participant to apply their expertise in the maintenance and development of their contribution (Millikin, 1994; Nadeau, 1994, Nicol et al., 1993) and at the same time access other resources the system provides. Attempts to collectively manage large data sets and information resources, useful to many users in the system, must be able to incorporate legacy information (Papazoglou and Zeleznikow, 1991) (and their applications to some extent) and resources.

Conclusions

Development of a framework to support on-farm breeding decisions, based on information and knowledge available from the dairy industry, requires synthesis among a number of areas of research. Investigating the support of dairy cattle breeding decisions involves knowledge of dairy breeding, decision-support systems and information systems domains. Information and knowledge can deal with uncertainty and equivocality, respectively, by clarifying the problem addressed, while a proper understanding of an information system should elucidate the use of both of these information products. Sources and uses of information and knowledge in the dairy industry are known but they have not been described as an information system. Apart from some specific applications, general information systems are not exploited in agriculture.

Information technologies, used to support information systems in the dairy industry, include management information systems, knowledge based systems and network technologies. Management information systems, particularly database management systems, have been developed to support the use of information on farms. With few exceptions the description of models for data systems in the industry is non-existent. Although recent research in agricultural systems shows interest in the use of knowledge-based techniques to support decision making, work should be carried out to find methods that build on prior research efforts into data systems. It seems clear that the use of decision-support systems will grow and those based on distributed information systems, such as the Internet, can make information delivery and decision support seem continuous.

Chapter 3. An Information Systems Approach to the Dairy Breeding Industry

A A Archer, R Lacroix, R I Cue and K M Wade.

A framework for integrating and analysing information requires a proper understanding of the information system within which it operates. This chapter introduces a method useful for systems comprehension and description. The approach was adapted to be useful for describing multi-organisational systems in general and was specifically applied to the dairy breeding industry. It facilitates the analysis of various dairy breeding activities.

Abstract

This paper describes a systems approach for analysing information systems which is then used to look at a DBIS. It discusses its performance from a specific perspective, and suggests how the system might be improved. The general systems approach describes system characteristics in terms of the goals, boundary, functions, structures, and performance. The performance can be assessed with respect to the goal described using attributes of the system expressed at an appropriate level of detail – examples of attributes are given. Using this approach, the goal of the DBIS was to support decision making on dairy farms for improving traits of commercial interest based on information and knowledge available from many sources in the industry. Several functions that create information were described and were categorised as on-farm breeding activities, herd output and support information. The components of the structure of the Canadian DBIS included farm, dairy herd improvement association, genetic evaluation, breed association, artificial insemination, research and government sectors. An analysis of DBIS performance from an on-farm decision maker's perspective showed that human-agents performed well as decision makers because of abilities such as heuristic and abstract reasoning, however they are limited in their data processing abilities. Recommendations for improving the DBIS's effectiveness ranged from automating and integrating the information among components before it flows to the farm, to unifying expected format, medium and time of information, as well as introducing knowledge-based techniques to support the human agent at the end of the information flow. The framework could be used to develop reference models to further study information systems and provide a means of improved decision making.

Introduction

Information comprises many types of aggregations of observed facts – ranging from signals to data to knowledge, that create understanding about a domain of interest. An information system is concerned with both the flow and the value of such products and can be defined as a complex set of components that is concerned with the maintenance, exchange and production of information with a specific end in mind for the overall system as well as the individual members. It goes, therefore, without saying that in order to use information technologies optimally, the whole system and its parts must be well understood. Systems theory is not new in its use as a model for real-world systems (Bunge, 1979) and has been applied to systems for organisations (within and across, virtually and physically) as well as to complete industries. A systematic methodology would assist the understanding of what the key parts are, their role(s) and how well such roles are carried out.

Churchman (1968) discusses a system's approach in terms of such concepts as a goal, a boundary, functions, components and performance measures while others have also included the notion of structure in the analysis (Bunge, 1979; and Ahituv et al., 1994). An industry-wide information system must include the flow of inputs/outputs from a performance viewpoint and many approaches can be used to evaluate system performance (Ballou et al., 1998; Burch et al., 1983; Churchman, 1968; Wybo, 1993; Todd and Benbasat, 1992; Goodhue et al., 1992; Ahituv, 1980; Ahituv, 1987; Ahituv et al., 1994; and Strong et al., 1997). However, the most suitable means of evaluating system performance will depend on the perceived goals and the perspective being considered (Malone and Crowston, 1994).

While information systems have been described in livestock production for the purposes of on-farm decision making (Tomaszewski, 1992; Tomaszewski, 1993; Dijkhuizen et al., 1996, Groeneveld and Gutzmann, 1990; Goldsmith et al., 1996; Groeneveld and Lacher, 1992; Jofre-Giraudo et al., 1990), a true systems approach was not applied, despite the fact that some of the specific requirements were obviously present. The primary objective of these previous cases was aimed at software development for the recording and/or analysis of data produced on-farm: they also focussed mainly on within-organisation software systems. The broader context of how information flows and is applied off-farm has not been seriously addressed. Although Groeneveld and Lacher (1992) do recognise a multi-organisational information system among swine producers and off-farm organisations that describes various information tasks and exchanges performed by different participants, their paper focusses mainly on software development and technical infrastructure needs while system description was clearly secondary. Tomaszewski (1993) and Misztal and Lawlor (1999) are similar in this regard, describing potential uses of the technology to improve on-farm dairy decisions but assuming that the information system is adequately understood. Applying a systems' approach first (to comprehend the underlying information system conceptually) and then subsequently developing technological support, based on identified needs, might improve the utility and acceptance of the technology in question.

A systems approach can be applied to the analysis of an industry-wide information system. Its use may prove beneficial for a more complete understanding and evaluation of such systems. The dairy breeding information system is an example that involves several participants on and off farm that exchange, store and process information.

The objectives of this paper are i) to describe an industry-wide information systems' approach – specifically to describe its goal, boundaries, functions and components; ii) to use this approach to analyse a specific dairy cattle breeding industry; iii) to discuss its performance from a given point of view; and iv) to postulate on possible improvements.

The General Characteristics of a systems approach

The essential nature of an information system is that it handles information that flows into as well as originates from the system and various operations within the system to create information products (Pietersma et al., 1998). Information products are created by adding value to information that flows through the system. Using a systems approach creates knowledge about how the system handles information by applying it to model the system.

In describing the general characteristics of an industry-wide information system (goals, boundary, functions and components) the approach can be either functional or structural – i.e., emphasizing the actions or the participants. Figure 3.1 is a simple representation of a multi-organisation information system with both functional and structural

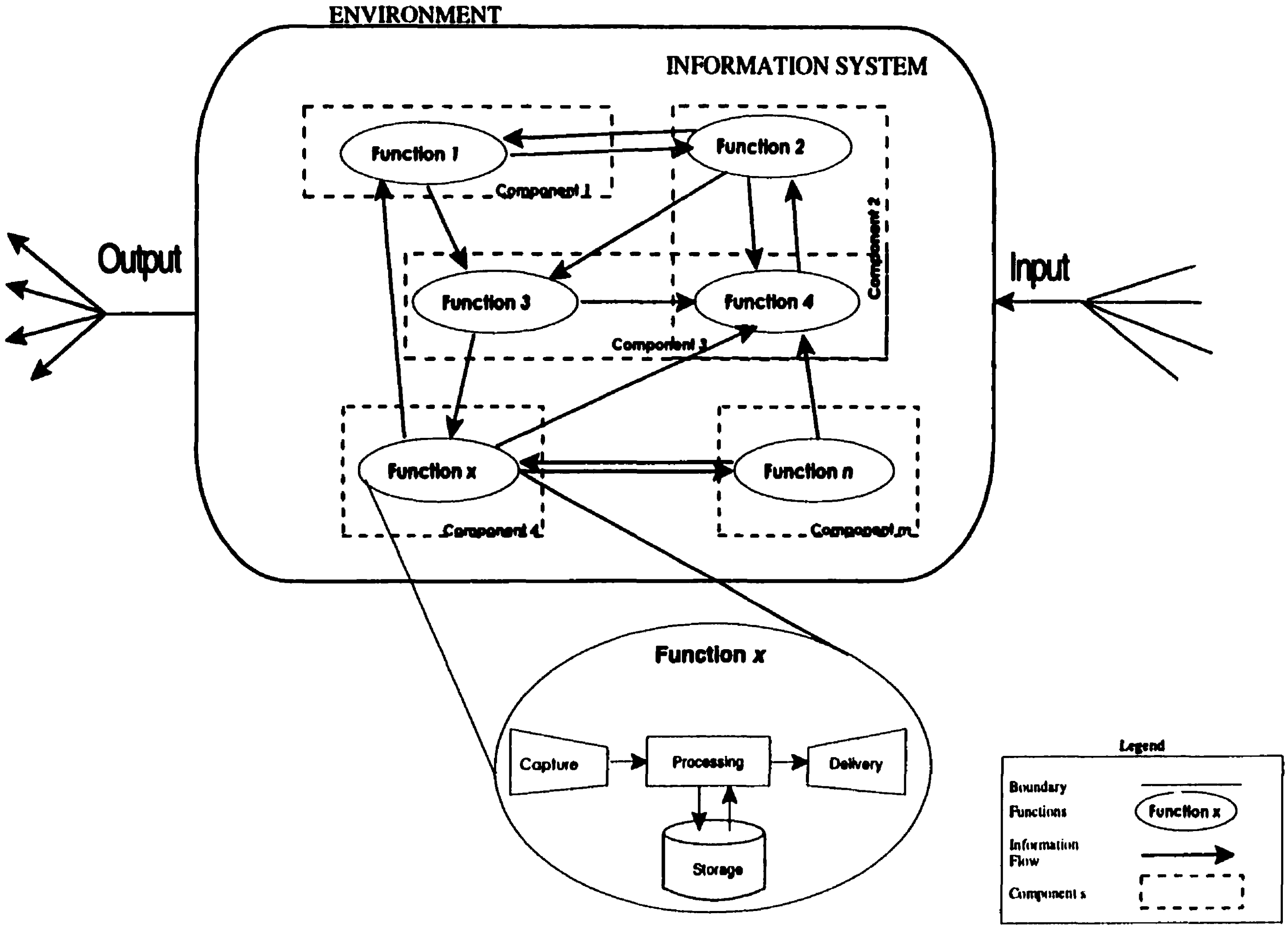


Figure 3.1 An illustration of the general characteristics of an information system.

characteristics. It will be used to illustrate most of the concepts in a systems approach.

Goal: The goal of the system is the reason the system exists (as described by the examiner with a specific vision). Malone and Crowston (1994) point out that in situations where resources are shared in a coordinating effort, they must be managed to best achieve the over all objectives of the system. The goal explains why particular operations are performed together in the first place; a delineation of the overall expectation of the participants. An implicit or explicit goal will exist in all systems (Ahituv et al., 1994) and identifying a goal is the first step to describe a system (i.e., which operations belong to the system and which belong to the environment around it). In Figure 3.1 the rectangular shape represents the boundary; the set of functions and components within it belong to the system.

Boundary: The boundary sets the scope of a system's operation. Within the boundary the functions' descriptions are directly derived from the goal of the system: that is, improving system performance will directly affect the performance of these functions. On the other hand, little impact on the characteristics or behaviour of external entities will occur (Churchman, 1968). Of course reciprocal influence does occur between the system and the environment since there are many information and non-information systems and entities in an industry and the world at large with which a system can interact. For example a system may be a subsystem within a greater system or some components within it may be components of other systems. While the boundary should be clear it is not fixed: it depends on the goal of the investigation (Ahituv et al., 1994), which if changed, could also change the boundary. Also, improving the performance of the system could concurrently change the functions or the components.

System Functions: System functions are important, specific operations that relate directly or indirectly to the described system and can be determined by looking at the purpose of individual functions and relating them to the overall goal of the system (Churchman, 1968). Conversely, one can determine how much effect the goal of the system has on the function(s) in question. The functions in an information system (e.g., 1- *n* in Figure 3.1) are based on 4 types of activities: the 1) capture, 2) processing, 3) storage and 4) delivery of information products. This is a greater level of detail (aggregation) than the current level being used to describe the information system. Function *x* (see Figure 3.1) is enlarged to show the relationship among these activities within functions - a level of detail that is useful for describing system performance.

Processing is the logical manipulation of collected information to satisfy participants' needs. Manipulation can take many forms including receipt, cleaning, integration, aggregation, transformation and decision-making. Processing adds value to information from input data by using human knowledge to create information or knowledge-based software resources which are beneficial to users of the system (Ahituv, 1987; Ahituv et al., 1994; McFadden and Hoffer, 1994; and Schmoldt and Rauscher, 1994). The processing can be human or computer based or both (Ahituv, 1987; Dhar and Stein, 1997) and occur at many different levels of planning (Pietersma et al., 1998). Similar information can exist in multiple media and formats – their creation themselves being processes – to facilitate the needs of different processing points or processors.

System Structure: The *structure* of the information system refers to the inter-relationships of the components in the overall system (Ahituv et al., 1994; Burch et al., 1983). At the activities level of detail, components contain structures that perform the actions, already described in Figure 3.1 (i.e., reception, processing, storage and distribution). The components (e.g., 1, 2 ... *m*) in Figure 3.1 are the different entities that perform the required functions. A single component of the system may be composed of one or several participants. Components *1* and *2* respectively, show that a component may be responsible for one or more functions. The functions carried out by a component require inputs, which may either be outputs produced by other functions or originate from other components within the system.

In a functional analysis the functions of the system are the focus of system description and may be considered fixed and determined by the goals of the system under analysis, however the components may vary from one instance of the information system to another. The information system in Figure 3.1 has x functions among n components. Another instance of the information system in time or place would still have x functions, but the number of components could change. Structural differences emerge when many instances of the information system are analysed. If by the goal a structural analysis is emphasised, the components may be fixed and the functions within the components may vary.

Performance Analysis: The performance of the information system under examination indicates how well the system accomplishes the goal described by the examiner. Performance can be assessed using appropriate measures of significant attributes of the information system chosen with respect to the goal described and a level of detail corresponding to that goal (Churchman, 1968; Ahituv et al., 1994; Burch et al., 1983; Ahituv, 1980). As well as considering how to assess the system one must also consider its importance in achieving the goal of the analysis. For example if the goal of the system analysis is to describe the information system, system performance is not at all important. However, if the goal of the analysis is to evaluate the information system efficiency, the performance is critical to the analysis.

Attributes of the information system can be considered as *aspects*: they apply to the characteristics of the system described including functions, structure and information. The appropriate level of detail to investigate the attributes will likely be consistent with the level used to describe system characteristics in the analysis. There are four classes of attributes that are commonly cited in the literature: timeliness, format, cost and value (Ahituv et al., 1994; Ahituv, 1980; Burch et al., 1983; and Ballou et al. 1998). Table 3.1 gives an example of how these classes of attributes could be represented at the activities level of detail for function, structure and information. This is not an exhaustive list – other attributes may be described – however it shows a number of attributes of the characteristics of the system that may be investigated. Again, the attributes used in evaluating the system will depend upon what in the system is being examined according to the goal of the analysis. These attributes could be directly used or combined, either quantitatively or qualitatively to describe how well the system is functioning.

Determining how to measure the attributes of the system used as performance indicators is possibly the most difficult part of performance analysis. The performance might be characterised in terms of efficiency, quality, effectiveness, utility, optimality, etc.

Table 3.1 Examples of the attributes of functions, structure and information at the activities level in an information system.

		Timeliness	Content	Format	Cost
Capture/Reception	Functions	input time	consistency, reliability	input suitability, flexibility	time, accuracy, reliability
	Structure	delivery time	consistency, reliability	compatibility, integratability	time, expense, integrity
	Information	age, volatility	integrity, accuracy, reliability, aggregation, exhaustiveness, relevance, redundancy	unity, transferability, portability	expense
Process	Functions	time	consistency, reliability	consistency, integrity	time, value
	Structure	time	consistency, reliability	suitability	time, expense
	Information	quality	aggregation, accuracy, reliability, integrity, completeness	aggregation, clarity	expense
Storage	Functions	update/access time	reliability	optimality, exhaustiveness	availability, usability
	Structure	update/access time	reliability, effectiveness	accessibility, security, media, design	expense, time
	Information	accessibility, availability	redundancy, relevance, exhaustiveness, aggregation, reliability	aggregation, optimality, efficiency, completeness	expense, opportunity
Distribution	Functions	output time	consistency, reliability	flexibility, output suitability	time, expense
	Structure	delivery time	consistency, reliability	clarity, design, media	expense, integrity, time
	Information	age, volatility	integrity, completeness, relevance, redundancy	unity, transferability, portability	expense

It could be qualitative, quantitative or a combination of both. The measure of performance is necessarily subjective in some respects since its basis is from a given perspective and a limited description of what is valuable (Ahituv, 1980).

Dairy Breeding Information System Characteristics

The information that serves dairy cattle breeding activities is multi-organisational and industry-wide. In this section an example is given to illustrate how the systems approach could be used to analyse an information system for dairy breeding information. This system shall, henceforth, be known as the dairy breeding information system (DBIS).

Goal and Boundary: The goal of this analysis will be to describe the information system that supports dairy breeding decision making on dairy farms for traits of commercial interest based on information and knowledge available from many sources in the industry. Figure 3.2 is an example of the functions and components included within the boundary of such a system. On-farm, the functions included are cow culling, sire selection and mating. Off-farm the functions include sire analysis; genetic evaluations; milk recording; extension and instruction in dairy cattle breeding, selection, culling; research in important traits, effective programs, genetic progress, statistical methods, conformation analysis; registration and herd-book keeping, etc. Although arguably important as well, this boundary ***excludes*** other functions that do not contribute directly to this goal such as nutrition and feeding , health management, and other agricultural systems.

DBIS Functions: The relationships among the functions in the system can be clarified by considering them at a "macro"-level, where types of functions are grouped together, as well as by each function individually. Table 3.2 shows the functions of the DBIS, described in Figure 3.2, grouped according to classes, and based on the types of information they produce. Breeding functions pertain to decisions made on-farm about the selection and breeding of individual or groups of animals - primarily human based at this time. Herd output functions produce information pertaining to the value of animals for their own present or future production, or that of their progeny. With herd output information the link to the source of the information is maintained while support information is generalised and the link is not sustained. Support information functions create knowledge and information that is generally applicable to individual or groups of animals.

Table 3.2 "Macro"-Level of Detail for DBIS Functions Classification

Function Type	Individual Functions
Breeding Functions	sire selection, cow culling, cow replacement, sire/cow mating
Herd Output Functions	milk recording, milk analysis, classification, milk production genetic evaluation, type trait genetic evaluation, animal identification, pedigree analysis, sire analysis,
Support Functions	Education, training, extension, instruction, regulation, conceptualisation, modelling, legislation, policy making

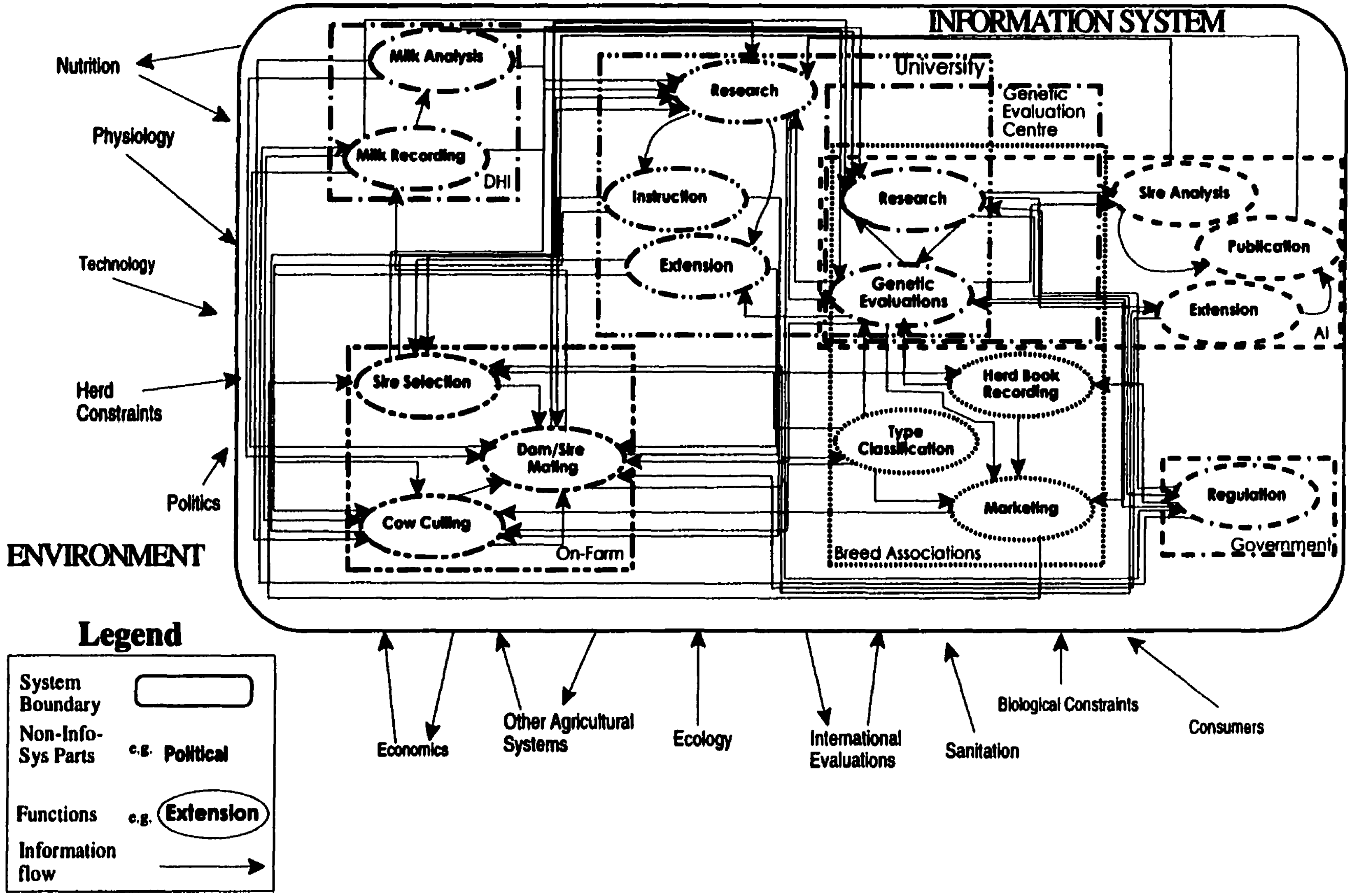


Figure 3.2 An example of the dairy breeding information system (DBIS)'s boundary, functions and influences.

Functions and Flow of Information in a DBIS: In the DBIS, information flows through components to support the on-farm decision-making. Figure 3.3 is a data flow diagram using cow culling as an example of the inter-dependence of decision making on the information moving through the DBIS from an on-farm point of view. It also shows the inter-relationship among functions at the macro-level of detail. Starting with production data on the farm – the basis for any animal agriculture information system (Groeneveld and Lacher, 1992), milk records are processed into milk performance records (Etgen, et al., 1987; Fiez, 1998). These records can be transformed into monthly, yield-to-date and 305-d lactation records (Christensen and Fehr, 1993) which indicate the herd and individual performance. Laboratory milk sample analyses for milk fat, protein and SCC are useful performance measures of milk quality. Herd performance reports on individuals or groups of animals can be used on farm for herd improvement decisions (Etgen et al., 1987).

Records on the physical characteristics of animals on farm, called type or conformation traits, are collected and used to maintain breed standards. Type traits may also be used for long term production and genetic purposes to make cow culling decisions (Christensen and Fehr, 1993). Different breeds will vary, depending on the characteristics and number of traits being considered.

Cow culling (as well as sire selection, and mating - see Figure 3.3) rely on the information above and, particularly, on genetic evaluations of animals. Standardized milk records of the productive ability of a cow are used to assess the potential genetic value (genetic evaluation) of economically important production traits by statistical transformation along with pedigree information (genetic evaluations are also computed for type traits). The measure of genetic potential can be expressed in different ways which can vary depending on the type of breeding decision being taken. For example, the genetic potential can be given as an expression of the dam's contribution to its offspring if one wants to investigate the pedigree or it could be transformed into an index (transformations of the computed genetic evaluations), combining several important economic traits together, and used directly as a culling tool. Differences in how genetic evaluations are expressed vary among countries for many reasons.

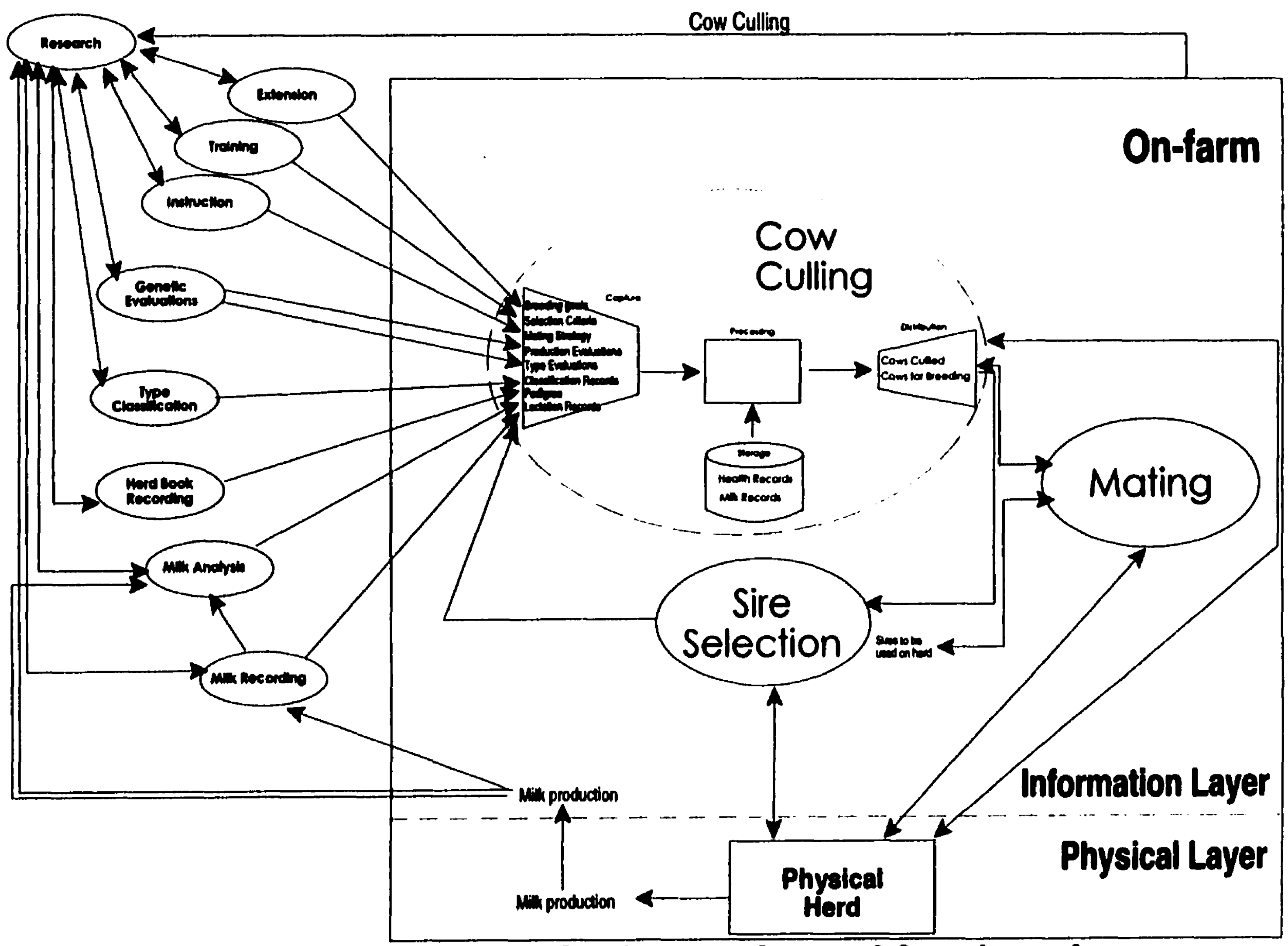


Figure 3.3 A data flow diagram of herd output and support information products exchange in the DBIS to support on-farm culling decision-making.

Genetic evaluations of individual cows are statistically combined to produce estimates of genetic merit of their male relations for these same traits; known as sire proofs. Maintaining the herd books is needed for accurate genetic evaluations of animals. These functions include the identification of animals, animal searches, and validation of pedigrees. Young bulls are offered for service based on projection of their pedigrees. Information for ranking bulls is transformed into several index formats which the decision-maker uses as a technology for more efficient selection. Cows culled and bulls selected may have an affect on each other and on mating as well.

Support information is also necessary for decision making. Research, regulation and instruction functions are all required to support the needs of on-farm dairy breeding decision-making. Support information communicates how knowledge produced by the other functions should be employed; they give a context for the other functions. For example in Figure 3.3 the arrows from the research function to the other functions indicate that they contribute to all functions to some extent. These functions then contribute to the decisions made on-farm. Likewise information is captured by research functions from all other functions. Research activities about various aspects of improving and utilising the genetic base of dairy production – communicating chiefly through the publication of resulting knowledge. Research functions capture data and information produced in the system, analyse it and produce concepts and models. The information produced may be in the form of generally applied information (e.g., statistics), knowledge on how to apply information (e.g., limitations of models or characterisation of new traits) or knowledge for improving the processing of functions (technology).

Research and instruction are dependent functions since the products of research are often communicated through instruction and researchers are "created" through instruction. For example, research about a specific trait or technique of estimation of that trait is also communicated through teaching and training of participants in the DBIS. By extension, training and institutional instruction knowledge about culling can be discovered and correctly applied; what was learned is disseminated.

DBIS Structure: A DBIS is required to perform the functions described above. Referring to Figure 3.2, the functions are performed by 5 different components. Many of the functions are performed by more than one component. A clearer view of the structure of the DBIS is shown in Table 3.3 of the decision-making, herd output and support functions. An indicator was included horizontally beside each sector on the Table for each function in which they participated. Where a single indicator appears in any vertical section one sector is solely responsible for a given function. Where there are many indicators appearing vertically there is shared responsibility for the function. The structures of the US, Dutch, New Zealand and Danish dairy breeding information systems are shown demonstrating differences in the ways that a DBIS may be organised in different countries.

The components' names follow organisational sector names in the industry: on-farm, dairy herd improvement associations (**DHI**), genetic evaluation centres (**GE**), breed associations (**BA**), artificial insemination (**AI**), research institutions, and government components (Table 3.3). The DHI support milking performance functions (i.e., milk recording and component analysis). In the Table these functions are supported by a single sector. However, a group of participants representing different sectors support the evaluation functions (including BA, GE, AI, and research institutions). This is indicated from Table 3.3 where the indicators lined up vertically show that more than one component structure is involved with one activity. Each sector exercises control over the performance of evaluation functions in different dimensions (e.g., GE at the national level for multiple breeds; BA at the national level for a single breed; and AI at the regional level for all

Table 3.3 A selected list of national DBIS structures.

Country	Functions	Culling	Sire Selection	Pairing	Recording	Component Analysis	Classification	Production EBVs	Conformation EBVs	Herd Book	Sire Analysis	Extension	Research	Instruction	Training	Regulation
Canadian DBIS	On-Farm	●	●	●												
	DHI				●	●							●			
	GE							●	●		●					
	BA						●		●	●			●			
	AI								●		●	●	●		●	
	Research Institutions											●	●	●	●	
	Government											●				●
USA	On-Farm	●	●	●												
	DHI				●	●		●				●				
	BA						●		●	●						
	AI										●				●	
	Research Institutions											●	●	●	●	
	Government							●	●		●	●	●		●	
the Netherlands	On-Farm	●	●	●												
	DHI/AI/GE				●	●	●	●	●	●	●	●	●		●	
	Research Institutions											●	●	●		
	Government												●			●
New Zealand	On-Farm	●	●	●												
	DHI/AI/GE				●	●		●	●	●	●	●	●		●	
	BA						●			●	●					
	Research Institutions											●	●	●		
	Government												●			●
Denmark	On-Farm	●	●	●												
	DHI/AI/BA/GE				●	●	●	●	●	●	●	●	●		●	
	Research Institutions											●	●	●		
	Danish Cattle Husbandry															●
	Government												●			●

breeds). The functions support varying informational needs for the aspects of evaluation functions that suit their clientele. Beyond the required interdependence, created by data input/output dependencies (see Figure 3.3), another layer of interdependencies among the participants exists since they must share responsibilities to perform DBIS evaluation functions.

The herd-book functions are supported by BA which is also responsible for performing the classification functions in the DBIS. These functions are the exclusive responsibility of these participants, however there are several participants in this sector since there are many dairy breeds in the industry. The AI are not exclusive to breeds but represent information on many sires regionally. There is a single GE which is responsible for all production evaluations and performs conformation evaluations in conjunction with BA, AI and DHI.

The support functions are not partitioned among sectors in the DBIS with the exception of regulation enacted by government bodies. Regulatory functions between national and regional governments have jurisdictional differences that affect individual sectors differently depending on their mandate and role. Support functions are performed by many individual sectors. Extension and training functions allow sectors to promote the utility of the information services they provide. Research is conducted by participants in each sector to provide information, knowledge and technology that improve other functions for which they are responsible occurring within a sector or among sectors. Research may be a shared responsibility (similarly to type evaluations), i.e., a number of sectors sharing resources to engage in mutually beneficial research activities.

The functions of the DBIS in other places may have a different structure (see Figure 3.3). In the US the number of sectors and their roles are similar to the Canadian DBIS described. Two notable differences are that the federal government is more involved in producing the genetic evaluations, and producing and disseminating sire information. The European and New Zealand DBIS have fewer sectors showing a greater integration of the functions (Holmes and Wilson, 1984; Pedersen et al., 1996; Borchersen, 1996). In Denmark and New Zealand this greater integration has allowed these small countries to make substantial genetic gain for traits in their particular markets. The integration of milk recording and genetic evaluation information functions has allowed The Netherlands to pursue a high level of genetic improvement, producing among the best bulls for protein in the world. The US strategy is also highly successful in producing a high degree of genetic gain and serving the demands of US dairy producers. Different DBIS structures have evolved according to different needs, markets and social structures.

Performance Analysis: There are many ways of improving on-farm decision making. The following section, based on the functional analysis used, will discuss effectiveness in practical, qualitative terms as one might currently observe them in a DBIS. Attributes pertaining to on-farm decision-making effectiveness are considered mainly at the activities level of detail. The use of some information technologies is proposed as means of improving effectiveness.

The current cow-culling, sire selection and mating decision-making practices will be considered the standard of effectiveness in this system. It is assumed that the practices of these functions are sustainable and afford the decision-maker at least the minimum satisfaction with their results. A description of current practices follows. Information products are currently exchanged (captured and delivered) through mail, courier, FAX, telephone, electronic networks and by human agents (Newman, 1999). They are stored in physical media such as paper filing and electronically in computer floppy and hard disks or in human memory. Given that the amount of computer usage on-farm ranges from 10- 25% (Rowe, 1997; Dukas, personal communication 1997; Thompson-James, 1999) the standard means of on-farm processing is by human agents.

The following describes how attributes useful for evaluating effectiveness (see Table 3.4) may be used at the level of activities considering the strengths and weaknesses of the human agent involved in the processing activities of decision functions. These decisions are reasonably complex, potentially involving information inputs and knowledge from multiple sources, while employing individualistic strategies that may not be well described. Human-agents as decision makers can be adept at heuristic reasoning and abstract processing, useful for these types of decisions, reliable at integrating the captured and stored information to make reasonably reliable decisions. They are also flexible in accepting a range of suitable formats for knowledge (i.e., coming from a combination of instruction, extension, expert advice or 'know-how') and information (e.g., recognising relationships among sources for records for the same animals). They do have limitations which can affect the timeliness and

Table 3.4 Attributes used to assess the effectiveness of on-farm decision making functions at the activity level of detail.

	Cow Culling; Sire Selection; Mating			
	Timeliness	Content	Format	Cost
Capture	input time	consistency, reliability	input suitability, flexibility	time, accuracy, reliability
Process	processing time	consistency, reliability	consistency, integrity	time, value
Store	update/access time	reliability	optimality, exhaustiveness	availability, usability
Distribute	output time	consistency, reliability	flexibility, output suitability	time, expense

content of decision functions. Their limited capacity to consider a large number of multiple parameters simultaneously and perform rapid data processing may increase the time required to capture incoming data, retrieve stored information and actually process culling, sire selection or mating decisions. Therefore, consistency among individuals or with one decision-maker over time may not be optimum. In addition, human agents are not always consistent when capturing, recording, retrieving information or recalling events (i.e., from memory, multiple sources). Multiple decision makers may compound content inconsistency cause reliability problems and create inconsistencies in formats used for information representation. They may also have varied expectations of the amount of information content required for reliable decision making. These factors conspire to compromise the reliability of capturing, storing, processing and delivering activities.

Human agents can also introduce errors into the available information or simply mis-use it. Simple mistakes such as incorrect entry when capturing, delivering or storing activities may drastically reduce the effectiveness of the decision-making process. For example, a simple mis-typing of the registration numbers for a specific herd may render the information useless or increase the time required to process and access the correct information even if farm-management software is used to assist decision-making. Another type of error consists of undervaluing inputs or not applying them correctly in the decision process due to a limited ability of the human agent to consider multiple factors simultaneously, which can lead to inconsistency in reliable decisions making. The stated ranking importance placed on information may differ from the actual ranking used at the time of decision. Research into judgement (Ramaprasad, 1987) has shown that presentation affects one's ability to perceive or remember objects, thus affecting their utility. Thus, if the content of the information in the DBIS is reliable and consistent but the format is not readily accessible or the decision maker questions its integrity, the perceived value of the information is low and may result in under-utilization or misinterpretation (Burch et al., 1983).

The effectiveness of decision-making can be affected by retrieving and storing information products which is dependent on the consistency and reliability of the content stored, the format's optimality and exhaustiveness. The update and access time are attributes of storage methods. Physical filing systems are ineffective methods of storage for these decision-making functions because of the volume of the information and inadequacy of update and access time. Electronic storage media (e.g., floppy disk, hard disk, and magnetic tape) are much more efficient. However, if the producer must update the information manually, because the information arrives on-farm through non-electronic means (not available or client not prepared to receive it that way), little is gained. Integration of the formats of information originating from independent sources will improve its optimality for decision making.

Recommendations for system improvement

The DBIS, as described, could be improved with respect to the effectiveness of on-farm decision making by minimising the potential for mis-handling through the integration of the flow of information onto the farm and among decision-making functions. Creating greater unity within the format of information among functions contributing to on-farm decision-making functions may increase on-farm use of information due to greater clarity of relationships across individual components. Unifying the expected format, the medium for its delivery and time of information presentation could increase effective capture and storage of information, while minimising the time and handling needed for these activities on-farm due to improved consistency and reliability (Koorn, 1996; Kroeze et al., 1996; and Tomaszewski, 1993). The weakness of human agents in some elements of processing activities and on-farm flow of information products among activities could be addressed by creating tools that automate these activities, allowing farm managers and those who assist them to concentrate more of their time on activities that suit their abilities (i.e., decision making). The integration of information products flowing onto the farm would improve the transparency among sources. This will minimise the confusion caused to the human agent by masses of information and a lack of human computational power allowing them to focus on complex and abstract reasoning which are their strengths.

One might expect high integration of information product format to lead to the integration of the components off-farm, similar perhaps to a European or New Zealand DBIS structure. Some examples of the potential benefits of integration include better information handling throughout the whole system, greater clientele for information produced, greater efficiency among functions performed off-farm, and the emergence of new applications of information, now possible through combined information and expertise. There is currently a high degree of cooperation since many of the off-farm functions themselves have been observed to be interdependent. On-farm, there would be fewer sources to gather information from and perhaps more support for electronic network solutions such as has occurred in The Netherlands (Koorn, 1996; and Kroeze et al., 1996). While this is possible, trends towards corporate independence of sector participants in Canada and the US, as well as long-time legacy formats and content, present serious challenges towards such a direction. The on-farm decision maker may also be resistant to changing traditional practices because of accustomed practices and pre-developed strategies. Hall (1991) suggests that intermediary positions exist between full integration and independence. The introduction of an electronic network, may allow the DBIS to occupy such a middle ground with a common means for the movement of information products among them while the participants maintain their independent control of functions.

The effectiveness of on-farm decision-making can be improved at the activities level of detail through the use of knowledge-based information technologies. Knowledge can be procedural and declarative (Schmoldt and Rauscher, 1994; Hogeveen et al., 1994), i.e., it can tell you "how to do something" or "what something is". Decision support systems (DSS) can use knowledge procedurally to structure and organise knowledge and information increasing its utility in a decision space. Computer-based processes perform transactions that are routine, simple and well modelled (Dhar and Stein, 1997). They enhance the abilities of decision-makers by compensating for their weaknesses in processing: handling multiple information needs in the decision space; increasing processing power and speed for rote processing (e.g., statistical calculations); and applying knowledge at an expert level to the problem. Declarative knowledge can be applied in a DBIS to determine the type end-user, data needs, and how best to present the information. These types of support will facilitate the decision-making process by reducing the time required to integrate the information exchanged and stored and improve the perception of information.

Conclusions

A systems approach was developed to describe industry-wide information systems according to its goal boundary, structure, functions and attributes. It can describe functions belonging to the information systems, show structural differences that exist among them over time and regions, and help with its performance analysis at varying levels. It expressly analyses DBIS but applies to industry-wide systems generally. The Canadian DBIS delivers genetic evaluation functions through a greater number of structural components compared to many other national DBISs. The analysis of the system shows that current flow of information through networks is adequate but could be improved to better serve on-farm decision-makers.

The framework could be used to develop reference models to further study information systems and provide means of improving decision making, e.g., anticipating the effects of technological changes to the system. The models should map how sources, sinks and pathways of information flow evolve through technological changes, and how these changes may affect the functions performed, structural changes, or the value of information products. Technologies such as DSS could be considered, while an evaluation of the benefits and limitations of DBIS deployment over the Internet represents the next logical area of research (although the development of quantitative performance measures as well as a greater level of detail in system components may be required for reference models).

Acknowledgements

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, les Fonds de l'Autoroute de l'Information du Gouvernement du Québec, and the Cattle Breeding Research Council of Canada.

References

Ahituv, N. 1980. A Systematic Approach Toward Assessing the Value of an Information System. Management Information Science Quarterly, Vol 4, No 4 (December): 61-75.

Ahituv, N. 1987. A metamodel of information flow: a tool to support information systems theory. Comm. of the ACM v 30, Sept:781.

Ahituv, N, S Neumann and H Norton Riley, eds. 1994. Principles of Information Systems for Management 4th Ed. Business and Educational Technologies: 22-36, 77-121.

Ballou, D, R Wang, H Pazer, and G Kumar-Tayi. 1998. Modeling Information manufacturing Systems to Determine Information Product Quality. Management Science 44(4): 462-484.

Borchersen, S. 1996. The Federation of Danish A.I. Societies. Pages *in* 34-37 Annual Report, National Committee on Danish Cattle Husbandry.

Bunge, M A. 1979. Ontology: A world of Systems (vol. 2). Dordrecht. D. Reidel.

Burch, J G, F R Strater, and G Grudnitski. 1983. Information Systems: Theory and Practice, 3rd Ed. John Wiley & Sons, New York, NY.

Churchman, C W. 1968. The Systems Approach. Delta Books, New York, NY.

Christensen, D A and M I Fehr. 1993. The Dairy Industry. Pages 75-92 in *Animal Production in Canada*, J Martin, R J Hudson and B A Young, eds. University of Alberta, Faculty of Extension, Edmonton, Alberta.

Dijkhuizen, A A, R B M Huirne, A W Jalvingh, and J W Visscher. 1996. Towards Integrated Information Systems for On-Farm Decision Support. Pages 54-60 *in* Proc. ICCTA '96/VIAS/NNAA Congr. on ICT Appl. Agric., Wageningen, NL.

Dhar, V, and R Stein. 1997. Intelligent Decision Support Methods: The Science of Knowledge Work. Prentice Hall, Upper Saddle River, NJ, USA.

Etgen, W M, R E James and P M Reaves, Eds. 1987. Dairy Cattle Feeding and Management. John Wiley & Sons, New York, U.S.A.

Fiez, E A. 1998. Flow of Records Within NCDHIP. U.S. National Dairy Database.

HTTP://www.inform.umd.edu/EdRes/Topic/AgrEnv/ndd.

Goldsmith, P D, J-F Forest, R Lacroix, and K M Wade. 1996. A Framework for Information Cataloguing and Management Information System Design. An Economic Perspective. Page 688 *in* Proc. 6th Int. Conf. Comput. Agric., Cancun, MX.

Goodhue, D L, Wybo, M D and L J Kirsch. 1992. The Impact of Data Integration on the Costs and Benefits of Information Systems. Management Information Science Quarterly, Vol 16, No 3 (September): 293-311.

Groeneveld E and T Gutzmann. 1990. Strategic Considerations in the Choice of Basic Software and Hardware for Information Systems in Animal Production. Computers and Electronics in Agriculture, 5: 135-149.

Groeneveld E and P Lacher. 1992. An Open information System for the Swine Production and Marketing Industry: Its Scope, Topology and Telecommunications Strategy. Computers and Electronics in Agriculture, 7: 163-185.

Hall, GM. 1991. Strategy, Systems, and Integration: A Handbook for Information Managers. Tab Books: Professional and Reference Books. Pa, U.S.A.

Hogeveen, H, M A Varner, D S Bree, D E Dill, E N Noordhuizen-Stassen, and A Brand. 1994. Knowledge Representation Methods for Dairy Decision Support Systems. J Dairy Sci 77:3704-3715.

Jofre-Giraudo, E, D H Streeter and W Lazarus. 1990. The Impact of Computer Information Systems on Dairy Farm Management Decisions. Agribusiness, Vol. 6 No. 5: 463 - 474.

Koorn, D S. 1996. Electronic Data Interchange to And From Dairy Farmers. Page 40 in Proc. ICCTA '96/VIAS/NNAA Congr. on ICT Appl. Agric., Wageningen, NL.

Kroeze G H, C Lokhorst, H J van de Beek and F de Vries. 1996. Proceedings - ICCTA '96/VIAS/NNAA congress on ICT applications in Agriculture: 611-616.

Malone T W and K Crowston. 1994. The Interdisciplinary Study of Coordination. ACM Computing Surveys, Vol. 26(1): 87-119.

McFadden, Fred R, J A Hoffer. 1994. Modern Database Management. 4th ed. Benjamin / Cummings Publishing Company, Inc., Redwood City, CA.

Misztal, I and T J Lawlor. 1999. Supply of Genetic Information – Amount, Format, and Frequency. Symposium: Information Technology for Animal Breeding Purposes. J Dairy Sci(82): 1052-1060.

Pedersen, J, G A Pedersen and U S Nielsen. 1996. Functional Traits in Danish Cattle Breeding. Pages 30 - 33 *in* Annual Report; National Committee on Danish Cattle Husbandry.

Pietersma, D, R Lacroix, and K M Wade. 1998. A Framework for the Development of Computerized Management and Control Systems for Use in Dairy Farming. J Dairy Sci(81): 2962 - 2972.

Rowe E. ed. 1997. Computer use is routine for them. Hoard's Dairyman, 1997(Jan): 50.

Schmoldt, D L, and H M Rauscher. 1994. A knowledge management imperative and six supporting technologies. Comp. Electron. Agric. 10:11.

Strong, D M, Y W Lee, and R Y Wang. 1997. 10 Potholes in the Road to Information Quality. Computer. 30 (Aug): 38-46.

Thompson-James, Margaret. 1999. Computer Use and Internet Use by Members of Rural Households. Rural and Small Town Canada Analysis, Vol. 1, No. 7 (May 1999), Statistics Canada. Catalogue no. 21-006-XIE.

Todd, P, and I Benbasat. 1992. The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. MIS Quarterly 16(3):373.

Tomaszewski, M A. 1992. Using Advanced Computer Technologies to Increase Extension Effectiveness. J Dairy Sci 75: 3242-3245.

Tomaszewski, M A. 1993. Record-Keeping Systems and Control of Data Flow and Information Retrieval to Manage Large High Producing Herds. J Dairy Sci 76: 3188-3194.

Wybo, M D. 1993. An Assessment of the Validity of the Subunit Interdependence Construct in the Context of Information Systems. Working Paper, McGill University, Faculty of Management, Montreal, Quebec.

Chapter 4. Exploiting the Internet to Support On-farm Decision Making In a Dairy Breeding Information System

A A Archer, R Lacroix and K M Wade

Having described the systems approach for investigating industry-wide information systems and having used it to analyse a dairy breeding information system, it became evident that automating and integrating certain functions in the system would improve system functioning from the point of view of on-farm decision making. The potential for using the Internet in a framework as a possible structure to integrate information sources and facilitate knowledge-based decision-support needed to be investigated. In this chapter the systems approach was employed to analyse a dairy breeding information system on the Internet. Having applied the systems approach in Chapter 3, there was a basis for comparing how well the Internet was being utilised for dairy breeding information as well as discussing of potential means of further exploiting it to serve as part of a framework for supporting on-farm dairy breeding.

Abstract

This paper examined the current state of the dairy breeding information system (DBIS) on the Internet as well as investigating means to improve its effectiveness through better use of the Internet. The examination of the DBIS on the Internet was based on the systems approach method – the goals, boundary and system structure and information functions were described and the performance was evaluated in terms of how to improve the exploitation of the Internet for on-farm decision-making. The information flowing onto the farm in the system was categorised either as support or herd output information which represents different aggregation levels of information and relates to different levels of intelligence density. Web sites, representing DBIS sectors, were examined for information services – either current or potential – available specifically on the Web. Support information is currently deployed more appropriately over the Internet than herd output information. Herd output information was available in high volume from most sectors via the Internet but the Internet could be further exploited to meet on-farm user needs as it allows for rapid, convenient information retrieval. The potential impact of current technology includes the delivery of information at more appropriate levels of intelligence density by providing farm specific herd output information to a given on-farm user when requested; filtering data to provide appropriate aggregates of information; on-line tracking to learn user preferences of information delivery; and changes to functions and structures of the DBIS to better serve user needs. As the attitudes of participants change with regard to their own role as on-line information consumers, and as the technology is used more transparently, the true value of certain information for making key decisions on-farm will become more apparent. Eventually, some functions may be implemented globally while others may be modified to support specific niches as sectors become more highly integrated. The exploitation of the Internet in the DBIS may be improved through the use of a common exchange mechanism and the employment of knowledge-based decision-support systems to organise, integrate and add knowledge to the information products.

Introduction

On-farm breeding decisions depend largely on information that is available from many off-farm industry participants. Information flows to these organisations from the farm, among these organisations, and from them back to the farm. Archer et al., (see Chapter 3) described a system's approach as a method of analysing an industry-wide information system according to the stated goal, boundary, functions, components and performance of the system.

The Canadian dairy breeding information system (DBIS) described by Archer et al. (see Chapter 3) is a functional analysis that uses the system's approach. The goal of that system was to describe information that originates from many sources in the industry and is available for the support of decision making on dairy farms for traits of commercial interest. The functions within the boundary of the DBIS (see Chapter 3) included on-farm functions (cow culling, sire selection and mating) and off-farm functions (sire analysis, genetic evaluations, milk recording, research, extension and instruction in dairy cattle breeding, conformation analysis, animal registration, and herd-book keeping). These functions were performed by 6 component sectors: the dairy herd improvement centre (DHI), the genetic evaluation centre (GE), artificial insemination centre (AI), the breed association (BA), government, and research centres. While the DBIS seemed to perform adequately, given current media of information exchange, the flow of information products on-farm may not be optimal because of the limitations of the on-farm decision-maker in handling the information products – multiple sources and large volumes are difficult for human agents to integrate and process in decision making.

Modern information highways (IH) have emerged as powerful tools which could enhance the performance of industry-wide information systems. The best known of these - the Internet - is becoming a popular means of information exchange in the dairy industry. As its popularity increases, the question of how to exploit the Internet intelligently to maximize the benefits of information handling in the dairy industry, must be addressed.

The Internet is a useful vehicle for information exchange. Information exchange can comprise one-to-one (individual to individual) and one-to-many (individual to group) information transfer (Lacroix et al., 1998a). Several services have been developed to express these forms of information exchange: e.g., Electronic mail (E-mail), file transfer protocol (FTP), TELNET, USENET groups and list-servers (BA et al., 1997; Davis et al., 1994; Egeberg and Rice, 1994; Fournier et al., 1997; Fowler et al., 1994; Innes, 1994; Lacroix et al., 1998a, Rimmington et al., 1994; Varner and Cady, 1993). Except for E-mail, the above services are all "pull through" (initiated by the user) on-line technologies (Lacroix et al., 1998a). Web sites transfer multimedia information on the Internet, allowing users to access documents as data, text, images and sound by a single graphical user interface (Winstead and Gershon, 1996). They are the foundation for the distributed information system known as the World Wide Web (WWW) which includes all of the services described above. Users can access Web sites examine available information and retrieve what they require. While the WWW is only a part of the Internet, its ease of use and accessibility have allowed it to dominate the Internet's potential utility for many industries. Dertouzos (1995) described a true "information super-highway" as being widely available, easy to use and the foundation for countless useful activities.

The WWW is becoming a means of transparent, inter-operating mass communication as the medium supporting other information technologies and supported by a critical mass of communication technologies. Intranets and Extranets, within and across organisational networks using the WWW, have become important technologies to support users in many industries (Lloyd and Boyle, 1998; BA et al., 1997; Covill, 1998). Technologies supported by the WWW include information transactions (e.g. supply chain management, enterprise resource planning, customer relationship management and legacy applications); commercial services (e.g. billing, customer service and merchandising); knowledge based search and analysis; and data warehousing (Brick and Henry, 1999; and Kimball, 1999). The improved efficiency of workflow, simplified and streamlined processing, along with a value-added focus and potential for real-time delivery, can create true e-commerce (Trustman and Meshako, 1999). The technologies being developed to support the WWW include Web PCs, Web TVs, as well as wireless devices (Web pagers, cell phones, etc) (O'Malley, 1995).

The use of electronic networks, including the Internet, to support agricultural systems has been investigated (Davis et al., 1994; Egeberg and Rice, 1994; Fournier et al., 1997; Fowler et al., 1994; Innes, 1994; Rimmington et al., 1994; and Varner and Cady, 1993). For example, bulletin boards have been used as a local or regional electronic network to exchange and maintain extension information among agricultural information systems. Tomaszewski (1992 and 1993) and Jones and Spahr (1991) discuss the establishment of electronic modes of exchanging milk records and other types of information among farms and DHI and other off-farm sources. They suggest that there would be greater advantages to the movement of information from the perspective of costs, time and format of capturing and retrieving the data. The system could be anchored by an electronic data interchange (EDI) standard that would improve efficiency and reduce costs by providing a predetermined format and sequence for exchange (Koorn, 1996; and Silver and Silver, 1994). The Internet has been employed to support extension offices and basic data-type information such as weather reporting and futures' prices (Fowler et al., 1994). These types of services are "push through" information services. Dairy-L (Varner and Cady, 1993; 10) and LaiToile (16) are one-to-many types of information services for English and French producers respectively. They represent discussion forums where participants share information in the form of ideas, know-how and references to assist each other.

Organisations that support dairy producers are exploring how best to provide services to their clients via the Internet. Investigating the deployment of the DBIS over the Internet would contribute to the understanding of dairy systems on-line and improve the exploitation of the Internet. The objectives of this article are to examine the deployment of a DBIS on the Internet; to study the impact of the Internet on the DBIS; and to investigate means to improve the DBIS through better use of the Internet. This necessitates an evaluation of the current use of the Internet as well as its potential impacts on dairy information systems. Based on the prior analysis of the DBIS by Archer et al. (Chapter 3), it should be possible to determine how to improve the benefits from using the Internet. The deployment over the Internet is a means by which this system could employ information technologies to improve on-farm decision making.

Materials and Methods

The examination of the DBIS on the Internet was based on the system's approach method developed by Archer et al. (Chapter 3) and Figure 4.1 is an adaptation from that work showing the functions, structure and flow of information among them. The system characteristics examined were the goal, boundary, functions, components and some attributes of the DBIS. This examination concerned dairy breeding thus the boundary included activities that are involved with the use and exchange of genetic information. The individual components include the GE, DHI, AI, BA, on-farm, research and government sectors. Functions in the DBIS can be considered according the information they produce. The DBIS was grouped as support information (i.e., information general to a herd / farm) and herd output information (i.e., information specific to herd / farm). Sectors having functions primarily concerned with herd output information include GE, BA, AI, and DHI and those primarily concerned with providing support information are research institutes and government.

For clarification an explanation of intelligence density, types of information products and their relationship and utility in decision making is useful using Figure 4.2 as an illustration. Dhar and Stein (1997) uses the term intelligence density as a heuristic measure of the amount of "decision support information" available to the decision maker from the information products. Information products for the decision-making process can be procedural ('how to') or declarative ('what is') in nature (Schmoldt and Rauscher, 1994; and Hogeveen et al., 1994), where the products at the knowledge end informs the participant about the decision-making *process* while less intelligence dense information inform the participant about the *state* of entities on-farm about which the decision will be made. In Figure 4.2 the information products (i.e., signals, data, information, knowledge) are arranged by increasing intelligence density moving from declarative to procedural knowledge. In general herd output information has a lower intelligence density and provides declarative information about farm management and herd performance while support information has higher intelligence density and provides procedural information aiding the decision-making strategy. To illustrate this further a dairy manager making decisions concerning replacement

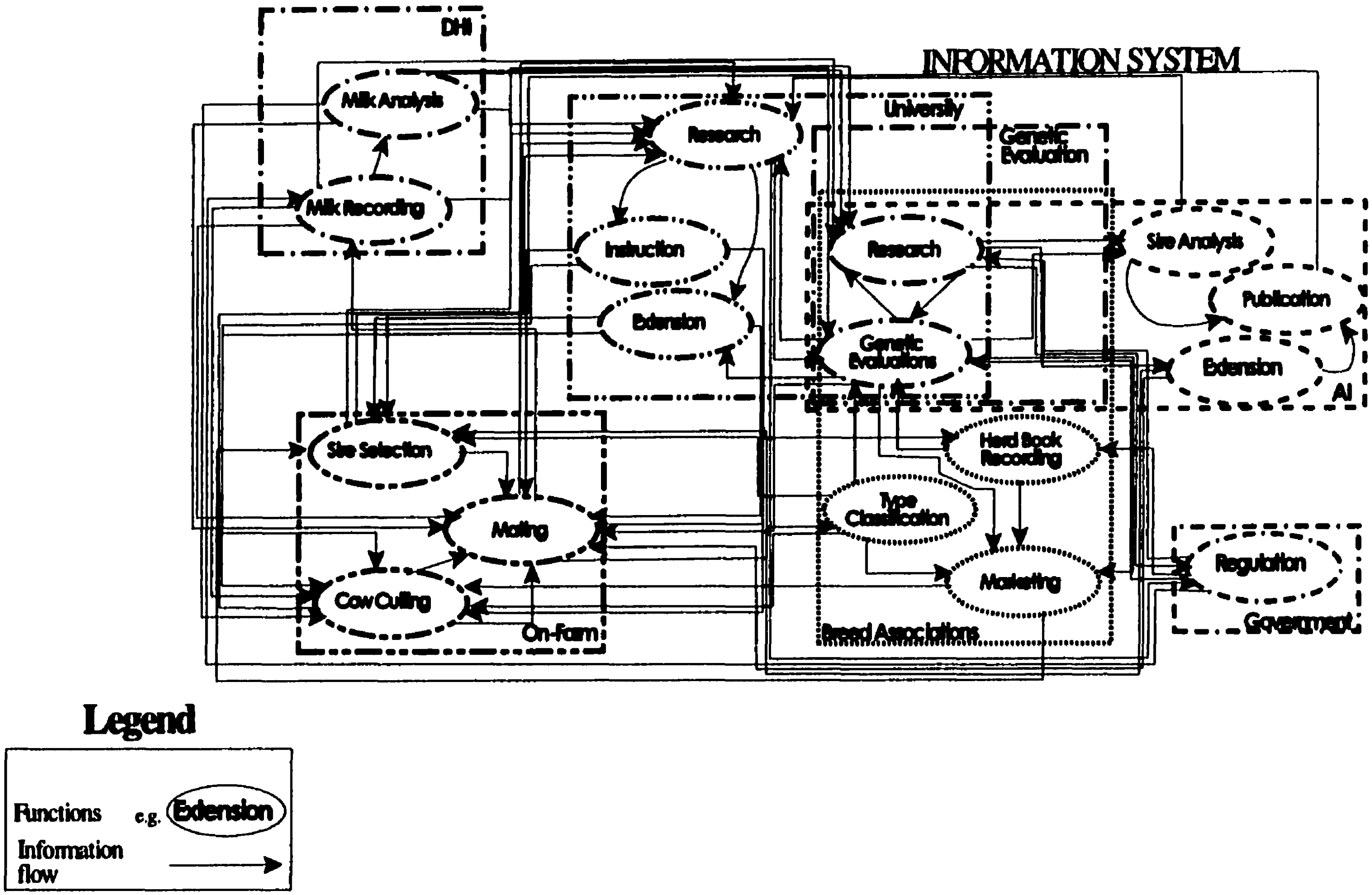


Figure 4.1 A Dairy Breeding Information System (DBIS)'s components and functions.

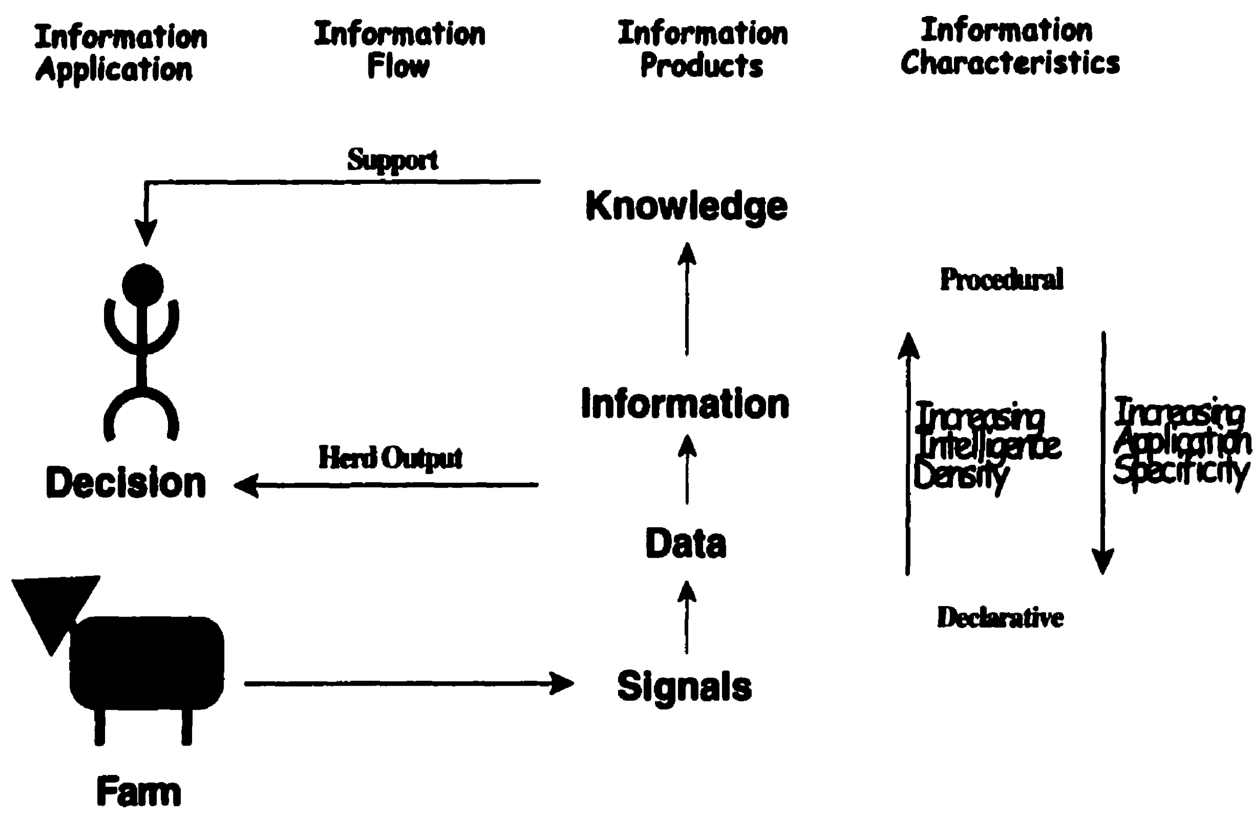


Figure 4.2 Intelligence density and application specificity of information products for on-farm decision making.

heifer requires information: an extension document (support) on selecting replacement heifers imparts procedural knowledge on *how to* select heifers, while farm data (herd output) impart declarative knowledge on *what* is the state of the heifers being considered.

Both support and herd output information products are required for effective on-farm decision making, but can be incorporated differently with respect to the on-farm decision maker based on their strengths in incorporating and using information (Chapter 3). The system was analysed functionally by investigating Web sites of organisations that participate in the DBIS examining the information products available through their Web sites and comparing this information with what is currently available from these organisations through other media. The sites were explored as to their use of Internet technology and the resulting usefulness in carrying out DBIS functions from the vantage of on-farm participants using a common PC based Web browser.

A sample of Web sites from DBIS were examined. Information services of the DBIS that were either currently or potentially available through the Internet, specifically through the WWW, were examined. The Web sites originated from the United States, Canada, and The Netherlands, however the primary focus was on the Canadian DBIS. The Web sites explored are not an exhaustive list that serve on-farm dairy breeding decision making, but do effectively demonstrate the potential of the Internet as it exists.

Discussion

Analysis of DBIS On-Line: Support and herd output information use Internet tools differently to implement their functions on the Internet. In the case of herd output information, FTP is often employed to exchanges data and information among DHI and GE, GE and AI, and BA and DHI. It is possible to upload milking records electronically from farms by a technician. Presently it is possible to download test-day results or reports for use in PC-based management software (Rowe, 1997; Spahr, 1993; 11, 18, 19).

The US Holstein Association and Holstein Canada (HC) have both automated herd book maintenance to register cattle by the Internet from farm (Lacroix et al., 1998a; 14, 15). Cow pedigrees can be accessed from Web sites using Web-based client-server queries as well. The latter function interfaces with an individual animal search function. Both BA allow any animal's genetic merit or pedigree in their herd-books to be accessed (14, 15).

The GE in Canada allows on-line CGI-based query access to its data-base of the Canadian dairy herd (Lacroix et al., 1998a; 7). This site also allows participants to verify animal pedigree and genetic merit (although downloaded documents are not official). Genetic evaluations are available from Canadian Dairy Network (the GE) as flat-file FTP down-loads. Other Canadian sectors (i.e., BA and AI) also provide genetics proofs either by creating links directly to CDN or by downloading and transforming the information from the site to their own specific needs (1, 3, 9, 12, 15).

The herd output information provided by AI generally has marketed genetic products (semen, embryos) on-line, providing profiles and high-lighting strengths (Lacroix et al., 1998b; 1, 9, 12, 20) of their products over others. Some sites take more advantage of the WWW's multi-media capabilities to publish attractive, high quality text and graphical information of the products and services they are offering. However, WWW tools might be better exploited, to, for example, sell semen or to contract breedings on-line.

With support information in the DBIS, many government and research sites have active on-line libraries and data-bases set up as catalogues of aggregated agricultural resources, using WAIS, Gopher and FTP sites (2, 3, 4, 5, 8, 17, 21). These sites inform users of research, extension bulletins, education and training opportunities, as well as new technologies in the industry. For example, extension papers, policy papers, and news releases are available from the National Dairy Database of the USDA (27) and multimedia documents and research-article archives can be accessed from Oklahoma State University's WWW Virtual Library (17). List-servers support a range of dairy production and improvement topics of general interest to a specified group of participants (Fournier et al., 1997; Varner and Cady, 1993; 6, 10, 13, 16) and E-mail provides specific technical support in the DBIS on the Internet. Other off-farm sites (DHI, GE, BA and AI) also provide support information such as technical information on the use of test-day information to generate lactation information from the Quebec DHI (19) or Canadian Dairy Network (7).

Support-type information from the DBIS appears better suited to the Internet than the herd output-type information at this present time: it flows to the farm in formats appropriate for a human end-user such as texts and images (Figure 4.2). The nature of the information being provided (with the exception of information such as commodity pricing or weather reporting) often allows its contents to have a "longer shelf-life" than herd output information. Support information tends to be associated with general application (i.e., general knowledge) and has broad relevance to many users. This makes publishing information to the Web with large DBIS-membership groups easier. At this level of general application the accuracy and reliability seem appropriate. Also, text and visual documents do not require much re-formatting adjustment for effective implementation. Improvements are possible, such as automating the search process – currently being done by some participants (17, 27) but not by others.

The use of currently available one-to-one or one-to-many Web tools has not made herd output data much more accessible on-line. Some functions are available on-line through current technologies but the quality of the information at the user-end should be more carefully considered. Strong et al. (1997) describe several issues that can impede the access of high quality information through the Internet. Some of the issues relevant to DBIS herd output information include inconsistent definitions and formats of information among systems; difficulty in timely access due to information volume; changing information relevancy as users' tasks and providers' environments change; and lack of automation across independent organisation sources. These issues make integration of large volumes of specific data difficult - a fact that is particularly evident with the provision of genetic evaluations.

Perhaps the most relevant limiting issue mentioned by Strong et al. (1997) limiting herd output information utility in the DBIS was flexible tools on the Internet that can aid the analysis of data provided across information sources. The current use of Web tools can access information from the various sources providing it. However relating this information together with other information extracted, to specific software problems the participant may have, is not yet available. For example, retrieval of the latest genetic proofs is possible but a producer cannot yet relate it with specific DHI herd information using on-line tools.

Unlike support-type information, significant investments may be required to overcome cost and format issues in challenges for DBIS herd output information on the Internet. Sectors should create or adapt Web tools to facilitate information processing and exchange for specific user needs.

Impact and potential impact of the current use of the Internet on the DBIS: The Internet can affect the DBIS's functions, structure, and effectiveness. The greatest immediate impact of the Internet should be on effective information delivery in the system. The time required to appropriate DBIS information products through the Internet is equal or far less than in other media, thus delivery time should be reduced. Furthermore, the pull-through nature of the information allows DBIS participants to access information at their convenience.

Pull-through access permits user-directed information access which can increase the intelligence density. Currently on-line herd output information is mostly accessed and downloaded on-line. However an on-farm participant should be able to, for example, request specific genetic evaluations from GE that pertain to his/her herd and information should be aggregated according to the goal preferences of the participant. Retrieving and transforming farm specific information reduces the burden to filter downloaded data on-farm, thus improving the opportunity for effective utilisation of the content that does flow to the farm. Redundant or trivial information can be avoided. Web sites providing such services would increase the intelligence density to an appropriate level, reducing the time and difficulty in decision making.

Such technologies could prevent the creation of an "information jungle" in the on-line DBIS where volume but not accessibility is high and there is a large number of information providers to consult. Herd output information flowing to the farm is declarative in nature, relatively low in intelligence density, specific to cows on-farm at a narrow point in the herd's life and high in volume (see Figure 4.2). An "information jungle" might lead to responses to on-farm information requests that do not provide relevant aggregates of useful information (Waksman and Harkin, 1996; and Strong et al., 1997) or to lengthy searches of many Web sites for information products required for a given decision-making event. Preventing a scenario of high volume from many sources of information will enhance decision-makers' strengths rather than their limitations (Chapter 3).

The impact of the Internet on functions of the DBIS is expected to be realised over a medium to long time horizon. Knowledge of both support and herd output functions will become more user-specific as sub-sets of user preferences are tracked over time; on-line preferences can then be provided. For example a DHI Web site could "learn", through collating information retrieval patterns through time, that a certain farm manager accesses herd average production data and reproduction efficiency data at a regular intervals, but, for example, has no interest in feeding information. The organisation could send this subset of information to the manager as an E-mail attachment, at predetermined intervals.

Some functions in the DBIS should be able to cross international boundaries and become global. For example, Canadian on-farm participants will be able to buy Dutch semen or US embryos. The Canadian GE may contract out to perform genetic evaluations for small or less developed countries lacking expertise or computing power. Providers of functions should consider a global market for their products while others, such as milk recording, by virtue of their regional focus might have difficulty doing so.

Teaching and extension could be improved through the Internet. Many Web sites serve merely as pointers to other means by which these functions are accomplished. However, the University of Ottawa has created a media bank of text, diagrams and photographs accessible, via FTP and WWW servers, by Francophone universities world-wide for french biology education (BIODIDAC project, 1998). Participants can modify, assemble and incorporate the information into up-to-date courses according to individual needs. Eventually multimedia documents and courseware will become more commonplace. No similar examples were found in the dairy industry from the sites surveyed; the need for Internet access in french in the Canadian dairy industry is clear (Archer et al., 1996; Lacroix and Wade, 1996b; Manguerra, 1997). A step-wise approach could be used to develop educational support resources for the DBIS. This would enable participants to benefit through distance education and technology transfer.

As more organisations are providing information through the Internet, greater coordination and cooperation among organisations across sectors may result in changes to the structure of the DBIS. Changes may occur to minimise costs and maximise resource use. For example three sectors, AI, GE and BA, provide genetic evaluation information (see Figure 4.1). Redundancy is thus created in the information provided. The DBIS might be better served if AI and BA shed their genetic evaluations functions or merged into a single sector that could provide the service of interpreting the evaluations. Data storage and processing could be integrated to provide uniform information products. Network technology has permitted the integration of DHI organisations in Canada in terms of storage and retrieval services (Biron et al., 1998). Integration of herd output information will lead to a reduction in time spent by participants at multiple sites. Changes should occur only as the effectiveness of information handling is improved.

Participants of the DBIS must be prepared for electronic exchange to take place. Resistance may arise due to fear of technology (Todd and Benbasat, 1992), security concerns, cost or lack of computer equipment and training. Attitudes about the derived benefits of the available information, its format and the medium, can encourage participants to access it. Participants must recognize the role and value of information to their endeavours. The kind of information provided at Web sites can be well or poorly adapted to the WWW (Archer et al., 1996). For example, some smaller BA provided poorly adapted, static information on-line, such as host name and contact information. Good use of the WWW (e.g., using multimedia or providing dynamic, rapidly changing information) was specific to a participant and not to the sector or function demonstrating insightful thinking about the potential use of the medium. Information is exploited well over the Internet when users perceive their role well (Dertouzos, 1995). Producers, researchers, and industry members must think of the information services they receive and produce as commodities or goods and assign commercial value to them (Goldsmith et al., 1996). While transition to this mind set can be difficult, the benefits of the transformation include better use of data and resources to produce specific information that is valuable to participants. Furthermore, transparent use of the WWW allows participant to focus on the role information plays in making key decisions in management strategy.

Views about the Internet's utility may be reflected by adoption of computer technology by producers (Newman et al., 1999). In the European Union 10% to 20% of producers have access to computers (Waksman and Harkin, 1996; and Koorn, 1996). Only 3% of the 29% of rural and small town residents in Canada (in 1995) reported typical use of the Internet for business purposes (Thompson-Jones, 1999). A survey of Quebec dairy industry participants suggested that computer ownership and use among producers in that province was much higher than the rest of Canada (as much as 77 %), however it was a limited survey (Lacroix et al., 1996). Many in this group may not have WWW appropriate PCs (only 9.3% reported using the Internet for work on the farm). The falling costs of computers, other means of connecting to the Internet (Web TV, Internet cellular phones, etc.) and the increasing importance of WWW will induce more producers to become connected (Lacroix and Wade, 1996a).

Improving DBIS exploitation of the Internet through available technologies: Various information technologies may be used to further improve DBIS exploitation via the Internet (Archer et al., 1996; Manguerra, 1997). Part by this requires DBIS members to analysing how to exploit the information in the system. Dynamic programs could be developed to filter, transform and aggregate the herd output information. They could be downloaded when the on-farm participant requests information on-line. For example, to automate genetic information retrieval, a dynamic program could filter out evaluations not from the specific farm. Or an "intelligent" agent could incorporate lactation and genetic evaluation records, regularly downloaded from the Web, into the on-farm electronic storage device, based on a knowledge of the sources, the relationship among the data, and the preferred storage format. The information stored would require less human processing to increase the intelligence density to an appropriate level at the time of decision-making. Participants should determine the information they need specifically, and possess the physical and organizational requirements for its integration into their data structure. Each participant on the server side should evaluate how exchange, storage, and processing and the specific functions of the DBIS could be improved by the Internet. Providing a family of information products to the participant by collaborating with other DBIS members within and across sectors, to maintain consistent presentation, would be key.

A framework for common exchange of herd output information would expedite timely exchange of accurate information over the Internet. Electronic data interchange (EDI) is a method to automatically exchange information between computer systems by structured standardized messages (Lokhorst and Kroeze, 1996; Koorn, 1996; and Silver and Silver, 1994). While development of a North American dairy EDI standard has been considered (Tomaszewski, 1992, 1993), standardization would be difficult due to high regional differences and the autonomy of organisations. The loss of autonomy throughout the system may not be agreeable to all participants (Goodhue et al., 1992). Data modelling is another method of common exchange that determines existing information resources and relationships among them (McFadden and Hoffer, 1994; and Harrington, 1998). Information retrieval from the implemented model by users is supportable via the WWW and accommodates heterogeneous data sources (McFadden and Hoffer, 1994). Common exchange is useful for herd output information in the DBIS, where 1) data are being reused; 2) speed is important; and 3) a high volume of data is exchanged among many participants (Spahr, 1993). Rapid transfer of information overcomes the "distance" between sectors and simpler and structured control of the data removes the necessity of human handling. The latter leads to optimized storing and processing, and more complete and accurate information aggregation and transformation (see Chapter 3). A common exchange framework could expedite on-farm functions such as production and reproduction monitoring, official herd reporting and milk supply prediction (Cue and Fletcher, 1998; Dindorp et al., 1996; Koorn, 1996; Lokhorst and Kroeze, 1996, Lescourret et al., 1993; and Lescourret et al., 1994). Combining a common exchange framework with other emerging technologies such as remote sensing, electronic identification of animals and automatic recording of events (Kalter et al., 1992) would benefit DBIS herd output functions. For example, on-farm data recording quality and content could be improved (Tomaszewski, 1992, 1993), information flow among off-farm participants could become more cost-effective, and DBIS functions' performance on-line, such as herd book recording, sire selection, and milk recording would be more transparent.

The use of decision support systems (DSS) in the DBIS via the Internet would increase the intelligence density of herd output functions on-line. These are knowledge-based frameworks that use human problem solving methods, such as heuristic reasoning, to pre-process information and help human decision making. The use of DSS could automate some procedural knowledge and allow them to be dynamically applied to on-farm decision-making instead of flowing onto the farm as support information products captured by the farm manager (see Figure 4.2). There are many DBIS activities that could employ DSS: advisory, strategic planning, and diagnostic processing functions (Strasser et al., 1997; Lacroix et al., 1996; Wade and Lacroix, 1997; Pietersma et al., 1998; Morimoto et al., 1995; and Spahr et al., 1988). Through the use of a common information exchange framework on the Internet, a DSS could, for example, be used diagnostically to advise producers on the re-calibration of economic weights for selection indices. The use of DSS would also be an appropriate way of implementing many human-processing functions. The Internet appears to be suitable for DSS development since both technologies rely on well-structured information-management techniques for their development (Manguerra, 1997). The Internet could collect, manage and distribute up-to-date sources of knowledge and information useful in supporting various decision-making efforts throughout the DBIS. Given that dairy producers are usually at a distance from experts in the industry, DSS distributed over the Internet should enhance the availability of expert reasoning to industry members when required.

Conclusions

The Internet is a useful medium for improving the effective and rapid flow of information for many DBIS functions, however it is not currently being fully exploited. The DBIS functions and structure may change over time reflecting the impact of the Internet and the changes should reflect a directed effort of organisations to serve information needs of on-farm users. As the Internet becomes a commonly used tool DBIS members must be prepared to exploit it in an appropriate manner: enhancing information exchange by deploying appropriate information on-line and creating applications that take advantage of

WWW faculties. Applications should reflect the nature of the support and herd output information produced in the system and anticipate needs of specific, user-defined demands for information. Computer based technologies used to accomplish this should organise, integrate and add knowledge to the information products providing a full range of necessary DBIS functions on-line for all participants. Increased usage of the Internet will reduce participant costs and improve the potential for highly developed and differentiated services required by all members of the dairy industry. Further studies in the DBIS should explore the basis for creating such technologies and applications.

Acknowledgments

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, les Fonds de l'Autoroute de l'Information du Gouvernement du Québec, and the Cattle Breeding Research Council of Canada.

References

Archer, A. A., R. Lacroix, and K. M. Wade. 1996. A network of information for the dairy industry. J. Dairy Sci. 79 (Suppl. 1):174. (Abstr.)

Ba, Sulin, Ravi Kalakota, and Andrew B Whinston. 1997. Using client-broker-server architecture for Intranet decision support. Decision Support Systems 19(1997): 171-192.

The BIODIDAC Project. 1998. <biodidac.bio.uottawa.ca/Info/biodidac.htm>.

Biron, Sylvain. 1998. Outils génétiques et expertise animale. Des notes d'un cour offert par le PATLQ, (Septembre), 1998.

Biron, S, B Farmer, R Moore, H Keurentjes and M Caron. 1998. Le contrôle laitier du troisième millionaire. Pages 113-135 *in* Symposium sur les bovins laitiers. Conseil des Productions Animales du Québec inc.

Brick, B, and J Henry. 1999. Enterprise Portals: Not Just Information Delivery. Intelligent Enterprise 2(16): 24-27.

Covill, Randal Jorde. Implementing Extranets: The Internet as a Virtual Private Network. Digit Press, Boston. 1998.

Cue, R C, and J D Fletcher. 1998. Predicting Milk Supply. Pages 11-15 *in* 1998 Annual Research Report, Department of Animal Science, McGill University. Montreal, Canada

Dhar, V, and R Stein. 1997. Intelligent Decision Support Methods: The Science of Knowledge Work. Prentice Hall, New Jersey, USA.

Davis, J C, G A Parmer, B R Smith, and E H Simpson III. 1994. Providing Internet connectivity for remote Alabama offices. Pages 718-721 *in* Proc. 5th Int. Confr. Comp. Agric. Orlando, FL

Dertouzos, M L. 1995. Communications, computers and networks. Pages 22-29 *in* Sci. Am., special issue: The computer in the 21st century, NY, New York.

Dindorp, U, M Flensted-Jensen, and A P Ravn. 1996. A standard format for data exchange - SFD. Pages 368-373 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Egeberg R D, and D G Rice. 1994. ExtNet: extension information delivery in North Dakota. Pages 702-706 *in* Proc. 5th Int. Confr. on Comp. in Agric. Orlando, FL.

Fournier, A, R Lacroix and K M Wade. 1997. Ouvrir les horizons grâce aux groupes de discussion. Le Producteur de lait Québécoise. 17(9):38.

Fowler, S R, J L Novak, and E H Simpson III. 1994. Modem-accessible information services for agric. and ext. Pages 617-621 *in* Proc. 5th Int. Conf. on Comp. in Agric. Orlando, FL.

Goldsmith, P D, Jean-Francois Forest, R Lacroix, and K M Wade. 1996. A Framework for Information Cataloguing and Management Information System Design: an Economic Perspective. Page 688-701 *in* Proc. 6th Int. Conf. Computed. Agric. Cancun, MX.

Goodhue, D L, M D Wybo, and L J Kirsch. The impact of data integration on the costs and benefits of information systems. 1992. MIS Quarterly 16(3): 293-311.

Hogeveen, H, M A Varner, D S Bree, D E Dill, E N Noordhuizen-Stassen, and A Brand. 1994. Knowledge Representation Methods for Dairy Decision Support Systems. J Dairy Sci 77: 3704-3715.

Innes, R J. 1994. Designing and implementing a low-cost county office network. Page 691-694 *in* Proc. 5th Int. Confr. Comp Agric. Orlando, FL.

Jones, L R and S L Spahr. 1991. IDEA: Intelligent Data Retrieval in English for Agriculture. AI Applications 5(1): 56-66.

Kalter, R J, A L Skidmore, and C J Sniffen. 1992. Distributed intelligence and control: the new approach to dairy farm management. Pages 171-176 *in* Proc. 4th Int. Confr. Comp. Agric. Ext. Prog. Orlando, Fl.

Kimball, R. 1999. Working in Web Time. Intelligent Enterprise 2(16): 28-32.

Koorn, D S. 1996. Electronic Data Interchange to And From Dairy Farmers. Pages 40-44 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Lacroix, R, K M Wade and A Fournier. 1996. Nouvelles en Vrac. Le Producteur de lait Québecoise. 17(4): 22-23.

Lacroix, R, J Huijbers, R Tiemessn, D Lefebvre, D Marchand, and K M Wade. 1996. Adding value to Dairy Herd Improvement information through the use of Fuzzy Logic. Pages 885-891 *in* Proc. 6th Int. Conf. Comput. Agric., Cancun, MX.

Lacroix, R, and K M Wade. 1996. Expected benefits from the use of the information superhighway in dairy herd improvement. Pages 175-178 in Proc. 30th Biennial Session of the International Committee for Animal Recording on Performance recording of animals: State of the Art. EAAP Publication No 87, 1996. Veldhoven, The Netherlands.

Lacroix, R, and K M Wade. 1996. L'Autoroute de l'Information en Agriculture: Pas Pour Demain, Pour Aujourd'hui. Le Producteur de lait Québécoise. 17(1):26-29.

Lacroix, R, K M Wade, S Des Marchais and R Marchand. 1998. L'Élevage des Bovin Laitiers et Internet; l'Information Génétique (Partie 1). Le Producteur de Lait Québecois, 18(10): 53-55.

Lacroix, R, and K M Wade. 1998. L'Élevage des Bovin Laitiers et Internet. Partie 2: L'Offre de Produit Génétique. Le Producteur de Lait Québecois, 18(11): 43-46.

Lokhorst, C, and G H Kroeze. 1996. Farm comparison by means of automatic data transfer.

Pages 225-230 *in* Proc. 3rd Polish-Dutch Bilateral Conference, Wageningen, NL.

Manguerra, H B. 1997. Internet-based decision support systems. Resource: Engineering and Technology for a Sustainable World, Vol 4(4):11.

McFadden, Fred R and J A Hoffer. 1994. Modern Database Management. 4th ed. Benjamin / Cummings Publishing Company, Inc., Redwood City, CA.

Morimoto, T, K Hatou, and Y Hashimoto. 1995. Intelligent Control for Plant Production System. Page 139-144 *in* Artificial Intelligence in Agriculture, A J Udink ten Cate, Ed. Oxford, Ox, UK.

Newman, S, Lynch, T and Plummer, A A 1999. Aspects of success and failure of decision support systems in agriculture. Pages 17-30 *in* Application of Artificial Intelligence, Optimisation and Bayesian Methods in Agriculture. H A Abbass and M Towsey ed. Queensland University of Technology Press, Brisbane (ISBN 1 86435 463 1).

O'Malley, C. 1995. The Wireless World: The Information Skyway is Coming Soon to a Wireless Device Near You, in Popular Science, Stepler, R L ed., 11: 56-62.

Pietersma, D, R Lacroix, and K M Wade. 1998. A Framework for the Development of Computerized Management and Control Systems for Use in Dairy Farming. J Dairy Sci(81): 2962 - 2972.

Rimmington, G M, H S Hawkins, and J E Hoare. 1994. LandcareNET--promotion of lateral information transfer. Pages 695-700 *in* Proc. 5th Int. Confr. Comp Agric., Orlando, FL.

Rowe E ed. 1997. Computer use is routine for them. Hoard's Dairyman, Jan:50.

Schmoldt, D L, and H M Rauscher. 1994. A knowledge management imperative and six supporting technologies. Comp. Electron. Agric. 10:11-30.

Silver, G A, and M L Silver. 1994. Implementing communications networks. Page 396 in Data Communications for Business. Boyd and Fraser Publishing, Danvers, MA.

Spahr, S L. 1993. New technologies and decision making in high producing herds. J. Dairy Sci. 76:3269-3277.

Spahr, S L, L R Jones, and D E Dill. 1988. Expert Systems -- Their use in dairy herd management. J. Dairy Sci. 71: 879-885.

Strasser, M, R Lacroix, R Kok, and K M Wade. 1997. An aggregated decision-support system for the recommendation of dairy cattle culling. J. Dairy Sci. 80 (Suppl. 1):149.(Abstr.)

Strong, D M, Y W Lee, and R Y Wang. 1997. 10 Potholes in the Road to Information Quality. Computer. v. 30, Aug. 1997, p. 38-46.

Thompson-James, Margaret. 1999. Computer Use and Internet Use by Members of Rural Households. Rural and Small Town Canada Analysis, Vol. 1, No. 7 (May 1999), Statistics Canada. Catalogue no. 21-006-XIE.

Todd, P, and I Benbasat. 1992. The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. MIS Quarterly 16(3): 373-393.

Tomaszewski, M. A. 1992. Using advanced computer technologies to increase extension effectiveness. J. Dairy Sci. 75: 3242-3245.

Tomaszewski, M A. 1993. Record-keeping systems and control of data flow and information retrieval to manage large high producing herds. J. Dairy Sci. 76: 3188-3194.

Trustman, J, and S Meshako. 1999. You Say You Want an Evolution. Intelligent Enterprise 2(16): 18-20.

Varner, M A, and R A Cady. 1993. Dairy-L: an electronic information exchange network for professionals advising dairy producers. J. Dairy Sci., 76: 2325-2331.

Wade, K. M., and R. Lacroix. 1997. The use of fuzzy logic membership values in ranking dairy cows for management purposes. J. Dairy Sci. 80 (Suppl. 1): 232. (Abstr.)

Waksman, G, and M Harkin. 1996. Green Telematics: Is Internet a Breakthrough for Farmers and Extension Workers? Pages 182-188 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Winstead, J, and N Gershon. 1996. Information Highways Internet and Web Services. Pages 21-28 *in* The Information Revolution: Impact on Science and Technology. Dubois, J-E and N Gershon, ed. Springer-Verlag, Berlin, Germany.

URL References

1 ABS Global. 1997. <http://www.absglobal.com>

2 Agriculture Canada Electronic Information System. 1997. <http://aceis.agr.ca>

3 Altagenetics. 1997. <www.altagenetics.com/dairy-menu.htm>

4 American Dairy Science Association. 1997. <http://www.adsa.uiuc.edu>

5 American Society of Animal Science. 1997. <http://www.asas.uiuc.edu>

6 Animal geneticists discussion group. 1997. <http://chuck.agsci.colostate.edu/>

7 Canadian Dairy Network. 1997. <http://www.cdn.ca>

8 Centre for the Genetic Improvement of Livestock. 1997. <http://www.aps.uoguelph.ca/cgil>

9 Centre d'Inseminasation Artificielle du Quebec, 1997. <www.ciaq.com/holstein.htm>

10 Dairy-L. 1997. dairy-l@umdd.umd.edu

11 Fiez, E. A. 1985. Flow of records within DCDHIP. <http://www.inform.umd.edu/EdRes/Topic/AgrEnv/ndd/dairy/FLOW_OF_RECORDS_WITHIN_NCDHIP.html

12 Gencor. 1997. <http://www.gencor.ca>

13 Genetic Evaluations Discussion Group. 1997. gedg@ansci.nw.uoguelph.ca

14 Holstein Association. 1997. <http://www.holsteinusa.com>

15 Holstein Canada. 1997. <http://www.holstein.ca>

16 LaiToile. 1997. http://infosys.AgrEnv.McGill.CA/laitoile/

17 Oklahoma State University: WWW Virtual Library. 1997. <http://www.ansi.okstate.edu/library/dairy>

18 Ontario Dairy Improvement Corp. 1997. <www.odic.com>

19 Programme d'Analyse des Troupeaux Laitiers du Québec. 1997. <http://www.patlq.com>

20 Select Sires. 1997. <www.selectsires.com>

21 USDA 1997. <http://cos.gdb.org/best/fedfund/usda/usda.form.html>

Chapter 5. Creation of a Unified Data Model for a Distributed Dairy Breeding Information System

A A Archer, R Lacroix and K M Wade

Chapter 5 presents a data model of the DBIS and presents a method for translating the prototype from the conceptual understanding of a DBIS into database terms. The model of the information system is represented in conceptual, logical and physical interpretations for domain experts, database developers and users respectively. In Chapters 3 and 4 a conceptual understanding of the nature of a DBIS emerged. In this chapter, the conceptual data model is more completely described and the logical and physical data models are explored. The method presented here is a tool that supports the over-all framework wherein the unified data model allows information to be integrated across organisations.

Abstract

A unified data model for industry-wide information in dairy breeding should improve clarity and utility of available information products and link information products to problems being addressed across organisations which can improve the decision-making capabilities of dairy-farm managers. The objectives were to develop conceptual, logical, and physical data models that were optimised from the on-farm decision-maker's perspective, and to be able to create views of the data which could be useful for decision making. An information engineering approach was used to develop the data model conceptually, logically and physically. Flat-file formats, validated by the Canadian Dairy Data Dictionary were used to develop the conceptual model while the logical model was a translation of the conceptual into relational database terms. Data used to populate the physical model came from three off-farm organisations (Québec DHI, Holstein Canada and Canadian Dairy Network). Structured query language (SQL) was used to create queries and views of the information from the database management system. The conceptual schema contained some redundancies regarding the storage of milk recording information in order to support a broad range of user preferences. The primary features of the logical schema included creating a cow entity which would allow users to access information about herd members irrespective of organisational source and to create entities that maintained the autonomy of the underlying source organisations that provide the information. Using SQL, two views of the data were created to show the capabilities of the physical schema to represent the DBIS. Data from the DHI test and lactation tables were retrieved according to current user expectations and data originating from several organisations were retrieved together. The implementation of the physical schema as a data base management system facilitates the retrieval and storage of information transparently and independently. The use of a unified data model, to support the dairy breeding information system, could promote the useful employment of the Internet for accessing legacy information from current source organisations and for facilitating new applications employing information from many organisation simultaneously by on-farm users.

Introduction

There is a high volume of information from many sources in dairy information systems available to support breeding decisions on-farm (Chapters 3 and 4, Lacroix et al., 1994). Decision support systems (DSS) are becoming a useful means of assisting breeding decision making on-farm. The information from the dairy industry has various forms and frequencies of delivery from the various organisations maintaining data. Prior treatment of the information is required before it can support on-farm decision-making; it must be analysed, integrated, aggregated and structured to be well used. A unified data model would provide a way for on-farm decision makers to analyse and manipulate the information in the dairy breeding information system (DBIS).

A unified data model of the information available should improve clarity and utility, and link information products to decision-making processes being addressed across the information system. Tasks that are tedious or not well suited for human agents in information management may be alleviated. It should also be the underlying basis for automation and networking of information flows, required for Internet suitability and distributed access to information. This could lead to value-added information products being available to users, through the World Wide Web as required (Trustman and Meshako, 1999).

A process for analysing and creating a unifying data model for information systems in the dairy industry needs to be described. A standard approach, known as information engineering (IE), is useful for data modelling and database system development (Harrington, 1998; and McFadden and Hoffer, 1994). The IE process was used by Lescourret et al. (1993, 1994) for database systems in the dairy industry, including the information description, conceptual, logical and physical model development for on-farm management activities. They provided frameworks for utilising information about animal morbidity, production, reproduction and feeding (the latter in greater detail), but did not deal with breeding. Other approaches or variations on IE exist. Ballou et al., (1998) and Ballou and Pazer (1985) described structured methods to model information systems elaborately, including means of evaluating the information and following its flow through the system conceptually. Their approaches focussed on information quality of conceptual models rather than an overall plan for the development of a first generation prototype. Levy et al. (1998) and Kemp et al. (1996) described means of integrating database systems to facilitate querying of heterogenous systems, however a detailed data modelling methodology was not created. Frederiks et al. (1997) in their method focussed on conceptual/logical data modelling (which is part of IE), proposing a highly theoretic approach. Similarly, Lee's (1996) semantic data modelling approach suggests improvements to the logical data modelling as a variation of the IE process. Thus, IE process appears to be an appropriate, common approach for describing and implementing a data model.

The IE process involves several phases which have been described in the literature (Graves et al., 1996; Ganguly and Noordewier, 1996; Harrington, 1998; Lescourret et al., 1993; Lescourret et al., 1994; McFadden and Hoffer, 1994; and Sargent et al., 1996). Three important phases of data model development emerge based on the means of representing them: conceptual, logical and physical (see Figure 5.1). The conceptual model or conceptual schema (Harrington, 1998) analyses and describes the information system domain in terms useful to practitioners in the domain and can, be represented with data flow diagrams (Harrington, 1998; McFadden and Hoffer, 1994; Lescourret et al., 1993). The logical model or schema is a design of the conceptual model translated into substantive database formalisms (Graves et al., 1996). It can be represented by entity relationship diagrams (Harrington, 1998; McFadden and Hoffer, 1994; and Lescourret et al., 1993). The physical model or schema is implemented when data populate a database system using the logical schema. The physical schema can be represented using a database management system (DBMS) and information may be retrieved from the DBMS using structured query language (SQL) (Harrington, 1998; McFadden and Hoffer, 1994; Ganguly and Noordewier, 1996).

A DBIS in which information is created, resides and flows among many participants has been described conceptually (Chapter 3), however a logical and physical schema should be described to validate the DBIS as a data model. Using an IE process, a unified data model can be developed for dairy cattle breeding information in order to access and use information across the whole domain to support automated decision making.

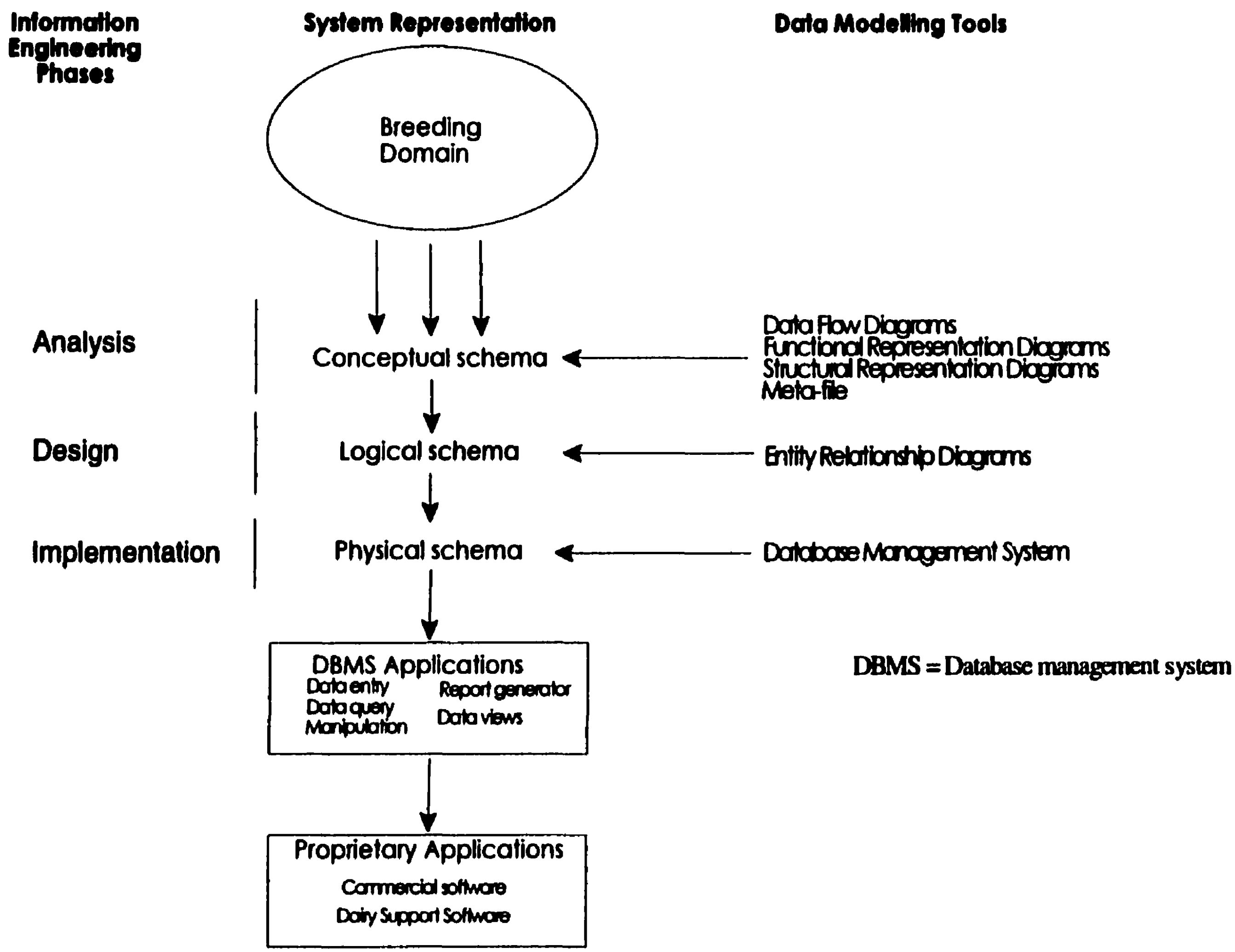


Figure 5.1 Information Engineering process

Our objective was to develop a data model for industry-wide information for dairy breeding decision making. A metafile, representing the information in a dairy breeding industry is described and builds upon previous conceptual modelling research by Archer et al. (Chapters 3 and 4). A logical model, using a relational approach, is created to represent the information. The implementation of the conceptual model will create the underlying tables in a physical data model that is optimised from the perspective of on-farm decision-makers focussing on creating views from information available to them related to the animals in their herds. Views of the data will be created to support on-farm dairy-breeding decisions to illustrate the effectiveness of the data model.

Materials and Methods

Three phases in the information engineering (IE) process are described: conceptual, logical, and physical modelling representing the conceptual schema, logical schema and physical schema respectively. The DBIS was described at the conceptual level with what can be called structural and functional representation diagrams (SRD and FRD respectively) and DFD by Archer et al.(Chapters 3 and 4). They described the general types of information available and identified which participants were providing and receiving them. Here the conceptual model is explored more completely using a metafile of information sources through the Canadian Dairy Data Dictionary (1989). The logical schema is described using an entity-relationship diagram and the physical schema using a DBMS. The development strategy involved moving among iterations of stages even if the prior stage was not complete; e.g., insights gained from the physical stage were used to iterate the logical stage.

The metafile lists information in the dairy breeding industry which is potentially available to producers through the Internet. Flat-file formats of information products available from web-sites of DBIS participants were used to determine the actual fields and the source of the information. Where similar information products were available from more than one source, information from all sources was included in the metafile according to the file formats available. The Canadian Dairy Data Dictionary was used to validate the fields and actual sources of the information included in the meta-file. The metafile included information from Canadian Dairy Network (CDN) -- the Canadian genetic evaluation centre; Holstein Canada (HC) breed association; Québec dairy herd improvement Programme d'analyse des troupeaux laitiers du Québec (or PATLQ); and artificial insemination centres (Gencor and Centre d'insémination artificielle du Québec).

The relational model, introduced by Codd in 1970 (Codd,1990; Harrington, 1998; Lescourret et al., 1993; Lescourret, 1994; and McFadden and Hoffer, 1994), was used to formalise the representation of the data model as a logical schema. The relational model groups entities as sets according common characteristics they share, similarly to the mathematical concept of relations. It is robust enough to express the complexity of the conceptual schema without compromising meaning in the domain. The phases in the IE process are followed irrespective of a particular logical schema.

Entities are things that exist in the real world about which data is stored and data stored about these entities are known as attributes. Once the entities are known, they can be linked together according to relationships described in the domain. Relationships between entities are often established through their primary indices or other indexed attributes of those entities called foreign keys (Harrington, 1998; McFadden and Hoffer, 1994). For example, in the dairy cattle breeding domain entities would be a cow or a lactation or a classification event. Relevant attributes for a lactation would be the lactation date and for a cow, its date of birth, sire or dam registration number. The cow entity has a relationship to a lactation entity since cows can have multiple lactations. The link is established through the cow's identification number appearing as an attribute of both entities. The logical model can be represented using the entity-relationship diagram.

Cardinality is the determination of relationships among entities in the data model. The cardinality among entities can be one-to-one (1,1), one-to-many (1,N), or many-to-many (N,N). Most of the relationships in a data model will be 1,N (Harrington, 1998, McFadden and Hoffer, 1994). Many 1,1 relationships are a single entity that has been made into 2 or they may be truly 1,N. One-to-one relationships can be created within a data model for example for reasons specific to the domain or to isolate part of an entity for security. Many-to-many relationships may exist in a data model, however since the relational model cannot represent them (Codd, 1990; and Harrington, 1998), N,N relationships are represented as two related 1,N relationships.

The relationships in the logical model must be normalised to remove data redundancy and to avoid update anomalies common in database design (Harrington, 1998; and Codd, 1990). If separate entity-relationship diagrams exist, modelling sub-portions of the data model, these should be merged, maximising integration across the domain. While there are five normal forms, we normalised to the third normal from which typically is adequate for most data model representations (Harrington, 1998). This entails removing repeating groups within entities, ensuring that non-key attributes are functionally dependent upon the primary key, and removing transitive dependencies that occur among entities.

The physical data model, or the physical schema, was developed using Microsoft Access 1997 © (a data base management system tool) to create tables from the entities, and links among tables according to the logical schema. The data to populate the physical schema originated from the McGill University research farm and came primarily through HC, CDN and PATLQ. Views (perspectives of looking at the data therein) were retrieved from the physical schema and were created using SQL.

Results

The results of the IE process (Figure 5.1) are given in the following section. The DBIS is modelled so that at the end of the process participants in the system can view, manipulate and retrieve the information contained within the system. In Figure 5.1 on the left the individual phases of the IE process are represented. The modelling tools used to contribute to each representational stage are shown on the right side and the stages are shown in the middle.

Conceptual Schema: The metafile developed for the DBIS is shown in Figure 5.2. The information is grouped according to the component sectors that produced them.. Information is maintained within each sector in individual files corresponding to specific information needs defined by the organisations within the sectors. For example, two aggregations represent the information from PATLQ: test day which represents information as a 'snapshot' of the herd at a given time, and lactation information which looks at the production of cows in the herd over an interval of time.

The information from CDN considers both sires and cows in the herd. One of the data sets concerns cow proofs (both production and conformation genetic evaluations) and 3 sets of data are concerned with national and international sire proofs. When creating the metafile, information products were duplicated if more than one source of information was offering similar products. While this is redundant from a design viewpoint there may be advantages to the industry member such as maintaining control of information. Advantages for certain users may be gained in the instance where they may only have access to one organisation in the system. Also, a user, needing information over varying historical periods, might be able to query more than one source of the information if the most immediate provider of that information does maintain the data in the form or time period required. Finally, redundant storage may facilitate applications where real-time access is required and traffic at a given provider is heavy.

Most information systems are not static; the Canadian system is under considerable change. The metafile does represent the information available to producers in the Canadian dairying system as it exists shortly before the time of reporting. The main issue is not just the specific information of one system or another, but the idea that based on an understanding of source data a description of available information can be developed to ultimately serve client needs conceptually.

Logical Schema: The perspective of interest for this research was on-farm decision-making. Information products shown in the metafile should be available from the data model in similar forms given. It is also assumed that access to the information is to be provided on an individual farm basis. Once a cow leaves the herd it was considered to be culled and cannot return.

The part of the conceptual schema that deals with cow information is represented as a logical schema in Figure 5.3. The entities and relationships are represented as tables (the rectangles) and links (lines between them) respectively. The parallel vertical bars crossing the relationship line symbolises that there is one and only one instance of a cow related to an instance of the other entity type.

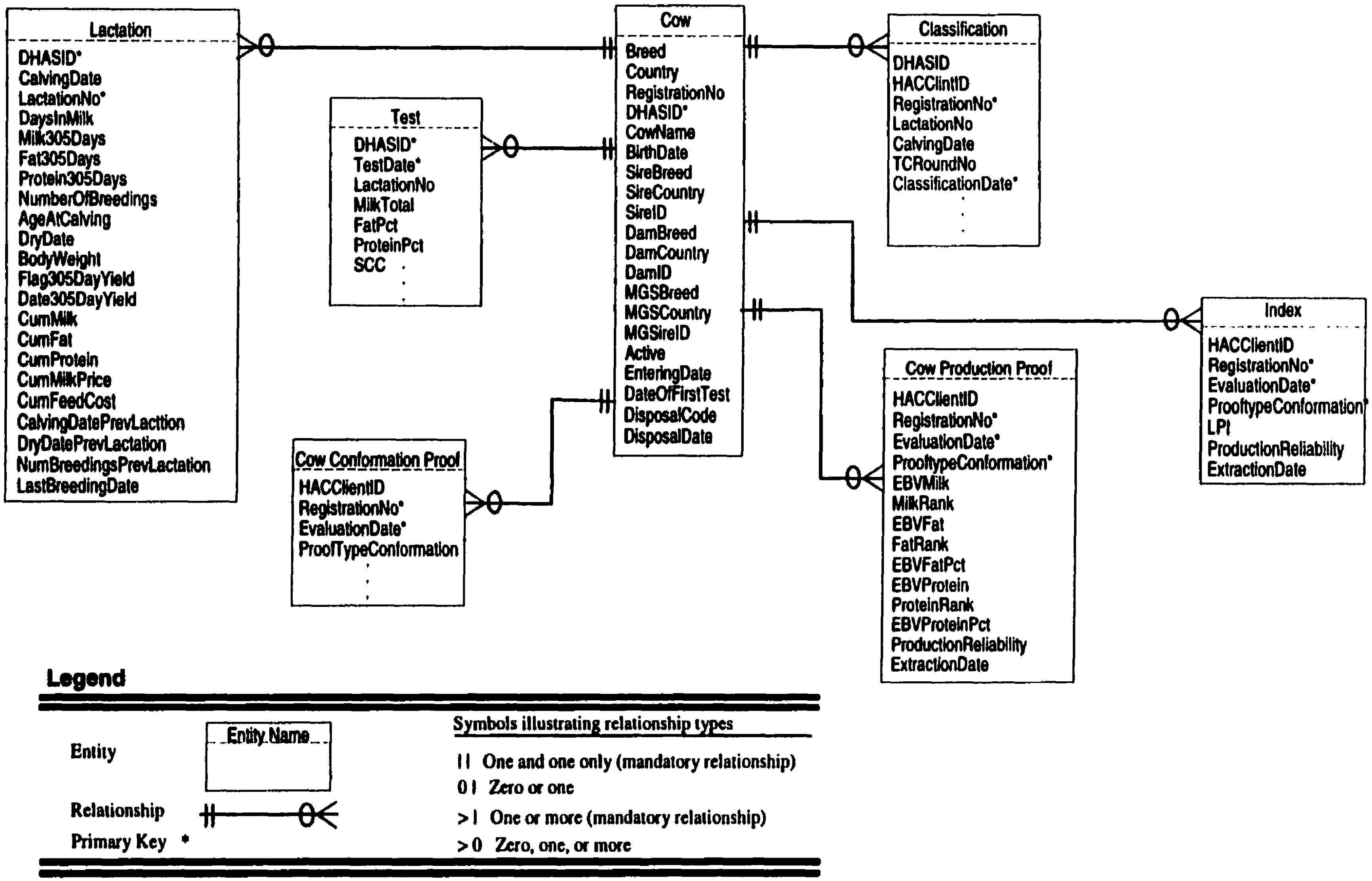


Figure 5.3 An Entity-Relationship diagram of DBIS cow information

The '0<' at the other end symbolises that there can be zero, one or many instances of that entity related to a cow. Not all of the attributes of the entities are shown for every table to avoid cluttering the diagram, as indicated by the dots at the end of those tables (see Appendix 1 for a complete list of tables and attributes on the cow side).

The basic entity on which all other available information is modelled is the cow. The attributes included will not change during the life of the cow on the farm for example its birth date, sire and dam identification numbers and the date it entered the herd. The primary key chosen was the PATLQ identification number, called DHASID, which is a unique non-repeating number within a farm. RegistrationNo uniquely identifies animals in Canada and could also have been used. Both identification numbers are included in the cow entity as they link instances of "cow to production" and "cow to evaluation" entities in the model.

Important events in the life of the cow that generate data with respect to breed improvement (Chapter 3) are included as separate entities in this model. From the metafile, milking records were available from test-day and lactation information, classification information from the breed associations, and estimated breeding values (EBVs) were available to create entities. The entities created reflect the representation of information according to logical reasonable groupings. This reinforces data integrity (by normalisation), balances the independence of source organisations and allows data integration at the same time. For example, one or more type attributes of conformation, important for some users, could be retrieved along with genetic evaluations through cow entity relationships. Fewer entities could have been created, such as a single conformation entity rather than the separate 'classification' and 'cow conformation proof' entities. One less entity would have been created and storage date, type classification round number, lactation and registration numbers of cows classified would not need to be repeated. This would, however, compromise the independence of HC and CDN who create and store the information. More than one entity was created for EBV information since both CDN and HC both store EBVs. A sound entity naming convention needed to be adopted to avoid problems with information retrieval for attributes with the same or similar names from two or more sources.

Production information is maintained both as lactation and test-day entities which might be considered excess storage or presumptive of user preference. It is reasonable to expect both views of the data to be important to users and keeping both entities balances efficient storage and efficient processing for the convenience of end users. The entity index, likewise, is maintained in the data model even though its attributes could be created from information available from other entities in the data model.

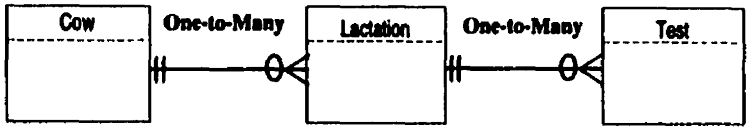
Maintaining both entities allows some distinctive cardinality features among cow, lactation and test day entities to be considered. Figure 5.4 shows ways that this relationship can be expressed among cow, lactation and test entities. The entities are represented as boxes and relationships as lines among them. The '0 <' means that zero, one or many instances of test and lactation are possible in these relationships and the double vertical line at the other end that there can be only one instance from this entity type corresponding to them. The relationship could be expressed as cow:lactation (1:N) and lactation:test (1:N) (see Figure 5.4a). This ignores the direct cow:test (1:N) relationship creating a hierarchy in the relationships and does not consider the test:lactation (1:N) relationship. In this model the relationships are represented as cow:lactation (1:N) and cow:test (1:N) (Figure 5.4b). While this minimises the lactation:test (1:N) relationship no hierarchy is created from the point of view of the central cow entity. Furthermore, with no evidence of a problem from the perspective of an on-farm view; the farm perspective is represented relatively well.

The structure of the logical model allows the greatest autonomy to the organisations where the information is maintained. Genetic information, for example, is a shared responsibility among a number of sectors (Chapter 3 and 4); the genetic information is grouped as production and conformation entities with the indices, which are derived from both, and also as a separate entity (Figure 5.3). If genetic information were grouped into a single entity, then physical schema deployment would result in a single table. It would then be difficult for HC, say, to restrict conformation EBV information to their clients only. In a logical schema with separate entities, organisations can moderate access to entities in the physical schema they control.

Perhaps the greatest problem with this logical data-model was multiple dependencies existing within the cow entity. If a cow in the herd is rented or sold and then brought back into the herd there will be two instances of the cow with different DHASID but with all other information being the same except for EnteringDate and DisposalDate. To the data model they are different, but in reality they are the same animal. This creates a problem in the DBMS for other uniquely indexed attributes within the cow entity, namely RegistrationNo, which is used to establish relationships with genetic evaluation and classification entities. Removing all but one of the conflicting instances overcame this problem, however this is not feasible for actual implementation. Usually the best approach is to create an auto-incrementing ID for the primary key for cow table not associated with any real data from the entity. But in this case the DHASID is supposed to accomplish this and the problem with RegistrationNo would not be solved. Another approach would be to create a complex primary key consisting of DHASID and RegistrationNo. The solution chosen did cause some information loss. The recently implemented Canadian Milk Recording Information Project has developed a standardized system of information recording on a national basis, and deals with this problem by having separate tables for animals, herds and herd members. A unique database code for each animal and herd members is generated by the DBMS which forms the primary key in the animal table and part of the primary key in the herd members. Among these tables RegistrationNo only appears in the animals table so the reference to animals in the herd members table does not violate uniqueness rules.

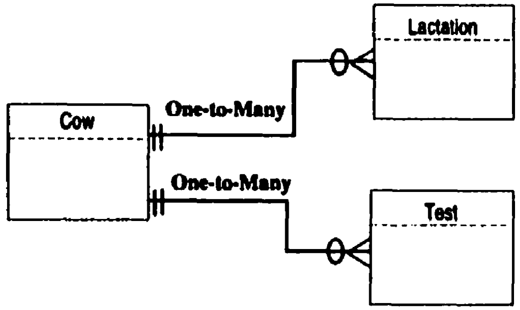
Information on male animals in the herd is included either as male relatives or as sires in pairings (see Figure 5.5). Sires or maternal grand-sires are the two male relatives included in the cow entity. Using herd-book records, pedigrees of cow instances could be pursued at greater depth, but only if it has relevance to on-farm breeding decisions. Service sires are included in the lactation and test day entity.

Figure 5.4a Cow:Lactation; Lactation:Test



One cow can have many lactations
One lactation can have many tests
One cow can have many tests

Figure 5.4b Cow:Lactation; Cow:Test



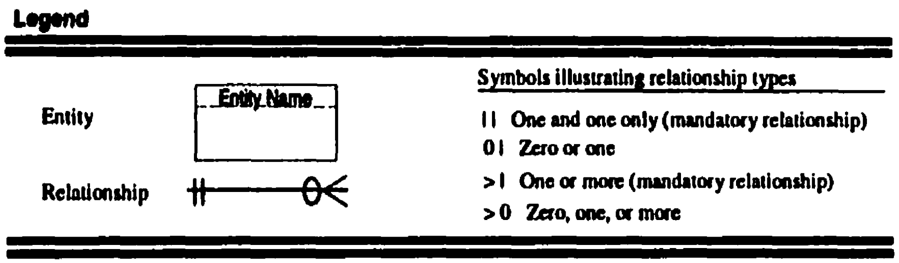


Figure 5.4 Entity-Relationship diagrams showing two ways of representing the relationships among cow, lactation and test-day entities.

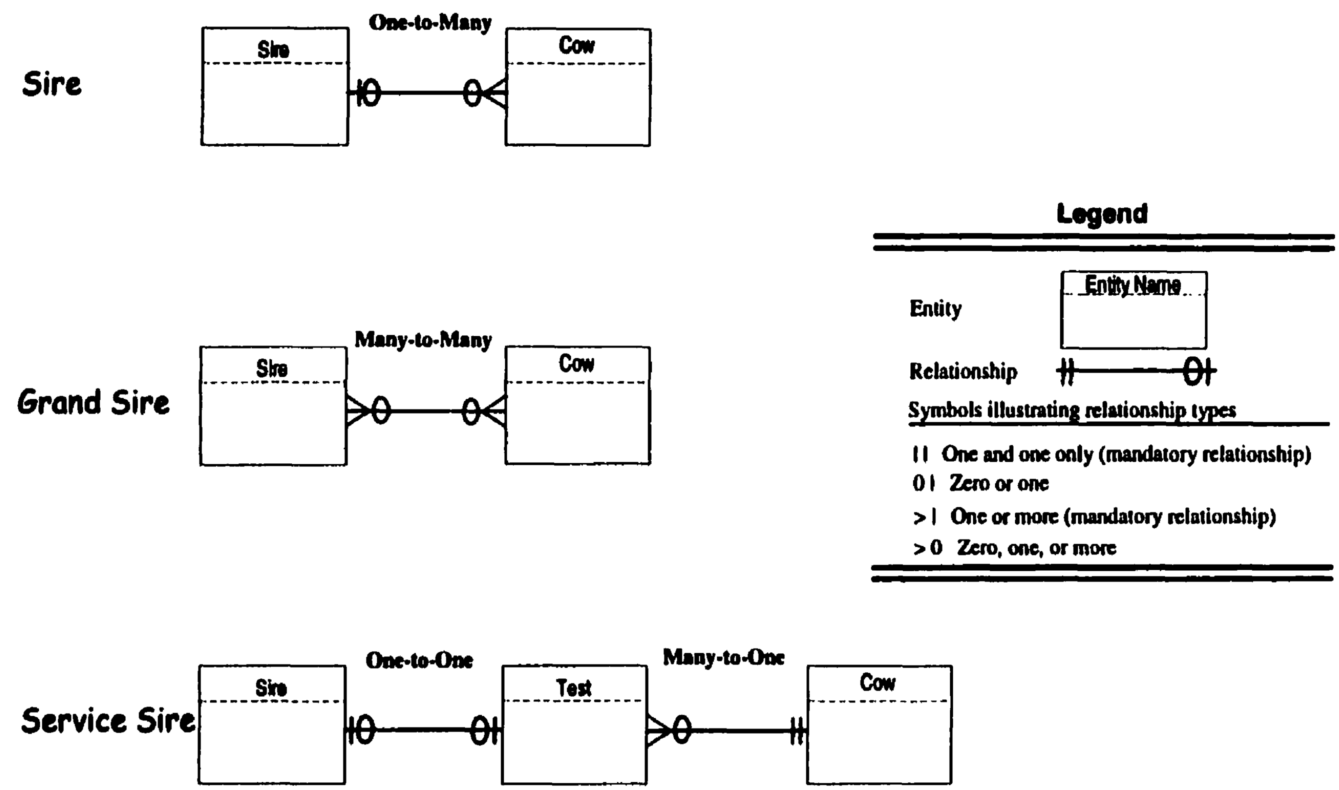


Figure 5.5 Entity-Relationship diagrams of the relationships between sire and cow entities in the data model

Physical Schema: Deploying the logical schema as a DBMS creates a physical schema (Figure 5.1). The tables in the physical schema were created according to entities in the entity-relationship diagram. Prefixes were added to them to identify the type of organisation of originating information. For example, the tables created for lactation information originating from PATLQ were named DHILactation and DHITestday. In the event that the same or similar information was being maintained by several participants the user could specify the source explicitly during selection.

From the physical schema instances of entities are manipulated through tools and applications (see Figure 5.1). Data retrieval applications, for example, are facilitated using SQL which can be used to retrieve information from within or across the entities represented in the data model. SQL is a high level language standard available with most commercial DBMS. The following complex SQL queries retrieve PATLQ SCC and related SCC or mastitis related information(see Figure 5.2) across organisations as Tables 5.1 and 5.2 respectively (the queries used to retrieve the information shown in the tables are shown in Appendix 2). Table 5.1, from a crosstab query, includes data from the test-day table (TestDay) with one data field from the lactation table (Lactation) and provides a herd summary that could be provided to a producer after a milk recording event. Organised by date of test, the information tells the producer at a glance the average days in milk, age at calving, number of cows in milk, value of milk and cost of feed respectively, and SCC of the herd over the past 14 months. SCC is further described on a parity basis. Blank spaces reveal an absence of information where no cows were represented: the DBMS considers these as null values which is not the same as zero.

Table 5.1 The Retrieval of SCC Milk production, calving and economic information from the milk recording tables (DHILactation and DHITestDay) using a SQL Crosstab query to summarise SCC data by parity.

Test Date	Ave DIM	Milk Value	Avg Feed Cost	Avg Calv Age	Cows in Milk	Ave SCC	SCC by Lactation Number ('000 / mL)					
							1	2	3	4	5	6
2/16/98	152.9	14.09	3.55	52.16	57	95.30	94.6	14.7	31.8	332.7	55.3	
3/16/98	165.6	15.69	3.69	51.18	62	96.10	63.1	97.3	53.0	235.8	97.3	
4/20/98	181.5	15.38	3.79	47.87	57	152.54	210.8	80.5	85.8	416.4	23.6	
5/19/98	199.4	14.49	3.53	48.02	56	112.34	206.8	87.7	46.3	154.0	22.4	
6/15/98	187.8	15.25	3.50	47.93	58	138.90	278.1	111.4	27.4	186.5	48.8	
7/20/98	195.8	10.56	3.21	50.01	59	149.58	70.6	177.5	115.1	325.5	70.2	110.0
8/17/98	183.9	9.69	3.65	48.58	59	223.24	57.5	237.4	186.5	614.2	65.8	40.0
9/21/98	175.1	11.59	3.27	48.76	53	203.04	57.2	240.8	324.3	269.4	79.2	59.0
10/27/98	169.3	13.70	3.42	47.71	57	650.47	139.8	1914.1	409.4	620.6	869.0	89.0
11/16/98	155.6	14.91	3.25	47.50	53	179.75	80.3	200.6	75.8	209.3	942.0	38.0
12/21/98	171.4	13.10	3.10	48.12	55	308.73	57.2	90.6	105.5	592.0	2322.0	73.5
1/18/99	179.2	11.34	3.08	47.14	54	260.37	183.0	60.4	155.6	175.5	2324.0	77.0
2/15/99	161.3	13.68	3.80	47.20	51	250.37	111.3	166.8	202.6	310.3	1301.0	92.5
3/15/99	154.7	14.32	3.95	46.95	51	321.59	246.5	72.0	398.2	264.7	1503.0	113.5

Table 5.2 A simple SQL retrieving milk recording, classification and EBV information together from the physical schema.

Reg No	DHASID	Parity	Expl Code	Class Date	Milk Total	SCC	DIM	Cum Milk	EBV Milk	F&L	Mam Sys	F Udder	R Udder	EBV F Udder	EBV R Udder
6367618	288	2	0	98/2/12	37.4	32.0	129	5198	520.0	80.0	81.5	81.5	81.5	-2	1
5756927	291	3	0	97/1/23	47.4	12.0	148	7597	1070.0	88.5	87.0	85.0	87.0	4	6
6367623	292	2	0	98/2/12	42.0	11.0	54	2156	114.0	74.5	67.0	74.5	63.0	1	-2
6534048	293	1	0	98/9/30	0.0	0.0	318	6058	623.0	78.0	74.5	78.0	71.0	2	-1
6520387	294	1	0	98/9/30	23.0	66.0	319	9506	1151.0	81.5	78.0	78.0	78.0	2	5
6534044	295	1	0	98/9/30	0.0	0.0	280	6552	367.0	83.0	74.5	78.0	74.5	4	2
6520384	296	1	0	98/9/30	25.6	19.0	272	7748	1305.0	80.0	80.0	81.5	80.0	5	5
6622323	297	1	0	98/9/30	21.2	14.0	257	5537	694.0	78.0	80.0	78.0	81.5	2	7
6622322	298	1	0	98/9/30	19.2	38.0	208	4322	-587.0	80.0	81.5	81.5	81.5	2	3
6534039	299	1	0	98/9/30	33.4	22.0	189	6517	499.0	83.0	81.5	81.5	81.5	7	5
6622325	300	1	0	99/5/12	28.6	47.0	181	4987	-8.0	78.0	81.5	83.0	80.0	3	-2
6495001	301	1	0	99/5/12	30.4	2063.0	94	3280	911.0	63.0	71.0	74.5	67.0	7	7
6739211	302	1	0	99/5/12	38.8	267.0	61	1974	947.0	80.0	58.0	63.0	58.0	4	9
6622329	303	1	0	99/5/12	35.8	26.0	81	2651	-42.0	67.0	67.0	71.0	58.0	5	-4
6695382	304	1	0	99/5/12	26.2	13.0	64	1587	94.0	85.0	81.5	83.0	80.0	1	-1
6670201	305	1	0	99/5/12	28.2	40.0	21	0	1030.0	81.5	81.5	81.5	81.5	9	6

Tables 5.1 and 5.2 are only examples of how SQL might be used to retrieve information. Herd information contained in the data model can be manipulated in various ways using several different attributes to add meaning and value. For example, SCC could have been constrained so that only a specific segment of the herd with high SCC, above a certain value, was shown. Or LactationNo could have been used to include only first parity animals. Several attributes in combination in a query may be used together. The information retrieved, and how it is displayed, are dependent on the user's ability to manipulate SQL. If well designed so that the logical schema accurately represents the conceptual schema, with enough expressive properties in the DBMS software, the physical schema will not be the limiting component. The data model creates the opportunity for expression but the user's needs must be clearly conveyed in the requests in order to produce it. A user's interaction with the data may be limited if they do not have a good knowledge of SQL. Methods could be devised to facilitate interaction with the data model.

SQL can retrieve information from several entities at one time and thus many organisations. Table 5.2 retrieved information originating from PATLQ, HC and CDN. In the SQL "FROM" statement, JOIN is used to establish relationships among entities to enable retrieval of information from many tables together by the attributes named (see Appendix 2). Only information related to the query will be retrieved from the database. Thus real time retrieval of information is improved and simplified for the user who can access the information they would require from underlying sources.

Genetic evaluations for milk production and, for example, fore and rear udder traits originate with CDN, however breed associations also maintain this information. In this case, if HC was the source of the EBVs rather than CDN, then BAProduction and BAConformation tables respectively would be included in the query. It is reasonable for the physical schema to make the distinction about sources if similar information is to be maintained. If the information was available from more than one source, priority of search for the information must be decided, perhaps by including creation and modification dates and allowing the system user to select their preferences. These issues would be sorted out through the domain knowledge available to create applications.

Discussion

The development approach used in this research allows for the separation of knowledge required for describing the DBIS (implicit in the DBMS – the back end) from deploying applications for its use (information retrieval and manipulation – the front-end). For example, domain experts and database developers can focus on back-end development issues of content, format and exchange (e.g. validity, completeness, transferability, unity, etc) for storage among distributed entities. Processing issues (e.g. integrity, aggregation, completeness, etc), which are front-end concerns, can be dealt with when developing retrieval methodology or dependent applications, or when performing decision making activities (Figure 5.1). An understanding and description of the DBIS (through the conceptual and logical models) would be in place when participants from various organisations attempt the latter. This division should improve productivity and work flow of on-farm user's decision making process (Trustman and Meshako, 1999) by facilitating the efficient incorporation of herd output information residing off-farm (Chapter 3). The focus can be on the quality of the decisions made and the value added therefrom; how to access and format the data becomes trivial. This gain in efficiency in the information system processing is partly due to the framework of information access provided by the database system software itself (i.e. tools provided with the system) but this gain cannot be separated from optimum design.

The physical schema directly facilitates the capture, storage and distribution functions and indirectly facilitates breeding processing functions in the domain (Chapter 3). It is a platform from which applications aggregate the information in the system (Figure 5.1), increasing the intelligence density of and adding value to the information provided to users for decision making. Applications can include information (from many source providers) across the whole breeding domain. Organisational concerns such as proprietary interests can be accommodated since entities can maintain information source. Security of information is of course an issue since there are many interests to be served in an information system where information is controlled and belongs to on-farm clients and several off-farm organisations, but robust logical schema, proper network architecture (if distributed), and the use of appropriate security would maintain proprietary control. New joint and separate applications for on-farm decision makers could be created.

A single logical schema is described, however, since the information described resides among several organisations and users geographically distant from each other, a distributed physical implementation would be necessary in the real world. Knowledge and information involved would require planned cooperative employment with the potential benefit of increased client bases for organisations and greater transparency of access to information products for clients. A control mechanism would be required to interpret issues including integration over a network, and priority of information (depending on decision space). Important advantages include local control, modular growth, lower communication costs and rapid response (McFadden and Hoffer, 1994). Challenges include processing overhead, slow response, and the potential difficulties of integrating legacy data systems including those based on data models different from the relational approach (McFadden and Hoffer, 1994; and Jacobs 1985).

The physical schema has distribution independence (Codd, 1990, McFadden and Hoffer, 1994): the logical schema can be implemented over any computing architecture and environment. This is suitable for the DBIS which, as an industry-wide information system, has many distributed users. On-farm clients of several organisations in the DBIS would be able to capture and store production information from many sources at one time using an integrated format. The unified data model supports collaboration among participants providing information to the producer. The underlying logical schema ensures integrated flow of production information among on-farm client and among organisations over a network, such as the Internet (Chapters 3 and 4) through a single channel.

Independence is manifested in another sense in that the user is not obliged to know how the information is stored to be able to access it, even across a network. One might think of the physical schema as existing at a different layer of abstraction than the applications that depend on it. Applications for on-farm breeding decision making can be built based on the model without needing to identify information sources. The user must only focus on what information they would like to access and not on where it is coming from.

Using Internet protocols, the physical schema could be developed as a distributed information source using the Internet infrastructure. A web-site supporting CGI, for example, would provide World Wide Web (WWW) access. At the back end the individual portions of the database may reside with and be controlled by the organisations supplying the information; each acting as a server for all data in the domain. The details of distributed management (i.e. database control and concurrency) would depend on the nature of the relationships among the organisations involved, and between the organisations and their clients. Countries in which the organisations are highly integrated could focus less on autonomy and more on homogeneity in its distributed management, while in countries with a high degree of autonomy and heterogeneity, transaction management may require asynchronous retrieval, greater server-side control of database changes and rigorous schema integration and global query management. A basis for agreement and common interpretations of much of the data in dairy breeding already exists in Canada because of cooperation among dairy organisations. This basis could be built upon to develop, for example, local libraries for definitions of data, access and security routines, and other common tools. With distributed network services, existing applications and commercial software used by participants could be accommodated, allowing integration even on top of a distributed database (Bernstein, 1996; Hart, 1997; Kador, 1996; Rymer, 1996; and Wiederhold and Genesereth, 1997).

The research described here intersects with many of the aims of the Canadian Milk Recording Information Project goals. The first phase objectives include the creation of data models to allow the integration of current regional milk recording programs and to investigate and implement appropriate DBMS; the second phase aims to develop and test a modular system and recommend an efficient method to exchange data (PATLQ, 1996). A unified DBIS data model accommodates the national integration of milk recording information but sees this as only a part of a greater information integration scheme. This is possible through cooperation among all sectors in the DBIS. The implementation of such a physical schema would benefit all DBIS members and the use of the Internet should be a cost effective, secure means of exchanging shared information.

Knowledge and expertise from each domain in the industry could be applied in different ways. Organisations within a single domain could individually support required functions already in progress (internally as operations, and externally as client information needs). Collaborative use of knowledge from several organisations across domains could be applied externally to support on-farm decisions requiring information from several parties. This is currently a grey area which can only be resolved in conjunction with participants in the information system who provide information products. Knowledge about the domain has been employed to create the DBMS and further knowledge should be employed to transform and provide the data made available through the system.

Conclusions

This research describes a process for the development of a unified data model for an industry-wide information system to support dairy cattle breeding decisions for on-farm managers. The results could provide production information to users through a DBMS which broadly supports un-predetermined user data choices for decision-making over a realistic domain. It may do this as a hidden layer supporting existing dairy breeding software applications within organisations or by fostering the creation of new applications among organisations in the breeding domain. In so doing it can act as a platform to provide production information in a networked information system invisibly without the technology becoming a barrier.

Using the WWW, the unified data model could disseminate information products of DBIS functions such as milk recording or genetic evaluations on-line. The information could be retrieved about specific decision-making events transparently from many organisations for the on-farm user. Further research should investigate the creation of a front-end to the physical schema to allow users to access the data without prior knowledge of SQL. Employing knowledge to the framework to create and link to applications that, for example, interpret queries to retrieve information would improve the data model as a decision-making tool for the domain.

Acknowledgements

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, les Fonds de l'Autoroute de l'Information du Gouvernement du Québec, and the Cattle Breeding Research Council of Canada.

References

Ballou, D, and H Pazer. 1985. Modelling Data and Process Quality in Multi-Input, Multi-Output Information Systems. Management Science 31 (2): 150-162.

Ballou, D, R Wang, H Pazer and G Kumar Tayi. 1998. Modelling Information Manufacturing Systems to Determine Information Product Quality. Management Science 44 (4): 462-484.

Bernstein, P A. 1996. Middleware: A Model for Distributed System Services. Communications of the ACM. 39(2): 87-98.

Canadian Dairy Data Dictionary, The. 1989. Canadian Dairy Information Exchange, Brantford Ontario, Canada.

Codd, E F. The Relational Model for Database Management. Addison-Wesley Pub. Co. Inc., Reading, MA. 1990.

Frederiks, P J M, A H M ter Hofstede and E Lippe. 1997. Information and Software Techhnology. A Unifying Framework for Conceptual Data Modelling Concepts. Information and Software Technology 39: 15-25.

Ganguly, S and M Noordewier. 1996. PROXIMAL: A Database System for the Efficient Retrieval of Genetic Information. Computed. Biol. Med. 26(3): 199-207.

Graves, M., E. R. Bergeman, and C. B. Lawrence. 1996. A Graph Conceptual Model for Developing Human Genome Center Databases. Computed Biol Med. 26(3): 183-197.

Hart, P. E., and J. Graham, R. 1997. Query-Free Information Retrieval. IEEE Expert 12(5): 32-37.

Harrington, J L. Relational Database Design Clearly Explained. AP Professional, Chestnut Hill, MA. 1998.

Kador, J. 1996. The Ultimate Middleware. Byte 1996(April): 79-83.

Kemp, G J L, J Dupont, and P M D Gray. 1996. Using the Functional Data Model to Integrate Distributed Biological Data Sources, Pages176-185 *in* Proceedings Eighth International Conference on Scientific and Statistical Database Management. June 18-20, 1996. IEEE Computer Society Press, Stockholm.

Jacobs, B E. Applied Database Logic Volume1: Fundamental Database Issues. Prentice-Hall Inc., Englewood Cliffs, NJ. 1985.

Lacroix, R, K M Wade, R Kok and F Hayes. 1994. Functional Characteristics of a Decision Support System for Dairy Cattle Management. International Summer Meetings of the American Society of Agricultural Engineers, Kansas City, MO. Paper No. 94-3027 ASAE, 2950 Niles Rd., St. Joseph, MI 49085-9659, USA.

Lee, Dong-Tsan. 1996. Semantic Data Modelling Using Linear Logic. Information Processing Letters 60: 19-27.

Lescourret, F, M Genest, J Barnouin, Michelle Chassagne, and B Faye. 1993. Data

Modeling for Database Design in Production and Health Monitoring Systems for Dairy Herds. J Dairy Sci 76: 1053-1062.

Lescourret, F, P Dunaud, J Barnouin, Michelle Chassagne, and B Faye. 1994. Computerized representation of feeding regimes in monitored dairy herds. Computers and Electronics in Agriculture 10 (1994) 151-166.

Levy, A, C Knoblock, S Minton and W Cohen. 1998. The Information Manifold Approach to Data Integration. IEEE Intelligent Systems and Their Applications 12(5): 12-16.

McFadden, Fred R., JA Hoffer. 1994. Modern Database Management. 4th ed. Benjamin / Cummings Publishing Company, Inc., Redwood City, CA.

Microsoft. 1997. Visual Basic Guide to Data Access Objects, Crystal Reports for Visual Basic. Microsoft Corporation, Document No. DD93015-1296.

PATLQ. 1996. "Working Together." PATLQ News bulletin, Number 35, May, 1996.

Rymer, J. R. 1996. The Muddle in the Middle. Byte 1996(April): 67-70.

Sargent, R, D Fuhrman, T Critchlow, T Di Sera, R Mecklenburg, G Lindstrom and P Cartwright. 1996. The Design and Implementation of a Database for Human Genome Research. Pages 220-225 *in* Proceedings Eighth International Conference on Scientific and Statistical Database Management, June 18-20, 1996. IEEE Computer Society Press, Stockholm.

Trustman, J, and S Meshako. 1999. You Say You Want an Evolution. Intelligent Enterprise 2(16): 18-20.

Wiederhold, G. and M. Geneserth. 1997. The Conceptual Basis for Mediation Services. IEEE Expert 12(5): 38-47.

Appendix 1

List of entities in the dairy breeding data model logical model with all attributes

The entities are set in bold on the right margin. A complete list of the individual attributes are included in the brackets. The primary key for each entity is underlined.

Cow (<u>DHASID</u>, Breed, Purity, Country, RegistrationNo, CowName, BirthDate, SireBreed, SireCountry, SireId, DamBreed, DamCountry, DamId, MGSBreed, MGSCountry, MGSireId, Active, EnteringDate, DateOfFirstTest, DisposalCode, DisposalDate)

CowConformationProof (HACClientID, <u>RegistrationNo</u>, <u>EvaluationDate</u>, <u>ProofTypeConformation</u>, EBVConformation, ConformationRank, TypeReliability, ProofTypeScorecard, EBVFrameCapacity, EBVRump, EBVFeetLegs, EBVForeUdder, EBVRearUdder, EBVMamSys, EBVDairyCharacter, EBVStature, EBVSize, ProofTypeSecondary, EBVFrontEnd, EBVChestWidth, EBVBodyDepth, EBVLoinStrength, EBVPinSettingDesirability, EBVRumpAngleAlphaCode, EBVPinWidth, EBVFootAngle, EBVBoneQuality, EBVRearLegsDesirability, EBVRearLegSideViewAlphaCode, EBVUdderDepth, EBVUdderTexture, EBVMedianSuspensory, EBVForeAttachment, EBVFrontTeatPlacement, EBVFrontTeatLength, EBVRearAttachmentHeight, EBVRearAttachmentWidth, EBVRearTeatPlacement, EBVDairyForm, EBVRumpAngle, EBVRearLegSideView, EBVUdderDepthAlphaCode, EBVForeTeatLengthAlphaCode)

Cow ProductionProof (HACClientID, <u>RegistrationNo</u>, <u>EvaluationDate</u>, <u>ProoftypeConformation</u>, EBVMilk, MilkRank, EBVFat, FatRank, EBVFatPct, EBVProtein, ProteinRank, EBVProteinPct, ProductionReliability, ExtractionDate)

Index (HACClientID, <u>RegistrationNo</u>, <u>EvaluationDate</u>, <u>ProofType</u>, LPI, PercentileRankLPI)

Classification (DHASID, HACClientID, <u>RegistrationNo</u>, LactationNo, CalvingDate, TCRoundNo, <u>ClassificationDate</u>, FarmName, FinalScore, FrameCapacity, Stature, HeightatFrontEnd, Size, ChestWidth, BodyDepth, LoinStrength, Rump, PinSetting, PinWidth, FeetLegs, FootAngle, BoneQuality, SetofRearLegs, MamSys, UdderDepth, UdderTexture, MedianSuspensory, ForeUdder, ForeAttachment, ForeTeatPlacement, ForeTeatLength, RearUdder, RearAttachmentHeight, RearAttachmentWidth, RearTeatPlacement, DairyCharacter, DairyForm, RumpAngle, RearLegSideView, FinalClass, RumpAngle-alphacode, RearLegSideView-alphacode, UdderDepth-alphacode, ForeTeatLength-alphacode, WryFace, UndesirableHead, WeakCrops, WeakBack, LacksStyle, NotWellSprung, LowLoin, NarrowHeart, Frail, AdvancedAnus, RecessedTailhead, HighTailHead, AdvancedTailhead, WryTail, ThurlsTooFarBack, WeakPasterns, Crampy, ToesOutRear, CoarseHocks, OpenToed, ShallowHeel, UndesirableStance, LacksUdderShape, Tilt, ReverseTilt, Bulgy, HeavyFore, FrontUnbalanced, FrontShort, FrontNotPlumb, FrontWebbed, FrontBlind, RearUnbalanced, RearShort, RearNotPlumb, RearTeatsTooFarBack, RearWebbed, RearBlind, CloseRib)

Lactation (DHASID, CalvingDate, LactationNo, DaysInMilk, Milk305Days, Fat305Days, Protein305Days, NumberOfBreedings, AgeAtCalving, DryDate, BodyWeight, Flag305DayYield, Date305DayYield, CumMilk, CumFat, CumProtein, CumMilkPrice, CumFeedCost, CalvingDatePrevLactation, DryDatePrevLactation, NumBreedingsPrevLactation, LastBreedingDate)

TestDay (DHASID, TestDate, LactationNo, MilkTotal, FatPct, ProteinPct, SCC, CondCode1, CC1Date, CondCode2, CC2Date, ExplanationCode, DaysInMilk, CumMilk, CumFat, CumProtein, Milk305Days, Fat305Days, Protein305Days, Urea, BodyWeight, NumBreedingsPrevLactation, LastBreedingDate, LastServiceSireRegNum, CalvingDate, DryDate, DaysInGestation, ProjTDRealFlag, MilkPrice, FeedCost, CumMilkPrice, CumFeedCost, FlagOffLact)

Appendix 2

The SQL queries used to create table 2 and table 3.

The queries used to retrieve the information in tables 1 and 2 respectively follow. The general SQL clauses required are in capital letters set in bold type. Capital letters in each clause describe SQL code within each clause of the query. The specific fields and tables chosen are shown in the following format: 'Table.field'. In the TRANSFORM and SELECT Clauses, a function native to the DBMS ('Avg') is used to transform the information retrieved from the data model.

SQL for Table 1

TRANSFORM Avg(DHITestDay.SCC) AS [The Value]
SELECT DHITestDay.TestDate, Avg(DHITestDay.SCC) AS [Total Of SCC], Count(DHITestDay.DHASID) AS [Total of Cows], Avg(DHITestDay.SCC) AS [Total Of SCC] **FROM** DHILactation INNER JOIN DHITestDay ON DHILactation.DHASID = DHITestDay.DHASID
WHERE (((DHITestDay.CondCode1)<>3))
GROUP BY DHITestDay.TestDate
PIVOT DHITestDay.LactationNo

SQL for Table 2

SELECT Cows.RegistrationNo, Cows.DHASID, DHITestDay.LactationNo, DHITestDay.ExplanationCode, Classification.ClassificationDate, DHITestDay.MilkTotal, DHITestDay.SCC, DHITestDay.DaysInMilk, DHITestDay.CumMilk, EBVMilk, FeetLegs, MamSys, ForeUdder, RearUdder, EBVForeUdder, EBVRearUdder
FROM (((Cows INNER JOIN DHITestDay ON Cows.DHASID = DHITestDay.DHASID) LEFT JOIN Classification ON Cows.RegistrationNo = Classification.RegistrationNo) LEFT

JOIN BACowProductionProof ON Cows.RegistrationNo = BACowProductionProof.RegistrationNo) LEFT JOIN BACowConformationProof ON Cows.RegistrationNo = BACowConformationProof.RegistrationNo

WHERE DHITestDay.TestDate BETWEEN #03/15/99# AND #03/15/99#

ORDER BY DHITestDay.DHASID;

Chapter 6. DAIRIE: A Dairy Integration and Retrieval Information Expert for use in On-farm Breeding Decisions

A A Archer, R Lacroix and K M Wade

This chapter presents a prototype graphical user interface called DAIRIE which is a front-end for accessing information in an integrated fashion from the underlying DBIS. The integration of information, described and illustrated with a prototype unified data model in the previous chapter, is facilitated from the on-farm user's reference point by making the access to the system transparent through the use of domain knowledge to create dynamic information retrieval. The objective of supporting decision making was implemented by using domain knowledge to describe dimensions in which information may be aggregated and the conceptual basis for using expertise to further aggregate specific information retrieved by DAIRIE.

Abstract

The objective of this research was to describe the DAiry Information Retrieval and Integration Expert (DAIRIE) prototype: a graphical user interface software developed to interpret on-farm user requests for breeding information, and to employ domain knowledge to create views of the data from a database management system (DBMS). This was further illustrated by using information related to mastitis from several organisations. The Microsoft Visual Basic 6 © development environment was used for creating forms and modules to assemble SQL queries that interacted with the unified data model of the DBIS. It was implemented on an individual farm basis – using Microsoft ACCESS © – and 33 dairy-farm databases were created with data from the National Breed Association (Holstein Canada), the local Dairy Herd Improvement Association (Programme d'analyse des troupeaux laitiers du Québec) and the National Genetic Evaluation Centre (Canadian Dairy Network). Users could select specific information from the data model and aggregate it over various time and herd levels. Knowledge of the domain was acquired primarily through the literature, texts and extension material, and was employed to structure the interface according to user preferences and for treating and presenting the information once retrieved. The resulting prototype consisted of modules that generate clauses of the SQL queries and knowledge implementation and, ultimately, was represented by three interactive forms (Data Selection, Data Aggregation and Data Display). The DAIRIE prototype supports user needs by dynamically retrieving herd output information that is available throughout the DBIS. Retrieved information is maintained in volatile data-spaces where aspects of it can be explored. The DAIRIE prototype can improve on-farm decision making by integrating and filtering information from source organisations, and by allowing it to be manipulated and stored together for specific decision-making problems. It has the potential to play a key role in such areas as decision support through new or existing applications. While the general concepts of dynamic information retrieval and data driven knowledge application were demonstrated to work well in this first generation prototype, knowledge employment was not richly implemented at this stage; the basis is, however, there for rapid future developments.

Introduction

On-farm dairy breeding activities depend on many sources of information in the dairy industry which are available in industry-wide information systems such as the dairy breeding information system (DBIS) described by Archer et al. (Chapters 3 and 4). Human limitations in interacting with many sources of information and processing large volumes of information can hinder a user in terms of on-farm decision-making.

The Internet is an appropriate candidate for an industry-wide information system in agriculture since it approaches a true information highway (i.e., easy to use, ubiquitous, and the basis for many activities (Dertouzos, 1995)). It has relatively low (and falling) costs for use and connection, maturing technologies that will allow remote users access to it, and is already being used by many participants in the dairy industry. If implemented through the Internet in a distributed, networked architecture a database management system (DBMS) would be available to a wide number of users without some of the constraints of other media (e.g., "pull through" of only relevant personal information rather than "push through" of general information).

Participants who are sources of information in the DBIS could incorporate the front end into a web-site as a tool for direct information retrieval for their clients (Chapter 4). A standard means of accessing relational DBMS is structured query language (SQL) (Harrington, 1998, McFadden and Hoffer, 1994, Microsoft, 1997). Yet the use of SQL may create a barrier to using the tool since most on-farm users are not proficient SQL users. A front-end – software component which aids users in accessing an underlying structure, is required. Figure 6.1 shows how this front-end could sit on top of a system for retrieving information and provide a single pathway for users on-farm instead of having to deal with many pathways for retrieving the same information. A front end mediation agent can be developed to receive user requests, interpret and transform them into correct SQL code, submit it to the DBMS, and return data useful for their breeding decisions.

Knowledge-based decision-support systems (DSS) can be used to aid on-farm users. Executive information systems (EIS) are a type of DSS designed to support managers' needs for information (Lacroix et al., 1994; and Young and Hugh, 1995).

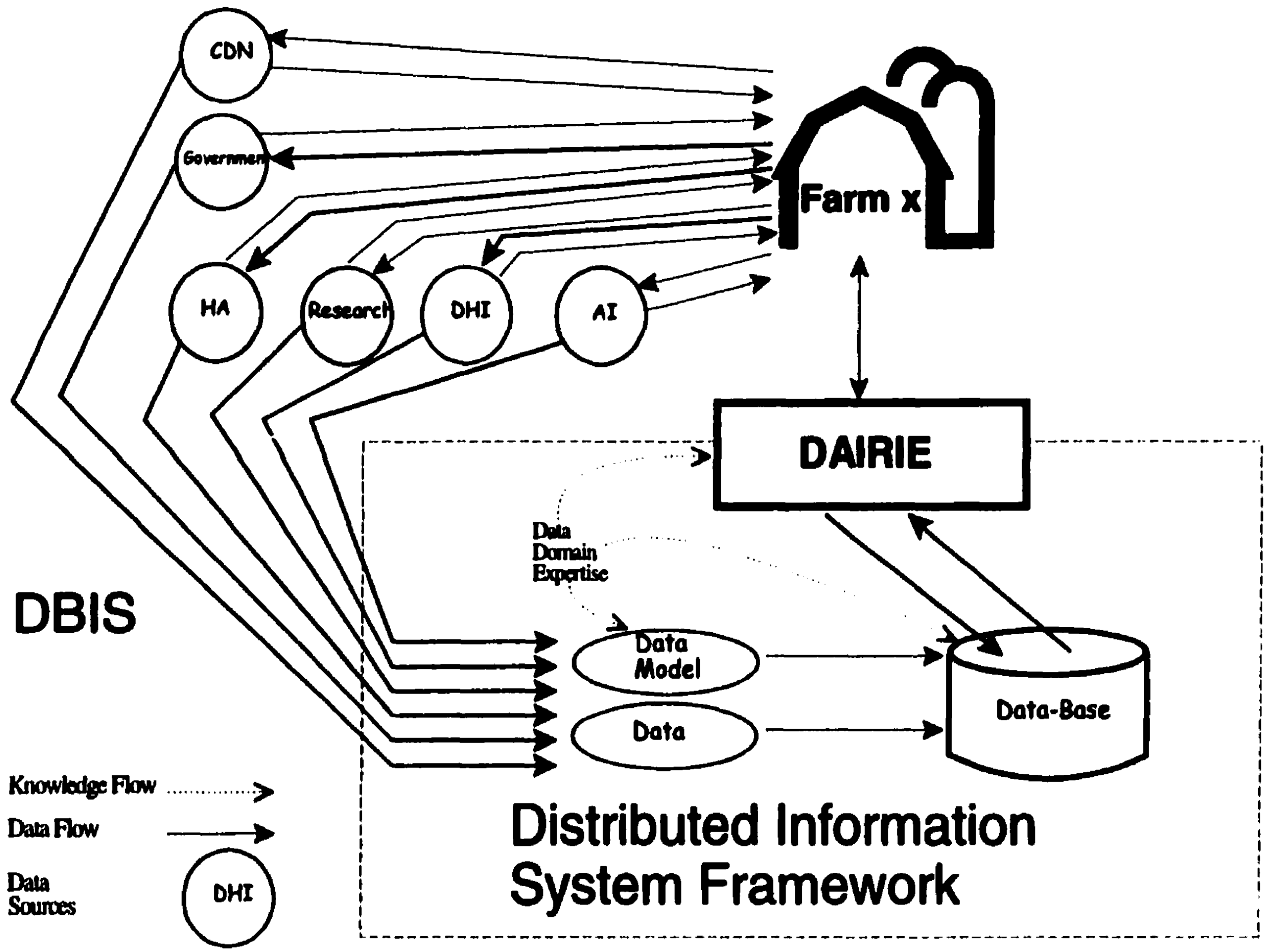


Figure 6.1 Component parts of a distributed retrieval framework

They are data driven systems, as described by Dhar and Stein (1997), in which knowledge is employed to support important, relevant selected data according to the user's preference in the domain: managing retrieved information, consistent with its utility for the user, and used to outline potential solutions to the problem of concern. They usually include a graphical user interface (GUI); integrated data access capabilities; *ad hoc* queries capacity; "drill down" capabilities; tools for data analysis and processing; and access to a variety of external data sources (Ivanovic and Budimac, 1999; and Nord and Nord, 1995). Executive information systems have been employed in a variety of management climates (Koh and Watson, 1998; Nord and Nord, 1995; and Young and Watson, 1995) and many commercial and research EISs have been developed (Dutta et al., 1997; and Varhol, 1995). Lacroix et al. (1994) propose the use of executive decision-making software to accumulate data and use knowledge to make complex computations as part of a framework for on-farm dairy decision-support. Ivanovic and Budimac (1999) proposed the use of an EIS in their framework for integrating information flows from databases and knowledge bases in agriculture. Their application is intended to support decision making for institutions. An EIS may be appropriate in supporting on-farm decision making. Implementing a data retrieval tool with many of the features of an EIS, supported by a DBMS distributed among many sources, as an integrated front-end would be a novel way to supporting on-farm decision making.

A unified data model has been described as a method of facilitating integrated, specific data retrieval from many data sources. It involves many participants in the DBIS through a relational database management system (DBMS) (Chapter 5) and can be used to support this work. In this paper Dairy Information Retrieval and Integration Expert (DAIRIE), a front end GUI mediation agent that interprets user information needs for on-farm breeding decisions, is described. It employs domain knowledge to create views of data from a DBMS which could be distributed and represents information from an industry-wide dairy information system. Using information from the mastitis domain as an example, DAIRIE shows how knowledge can be employed to manipulate queries provided for cow culling decisions.

Materials

The unified data model for the DBIS described by Archer et al. (Chapter 5) was implemented as a relational DBMS. The SQL version used was Microsoft Jet SQL which is consistent with the ANSI standard (Microsoft, 1997). The DAIRIE prototype was developed using Microsoft's Visual Basic version 6 programming environment (VB6). The data access technology of VB6 included methods for data access, data sources and data controls (Figure 6.2) to facilitate a structured interface for information manipulation, dynamic data binding, and data formatting objects.

The data used to populate the databases came from three sources in the dairy industry: Holstein Canada (breed association), Programme d'analyse des troupeaux laitiers du Québec (PATLQ) (dairy herd improvement) and Canadian Dairy Network (CDN) (genetic evaluation centre). The set contained information from 33 farms in the province of Quebec. Data included milk recording, classification and EBV records. A more complete explanation of the information contained in the data bases can be found in Archer et al. (Chapter 5). Thirty-three databases were created containing separate data for each farm.

Methods

The DAIRIE front-end is based upon SQL code which will now be briefly described. This project focussed on the selection capacity of SQL and comprised four main clauses: the SELECT clause sets the data fields retrieved; the FROM clause determines which tables from the database or databases should be included in the query; the WHERE clause poses restrictions on the fields selected; and GROUP BY determines the level of sorting of the data. All select queries must at least have a SELECT and a FROM clause. For example the simplest type select query would be composed as SELECT * FROM *table_n* (where *table_n* represents the name of the table or tables associated with the query and the asterisk represents all the data fields in that table).

A graphical method of representing the choices available to users is employed to develop the interface. Visual Basic 6 allows form objects and control objects placed on forms to be defined and created in most cases graphically (with minimal coding) to represent the data fields and dimensions of data retrieval requisite for phrasing SQL requests.

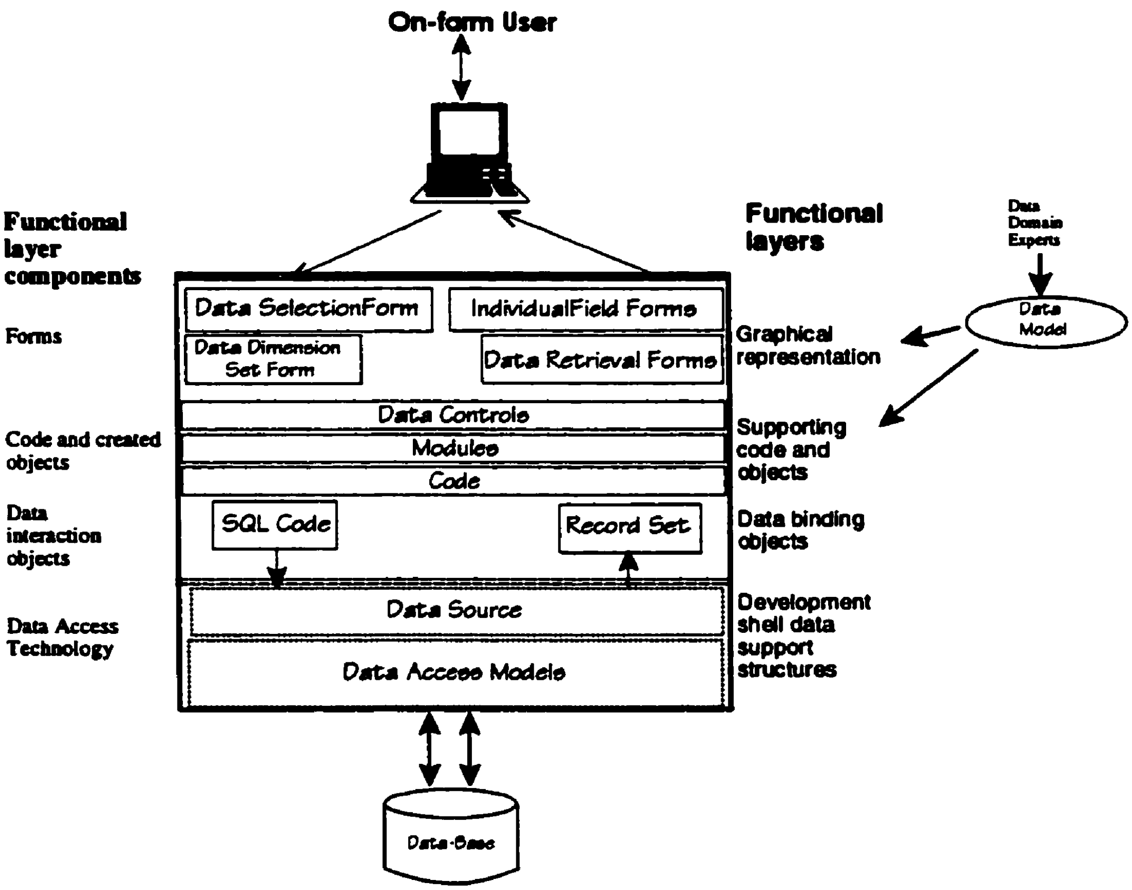


Figure 6.2 DAiry Information Retrieval and Integration Expert components and structure integrated withVB6 technologies to create, access and display information.

Code was written to support the objects and establish the connection to the underlying data base. Modules were created to translate user selections into the SQL code.

Using rapid prototyping principles, there were a number of criteria that were followed in the development of DAIRIE. Retrieval was on an individual farm basis.

Secondly DAIRIE allows retrieval of selected information over a variable time span. Creating a data-driven tool that provides information to support decision making must recognise that operational, tactical, and strategic decisions consist of different time horizons and information (Pietersma et al., 1999). With this in mind the user should be able to set the time interval for data retrieval and the level of detail should also be flexible according to user needs (i.e., the user ought to be able to view whole herd, groups or individual animal information).

Knowledge is employed in two ways: 1) structure the front-end for user preference in terms of important information and relevant dimensions and 2) treatment and presentation of the information once retrieved. Knowledge was acquired through an informal survey of the literature, texts, extension material and information provided by participants to on-farm users which were available in other media (e.g., paper mail-outs, electronic disks, etc.). Using this method a sub-selection of fields was identified which could be selected as an option button for selecting fields relating to mastitis. Knowledge about treatment and presentation of the information was captured by consulting with an expert as a trial user. Through these methods of knowledge acquisition the dimensions of data selection and their format in the software were validated.

Results

DAIRIE software: Figure 6.2 shows the structure of DAIRIE which supports data configuring and retrieval from an underlying DBMS. The layers are the graphical representation, supporting code and objects, data binding objects, and development shell data support layers. Forms, modules and code, data interaction objects, and data access technology, respectively, represent these layers. The forms are the visible elements of DAIRIE which the user must interact with and modules create the SQL code. The SQL code and record sets are the two types of data interaction objects described here to interact with the data access technology of the development environment to interface with the DBMS. The SQL code sends requests to the DBMS and record sets contain the retrieved information. Record sets are displayed on the created forms using objects called controls which are part of the code and created objects layer. The forms and modules are the developed components which rely explicitly on the unified data model and domain knowledge and are presented in the results section. The utility of the DAIRIE interface is demonstrated using information relating to mastitis because breeding decisions to control mastitis are complex problems requiring managers to manipulate information in many ways to make their decisions. Mastitis is a problem of economic importance to dairy producers affecting many herds and may make use of information and knowledge from several sources.

Interface: Three common forms were created: the data selection form, the data aggregation form for setting the time-period and herd level horizon, and the data display form for displaying the results. Figure 6.3 shows what each form contributes to the SQL query and progression through the forms exploits the focus of the user on different aspects of the data retrieval process to build the query and constrain the choice. There is a directed progression through the forms from the data selection form (1) to the data aggregation form (2) to the data display form (3) – see Figures 6.4, 6.5 and 6.6 respectively for greater detail of each form.

Data Selection Form: As an initial step the user is asked to select the data desired. Individual data fields are presented to the user for selection. Field names are grouped under 5 subheadings (Production Fields, Lactation Information, Production EBV Fields, Classification Fields, Conformation EBV Fields) corresponding to sub-groupings of the information in the database (see Appendix 1 for a list of the individual fields that one can retrieve from the underlying DBMS along side the text names used in the underlying source). The subgroups correspond to data definitions in script module.

The fields are placed on tab-sheets (see Figure 6.4) according to the sub-groupings described above. All the variables are kept on a single form but can be viewed according to logical groups. The arrangement avoids formatting difficulties that placing all the variables on one page might create for users to find and select the data that concern them. This also accommodates limitations that may arise from differences in individual user interface environments (e.g., monitor size).

Data can be selected directly by using the mouse to point to individual field check boxes and "clicking" on the required field. Data can also be selected using preset choices (including a "clear all" selection of all fields in a sub-grouping). For example, a mastitis button option is included to preselect points that have been predetermined as appropriate by the authors. Figure 6.4 shows the data selection form with the mastitis option selected; the production information tab-sheet is shown. Any grouping of fields selected can be further added to or subtracted from; the user has full control to select any combination.

Only information available to the herd being considered will be included in the data selection form. For example if a farm did not classify their herd neither the classification subgroup nor the classification option button are made available for that herd. DAIRIE determines which underlying tables are available to the user and adjusts the appearance of the GUI accordingly. Data selection necessarily precedes the setting of time length and herd level since the dimensions of the query are dependent on the data.

Data Aggregation Form: The data aggregation form allows the user to set the length of time and the herd level of interest. Figure 6.5 illustrates the form used to set relevant dimensions after the data fields are chosen. This form also allows the user to include common variables (individual animal data fields) in the query not available on the data selection form. The user must select the length of time dimension first then herd level is selected to support the underlying programme logic and settings for the time dimension.

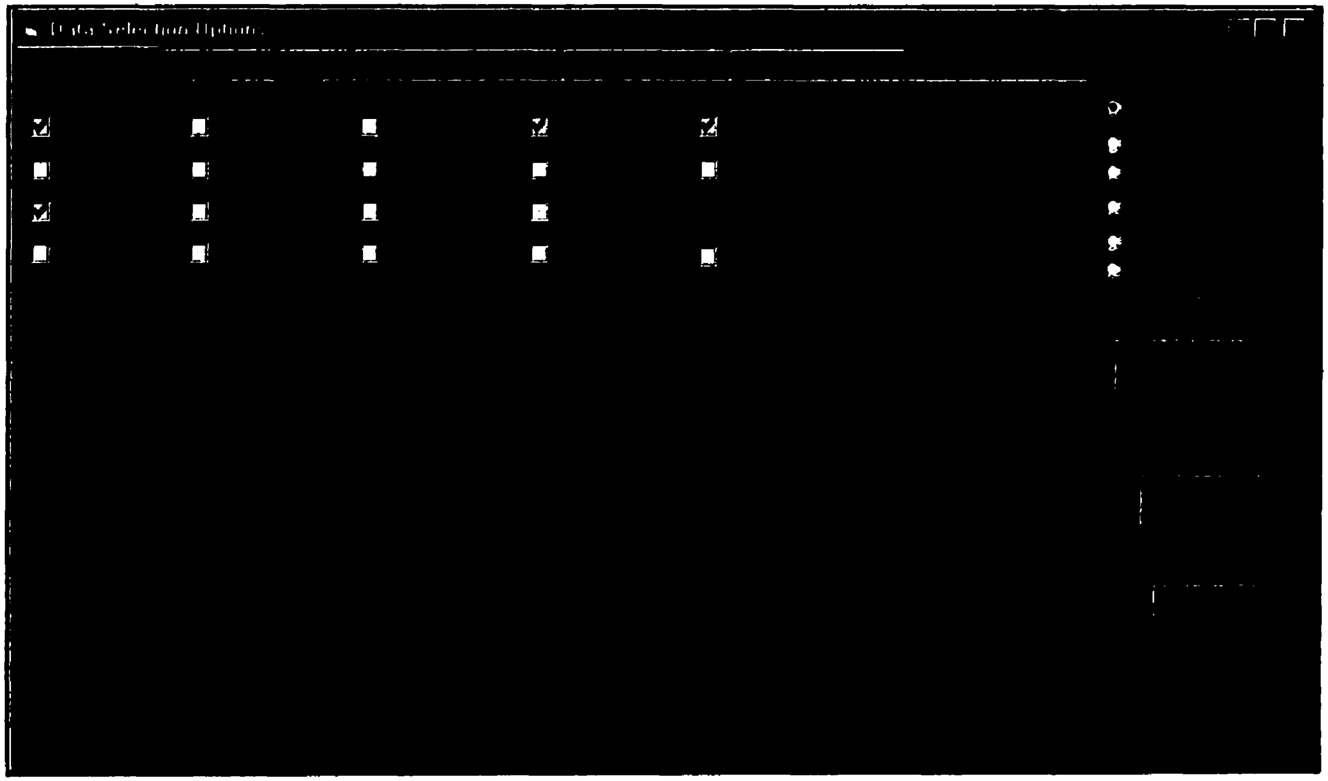


Figure 6.4 Data selection form production fields view when mastitis option is selected.

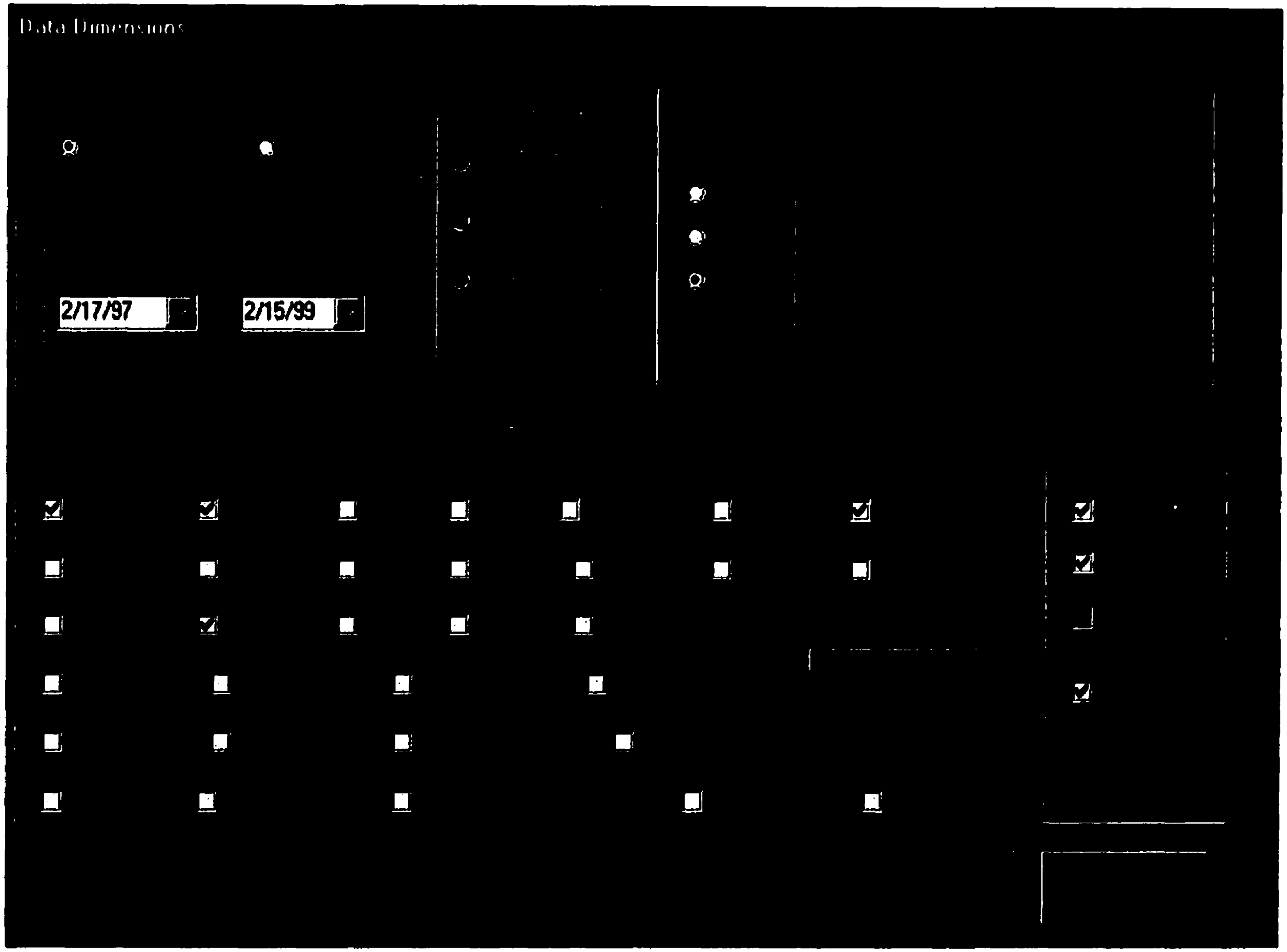


Figure 6.5 Data aggregation form with 'cow' level and some common cow variables selected.

Time length can be set by predefined lengths or be a length selected by the user (see Figure 6.5) and the user can toggle back and forth between these two methods. In this version of the prototype the predefined time periods are latest date, past year, or complete history. The predefined periods could vary in number and length of periods depending on the amount of information in the system and the significant periods identified by domain experts. Individual data sources have different time periods and DAIRIE must be able to reflect these differences according to the data elected by the user. A DAIRIE generated record set interacts with the data sources to identify which ones have been selected and can make the appropriate time periods available on the form.

The user could also determine specific time period by use of a pull-down list of the appropriate dates. Once this option is selected Figure 6.5 shows that the 'Start of Period' and 'End of Period' must be selected (one cannot proceed otherwise). The user can select any date which appears in the pull-down list -which will encompass the whole period of time available from the data source. This means that the user has the ability to select any time period of interest.

Table 6.1 shows the different time periods that existed among the date sources which could be used to populate the pull-down menus and set predefined time lengths. The dates in the time length dimension are dependent upon the data fields selected and the smallest increment available will be defaulted. Once data are selected from a subgroup the governing date variable will be included in the time dimension settings. The prototype validates which time periods are available and selects the shortest to populate the record set through a predefined SQL query.

Herd level (see Figure 6.5) can be specified over the entire herd for an individual animal or for groups of animals. The individual animal level allows one to focus on details about individual animal records, the herd level will allow exploration of details concerning all animals active in the herd, and the group function will provide information about identified groupings within the herd. 'Herd' level aggregates all the records in the herd into a single average for all parameters selected. All levels resolve information across the time frame indicated and only animals that are present in the herd during the specified period will be included.

It is possible to make sub-groupings of 'herd' and 'group' levels to illicit more detailed information on the 'herd' and 'group' level. The sub-groupings available in the prototype are the parity and the time that records were made (i.e., test date, classification date or evaluation date).

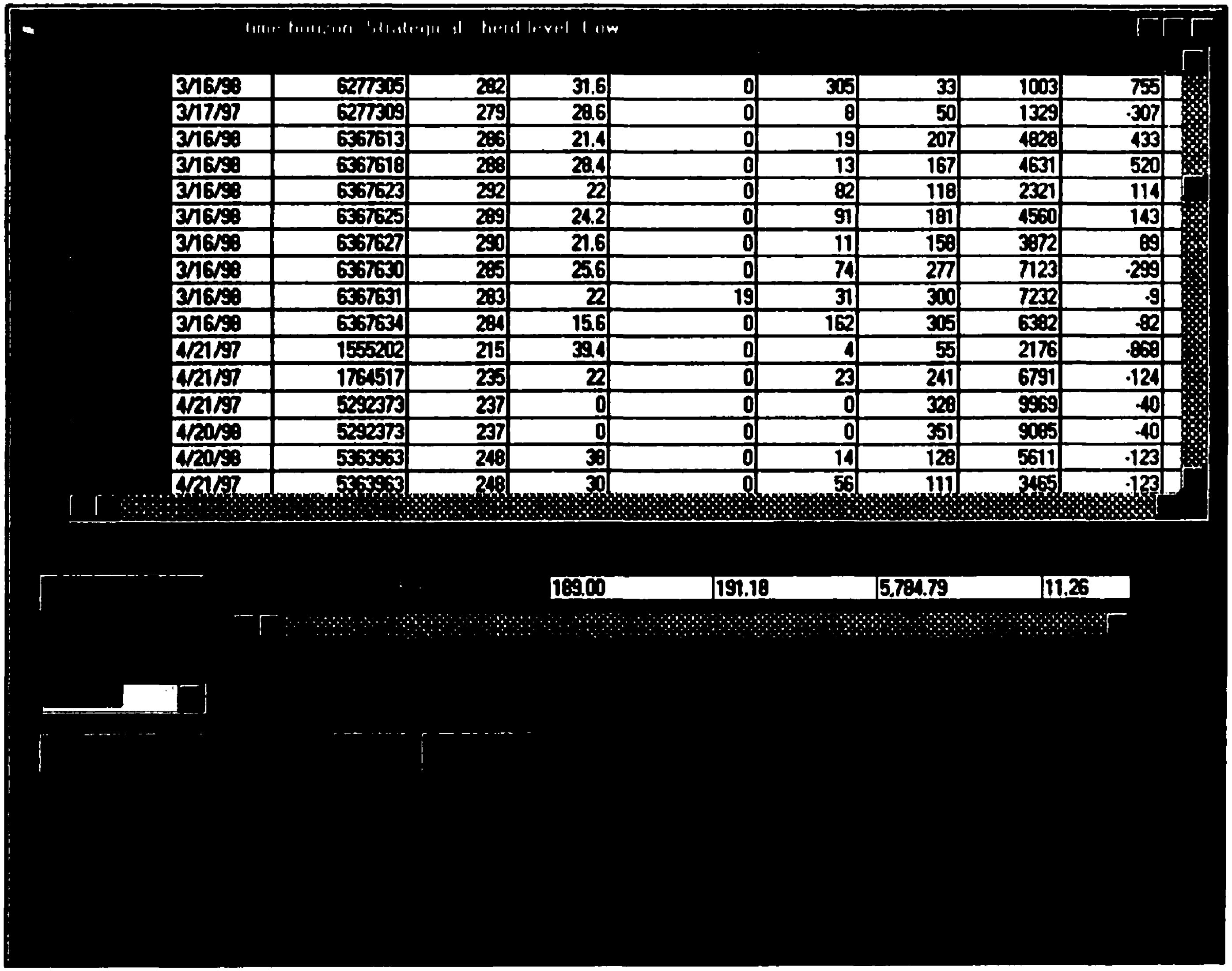
Table 6.1 Time intervals between information collected by type of information.

Data Type	Date Variable	Time Period (approximate)
Test	Test Date	30 d
Genetic Evaluation	Evaluation Date	90 d
Lactation	Calving Date	390 d
Classification	Classification Date	360 - 540 d

The mastitis option can be considered, where fields from production, evaluation (both production and conformation) and classification information are included. If, for example, 'herd' level were selected, test date, parity, the date of classifying, and evaluation date would be made available to form subgroups for the information query. Were parity selected, however, test date, the date of classifying and evaluation date would be available, but not parity since it is specified as the main grouping.

When 'cow' level is selected common variables can be selected from the 'Cow variables' area of the data aggregation form (see Figure 6.5). These common variables are fields with information that commonly related to many data sources. Information from all data sources which cannot be aggregated across 'herd' or 'group' level, such as registration number or days in milk, is included in the common variables area of the data aggregation form. The DAIRIE structure incorporates these data fields here where, if the 'cow' level is selected, they can be explicitly associated with individual animals. If data are selected using preset choices, common variables can be included. Again using the mastitis option as an illustration, explanation code is selected automatically. Other common variable fields can be included or, if explanation code is not of interest to the user, it can be deselected.

Data Display Form: Figure 6.6 is an illustration of the data display form where the results of the composed data query request are displayed. The retrieved record set is displayed similarly to a spreadsheet format by the main display of the form. The user can re-sort the retrieved information using a pull-down list of the number of columns desired.



3/16/98	6277305	282	31.6	0	305	33	1003	755
3/17/97	6277309	279	28.6	0	8	50	1329	-307
3/16/98	6367613	286	21.4	0	19	207	4828	433
3/16/98	6367618	288	28.4	0	13	167	4631	520
3/16/98	6367623	292	22	0	82	118	2321	114
3/16/98	6367625	289	24.2	0	91	181	4560	143
3/16/98	6367627	290	21.6	0	11	158	3872	89
3/16/98	6367630	285	25.6	0	74	277	7123	-299
3/16/98	6367631	283	22	19	31	300	7232	-9
3/16/98	6367634	284	15.6	0	162	305	6382	-82
4/21/97	1555202	215	39.4	0	4	55	2176	-868
4/21/97	1764517	235	22	0	23	241	6791	-124
4/21/97	5292373	237	0	0	0	328	9969	-40
4/20/98	5292373	237	0	0	0	351	9085	-40
4/20/98	5363963	248	38	0	14	128	5611	-123
4/21/97	5363963	248	30	0	56	111	3465	-123

189.00 | 191.18 | 5,784.79 | 11.26

Figure 6.6 Data display form displaying information when mastitis option selected.

Additional information on specific variables retrieved are included below the averages display window (e.g., the somatic cell count (SCC) button).

The averages display window (see Figure 6.6) is created from the retrieved record set at run time by averaging all non-null and non-zero data (deals with missing data) from each field. A control object called the Grid control was used to assign information from the record set to each cell and assign the field name as the column heading. Averages give the user an internal herd reference point to compare individual herd records shown in the MSHFlexgrid control when 'cow' level is displayed. This is not required for 'herd' and 'cow' levels since they are aggregated over a number of animals and averages are computed by the DBMS.

The MSHflexgrid and Grid display the principle retrieved record set however specific selected information can be displayed through associated record sets, and information about the query collected from the data selection and data aggregation forms. Somatic cell count and explanation-code buttons, displayed below the averages display window, access information on fields which are associated with the query but are derived from independent record sets.

The record set can be thought of as a transitory representation of the on-farm situation at one instance of its existence and permanent record can be created locally by the user in an ASCII text format. Thus, information for a specific problem or view of data according to user view from many sources can be placed together in one format in the same medium for further access, manipulation and consideration.

Modules: Modules were written code which perform specific functions in DAIRIE, independent of the forms. Underneath the forms seen by DAIRIE users, at the code and created objects layer (see Figure 6.2), the modules take the information input by the user, create the SQL code and display the retrieved record sets. The SQL module takes input from the user (through forms) into a set of functions to create the SQL query (Figure 6.3) – one function for each main SQL clause (i.e., "SELECT," "FROM," "GROUP BY," and "WHERE" clauses). The script module contains knowledge of the unified data model underneath DAIRIE without binding it explicitly to the fields described in the data selection form making DAIRIE more than a user interface. The script module also indicates which data sources are relevant to the current SQL query. Table 6.2 shows the definitions contained in the Script module which represents the sources and significant fields such as time and unique animal identifiers (e.g., classification date and registration number for the classification data).

Table 6.2 A delineation of Script Module contents which allows DAIRIE to interface with the database for the DBIS.

Logical	Name	ID Type	Date Type
CowTabl	Cows	DHASID; Registration No; Sire ID; Dam ID	Birth Date
TstDTabl	DHI Test-Day	DHASID	Test Date
LactTabl	DHI Lactation	DHASID	Calving Date
CnfrTable	Classification	Registration No	Classification Date
EBV CnfTabl	BA Cow Conformation Proof	Registration No	Evaluation Date
EBVProTabl	BA Cow Production Proof	Registration No	Evaluation Date
EBVIndexTabl	BA LPI	Registration No	Evaluation Date

Knowledge Implementation: Knowledge in the prototype was implemented around determining important (and relevant) user preferences of the dimensions of the data available and the utilisation of the information once retrieved. User preferences were developed, based on knowledge of information available, requirements for use, and constraints of information; guiding the user through the system to create functional and useful queries. This required an expertise in dairy cattle industry and information systems as well as knowledge-based decision support. Modules containing specific knowledge of individual fields were created to transform the information retrieved with respect to the dimension of the information selected. The SCC and the explanation-code modules, for example, support the SCC and the explanation-code buttons on the Data display form (Figure 6.6). By selecting these option buttons the user can further explore these fields to discover information that may be useful in the decision-making process.

Discussion

The DAIRIE prototype supports user needs to retrieve herd output information available in the DBIS dynamically. Only data selected by the user are presented; moreover it is presented across a specific span of time and at a herd level. Currently parity is the only group setting available between 'cow' and 'herd' level but other groupings could be added, e.g., dry cow, late lactation, feeding groups, etc. The data driven approach to support decisions uses knowledge selected pursuant to the data selected. Control of the data by the decision maker may improve the utility of information through an improved ability to perceive and remember salient features of herd output information.

Data retrieval on the farm can therefore, be more effectively unified and specific. There is a single gateway for information so the users only need to deal with a single interface to the DBIS rather than separate entities for each source of information. The potential is there to combine the knowledge and information among organisations providing data in new ways to discover knowledge. The filtering of information occurs before a user receives it and only information specific to the problem at hand is retrieved. The volatile data spaces can be stored directly in an on-farm PC to create permanent records.

Data retrieval activities that are tedious to perform or poorly suited to humans involved in the DBIS on-farm are minimised. The capturing and storage of information on-farm is automated, reducing the time for retrieval and storage, and the potential for introducing errors (Chapter 3). High volume information from many sources can be processed rapidly without human errors occurring, such as mis-keying or data processing errors (Ahituv, 1987; Chapter 3).

On-farm processing activities involving human agents, such as decision-making can make better use of herd output information flowing from DBIS participants. Problem-based, preset choices of data fields (e.g., the mastitis button) use domain knowledge to include fields that the experts consider important as well as allows users to add or remove fields. Once retrieved, the information can be considered together at the time of decision making. There is less potential for some information to be underutilised since all relevant information selected is kept together in the same ASCII file.

Knowledge was implemented, adhering to the data driven concept of decision support and modules were developed to apply knowledge from the source organisations in the domain. These modules create a knowledge set which takes advantage of relationships among data fields in different combinations. They reside independently from the fields and are employed only when a user selects a field to which the knowledge applies. This approach to using knowledge sources increased the intelligence density of information retrieved by DAIRIE without constraining the user to structure the information chosen in a specific way nor to look at information they would not normally consider. At the same time it does not allow knowledge to be misappropriated or misused to draw faulty conclusions about the information. More modules need to be created to broaden and deepen the knowledge applied from the DBIS and make DAIRIE useful for practical on-farm breeding decision support for output herd information available.

Implementation of the system must be propagated in distributed architecture over a network, such as the Internet. The prototype was developed on local a PC but the framework is extendible to a distributed environment. Both the DBMS and the VB6 programming environment are compatible with distributed architectures, making them suitable for applying the prototype over a network. For example, the interface could be imbedded in a Web page as a front-end to access the data in the system. Distributed network requirements raise issues of autonomy and heterogeneity of components, query parsing and optimisation, synchronisation and security (Attaluri et al., 1995; Ba et al., 1997; and Nicol et al., 1993). Without minimising these issues the focus of this study concerned investigating data specific issues for improving the effectiveness of dairy farm specific information systems; others have undertaken the job of dealing with these issues, which exist at another layer, more generally and thoroughly than our work at this time. Having clarified that our intentions were not to deal with these issues explicitly in this study, the development environment chosen does have the facility to support heterogeneous databases and different platforms, allowing future research of these dimensions. The development of an agent to parse queries could be a next step to exploring these issues; other issues could at least be partially addressed by employing components and technologies available through this, or other development environments, and technologies employed at other layers that are important in distributed architectures. Well-described approaches can be methodically applied as the framework is implemented in the dairy industry.

DAIRIE is expandable to include other information specific for breeding problems not shown here or even to include information outside the breeding domain. While the latter concerns fundamental domain definition and modular expansion of the data model of the DBIS (Chapters 3, 5), the former issue involves adding fields to the Data selection form available from the underlying DBMS (Figure 6.2).

Further information system automation may be achieved by integrating DAIRIE with decision-support systems to increase the effectiveness of the decision-making process on-farms. Both data selection and retrieval components could be implemented to expect requests from and deliver information to non-human processing agents. Misztal and Lawlor (1999) propose that an on-line breeding system could be supported with output such as DAIRIE produces. Legacy software provided by participating data source organisations -- such as, Vision 2000 by PATLQ -- could be dynamically linked with DAIRIE as a mediation agent on-line. A problem would be that the current DAIRIE may not be able to fully provide the information requested and this deficit would have to be solved through other online mediation agents such as additional modules added to DAIRIE.

Conclusions

The DAIRIE prototype was developed as a GUI to facilitate on-farm retrieval of information from the Dairy breeding information system. It improves effective integration of data retrieval and storage among information sources and their clients through an electronic medium to support an industry-wide DAiry information system. The independence of the information and knowledge sources from DAIRIE, and from each other are maintained. The information independence is primarily due to the script module which establishes the relationship between DAIRIE and the DBMS which supports it. The knowledge sources are transparently incorporated as individual modules and functions.

Through this tool human processing in on-farm decision making can make maximum use of information flowing in the DBIS while minimising some human limitations (i.e., retrieval, integration and manipulation). Domain knowledge can be employed without pre-determination of the structure of the data from many sources in the DBIS. The investigation of this prototype as an automated decision-support tool in real-time implementation on-line could yield information on its reliability, consistency and the timeliness of its delivery mechanisms, leading to greater confidence in its implementation to support end-users in the field. Investigations into the most appropriate format for presenting retrieved information, and the greater employment of different types of knowledge should further improve the prototype and effective use of all information available for decision making.

Acknowledgments

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, les Fonds de l'Autoroute de l'Information du Gouvernement du Québec, and the Cattle Breeding Research Council of Canada.

References

Attaluri, G K, D P Bradshaw, N Coburn, P-Å Larson, P Martin, A Silberschatz, J Slonim, and Q Zhu. 1995. The CORDS Multidatabase Project. IBM Systems Journal 34(1): 39-62.

Ba, Sulin, Ravi Kalakota, and Andrew B Whinston. 1997. Using client-broker-server architecture for Intranet decision support. Decision Support Systems 19(1997): 171-192.

Dertouzos, M L. 1995. Communications, computers and networks. Page 22 in Sci. Am., special issue: The computer in the 21st century, NY, New York.

Dhar, V, and R Stein. 1997. Intelligent Decision Support Methods: The Science of Knowledge Work. Prentice Hall, Upper Saddle River, NJ, USA.

Dutta, Soumitra, Berend Wierenga, and Arco Dalebout. 1997. Designing Management Systems Using an Integrative Perspective. Communications of the ACM 40(6): 70-79.

Harrington, J L. Relational Database Design Clearly Explained. AP Professional, Chestnut Hill, MA. 1998.

Ivanovic, M K, and Z D Budimac. 1999. Application of Agent-Oriented Workflow in Agriculture. Pages 337 - 343 *in* Proc. 2nd European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment, Sept. 27-30, 1999. Bonn, Germany.

Koh, Chang E, and Hugh J Watson. 1998. Data management in executive information systems. Information & Management 33: 301-312.

Lacroix, R, K M Wade, R Kok and F Hayes. 1994. Functional Characteristics of a Decision Support System for Dairy Cattle Management. International Summer Meetings of the American Society of Agricultural Engineers, Kansas City, MO. Paper No. 94-3027 ASAE, 2950 Niles Rd., St. Joseph, MI 49085-9659, USA.

McFadden, Fred R, J A Hoffer. 1994. Modern Database Management. 4th ed. Benjamin / Cummings Publishing Company, Inc., Redwood City, CA.

Microsoft. 1997. Visual Basic Guide to Data Access Objects, Crystal Reports for Visual Basic. Microsoft Corporation, Document No. DD93015-1296.

Nicol, John R, C Thomas Wilkes, and Frank A Manola. 1993. Object Orientation in Heterogeneous Distributed Computing Systems. Computer 26 (June): 57-67.

Nord, Jeretta Horn, and G. Daryl Nord. 1995. Executive information systems: A study and comparative analysis. Information & Management 29: 95-106.

Pietersma, D, R Lacroix, and K M Wade. 1998. A Framework for the Development of Computerized Management and Control Systems for Use in Dairy Farming. J Dairy Sci(81): 2962 - 2972.

Varhol, Peter D. 1995. Three routes to OLAP (ESSbase from Arbor Software, Express from IRI Software, and Lightship from Pilot Software). Datamation 41(Aug. 15): 57-61.

Young, Dale, and Hugh J Watson. 1995. Determinates of EIS acceptance. Information & Management 29: 153-164.

Chapter 7. Summary and Conclusions

A means by which information, belonging to various dairy organisations, and related to specific animals on dairy farms, could be used by farmers to improve breeding decisions made was described. The framework involved several distinct components: the comprehension and analyses of DBIS in terms of components and performance in networks (particularly via the Internet) with respect to on-farm decision-making. Modelling of the DBIS was based on the understanding gained from this analysis which led to the ultimate implementation of a mechanism to improve the interactions of farmers with the information flowing in the DBIS according to the developed model.

The information systems approach developed is useful for describing and analysing dairy-breeding information systems. As on-farm producers require more complex information structures, and the volume of supporting information increases, this approach could be employed to understand and respond to their needs. The systems approach may also be utilised to comprehend the role and needs of other participants in the DBIS. In general, the approach could be employed to analyse other types of multi-organisational information systems (including complex ones) and could even prove beneficial as an analytical tool for investigating system performance under various strategies for coping with the growing volume of information in the system.

The systems approach was successfully used to explore the exploitation of the Internet in the Canadian DBIS. The role the Internet plays can be used to improve the exploitation of herd output information. In order to accomplish this, DBIS participants should be aware of how to exploit WWW faculties and computer technology appropriately in order to understand specific user needs and provide a full range of necessary DBIS functions adapted to this medium. As DBIS users become accustomed to their role as information consumers, particularly through the Internet or other electronic networks, they will insist (and eventually rely) upon highly-developed services such as cooperative information provision and differentiated services.

The unified data model described here should be useful as a basis to exploit information in dairy breeding even more. Ultimately the unified data model developed should serve as a platform for offering common on-line services, regardless of the off-farm organisation that are accessed. Organisations should collaborate to clarify and improve information definitions, enrich and complete the metafile description of the domain to include all data available at logical and physical implementation levels. This would play a strong role in protecting proprietary information and assuring support for legacy information and applications. Data models are representations of reality and thus, the relationships between breeding information and other biological, economic and managerial aspects of dairy farming could allow the expansion of the model to include other important aspects of dairy production. On implementation, modular expansion of the model to include information for other on-farm management activities could be accomplished without disrupting user access. Given the similarities between DBMS and data warehouses, data warehousing should be explored as a suitable physical schema of the unified data model.

It has been demonstrated, through the development of the DAIRIE prototype, that it is possible to develop breeding decision-support software to retrieve, aggregate and analyse information driven by on-farm user requests. Domain knowledge can be employed to support dynamic queries for information without prior knowledge of the information structure. Implementing domain knowledge as modules, which can then be deployed to support various aggregations of information, is only one method of applying expert knowledge in this context. While the concept worked well, exploration in this generation of the prototype was limited: knowledge employment was not rich and should be pursued to a greater extent. Further study should investigate the effect that dynamic information selection and aggregation would have on the adoption of DSS and the acceptance of advice given. Further investigations of this method of knowledge support for dynamic information selection and aggregation should be pursued. The validation of the prototype is an iterative process that consists of testing functionality, validating data manipulation, knowledge incorporation and assessing suitability for end-users.

Other approaches to employing knowledge in the prototype could be investigated. The blackboard approach of employing knowledge sources with a control mechanism is another method that could be explored. In this context, where the decision activities are limited in number and recognised experts exist, it is possible that this well described method of opportunistic reasoning could work well. This should not necessarily be seen as a replacement to the aspects of decision support developed in this framework (e.g., the data driven approach for supporting farmers), but rather as an enhancement. For example, it could be used for a particular DSS on mating decisions, developed to work in conjunction with DAIRIE. More structured knowledge-based methods could also be employed (e.g., a specific module developed *within* DAIRIE or a module that uses the output from DAIRIE). Data requirements would be predetermined and the volatile data space would be called through views in formats and dimensions required.

The implementation of multi-organisational information systems for breeding, or any area of on-farm decision-making, is potentially a costly and lengthy process. It could, however, yield many benefits to all participants if the players make the proper investment. Off-farm organisations should be willing to make critical decisions, perhaps outside of their short term interests, in view of the potential benefits of collaborating in the creation of such supra-organisational cooperative frameworks (e.g., a means to offer valuable support presently to dairy producers over the Internet and an effective framework to develop useful knowledge-based applications). Short term financial disincentives and a possible low return on value-added information should be endured until users make the technical and attitudinal shifts and are able to appreciate the value of the information provided. The development of a suitable data model may be costly, require compromise and be time consuming. However, cooperation among off-farm organisations could lead to a broadened client base and cooperation with regard to costs for development of information-system support. The long-term benefits include optimal use of information infrastructures like the Internet to gain insights into user preferences, and the development of appropriate applications to serve information needs competently through emerging technologies such as on-line analytical processing or e-commerce.

Chapter 8. Contributions to Knowledge

This thesis has resulted in the following original contributions to knowledge:

1) **The development of a methodology that describes and implements information in the dairy production domain as an entity for decision support for on-farm users.** This methodology for approaching the implementation of decision support over a large and variable domain in the dairy industry should serve as a foundation to promote further development of decision-support techniques in dairy production domains. This methodology could be applied to any industry where the members are distributed and autonomous, have varying levels of experience in information systems available to them, and have access to and control of the information supported by the system described.

2) **The systems approach, useful for analysing information systems, was extended to be useful in analysing multi-organisational information systems.** Information systems in the dairy breeding industry are distributed among several independent organisations. The systems approach described here was adapted from prior applications as a method to understand and analyse information systems that span many organisations in a complex web of participants.

3) **Dairy breeding activities were described from an information system's perspective.** Information support and use were investigated as a distinct layer in the dairy industry, separate from the *real world* entities they represent and the *tools* that propagate them. A number of systems that involve the maintenance, exchange and production of information important for dairy breeding decision making were described and compared including a number of national dairy breeding information systems.

4) **Dairy breeding information systems that support on-farm decision-making were examined.** Using the systems approach, a number of dairy breeding information systems were investigated including those in Canada, USA, The Netherlands and Denmark. The role of the Internet was also examined for some of these industries.

5) **A unified data model of information available in the dairy breeding industry was developed.** Using a process called information engineering a formal methodology is given for developing data models that span the sources of information in an industry-wide information system. The unified data model itself creates a bridge from understanding dairy information systems to describing them in terms that are useful for the development of data systems.

6) **A tool was developed and demonstrated to integrate and retrieve information over a dairy breeding information system which can apply knowledge to support decision making.** The interface that was developed -- DAIRIE -- implements a method of accessing specific information products that are directed by the user and relate to an on-farm decision. The information retrieved is also not bounded by the source of origin. What is achieved follows the principle of insulating DAIRIE users from having to have intimate knowledge of how and where information is stored, and the underlying boundaries among the sources. Furthermore, it is based on principles of data-driven support: knowledge can be applied dynamically in the system, according to the data selected

Bibliography

Ahituv, N. 1980. A Systematic Approach Toward Assessing the Value of an Information System. Management Information Science Quarterly, Vol 4, No 4 (December): 61-75.

Ahituv, N. 1987. A metamodel of information flow: a tool to support information systems theory. Comm. of the ACM v 30, Sept:781.

Ahituv, N, S Neumann and H Norton Riley, eds. 1994. Principles of Information Systems for Management 4th Ed. Business and Educational Technologies: 22-36, 77-121.

Allore, H. G., and L. R. Jones. 1995. An Approach to Summarize Somatic Cell Score Trends for a Data-Driven, Decision Support System. J. Dairy Sci. 78(6): 1377-1381.

Allore, H. G., L. R. Jones, W. G. Merrill, and P. A. Oltenacu. 1995. A Decision Support System for Evaluating Mastitis Information. J. Dairy Sci. 78(6): 1382-1398.

Appleman, R, and J Noble. 1997. Using DHI Records to Manage The Dairy Herd. U.S. National Dairy Database. HTTP://www.inform.umd.edu/EdRes/Topic/AgrEnv/ndd.

Archer, A. A., R. Lacroix, and K. M. Wade. 1996. A network of information for the dairyindustry. J. Dairy Sci. 79 (Suppl. 1):174. (Abstr.)

Attaluri, G K, D P Bradshaw, N Coburn, P-Å Larson, P Martin, A Silberschatz, J Slonim, and Q Zhu. 1995. The CORDS Multidatabase Project. IBS Systems Journal 34(1): 39-62.

Ba, S, R Kalakota, and A B Whinston. 1997. Using client-broker-server architecture for Intranet decision support. Decision Support Systems 19(1997): 171-192.

Ballou, D, and H Pazer. 1985. Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems. Management Science 31 (2): 150-162.

Ballou, D, R Wang, H Pazer, and G Kumar-Tayi. 1998. Modeling Information manufacturing Systems to Determine Information Product Quality. Management Science 44(4): 462-484.

Bakus, G B C, G Th Timmer, A A Dijkhuizen, V R Eidman and R P King. 1995. The Impact of a Decision Support System for Strategic Pig Farm Planning on the Advice of Extension Officers. Computers and Electronics in agriculture, 12: 51-64.

Barrett, J R, and D D Jones, eds. 1989. Knowledge Engineering in Agriculture. American Society of Agricultural Engineers, St Joseph, U.S.A.

Barrett, S, and B Konsynski. 1982. Inter-Organization Information Sharing Systems. MIS Quarterly, Special Issue: 93-105.

Batchelor, W D and R W McClendon. 1992. A Blackboard Approach for Resolving Conflicting Irrigation and Insecticide Scheduling Recommendations. Transactions of the ASAE, 35(2): 741-747.

Bauer, M A, N Coburn, D L Erickson, P J Finnigan, J W Hong, P-Å Larson, J Pachl, J Slonim, D J Taylor and T J Teorey. 1994. A Distributed System Architecture for a Distributed Application Environment. IBM System Journal, 33(3): 399-425.

Bernstein, P A. 1996. Middleware: A Model for Distributed System Services. Communications of the ACM. 39(2): 87-98.

Biron, Sylvain. 1998. Outils génétiques et expertise animale. Des notes d'un cour offert par le PATLQ, (Septembre), 1998.

Biron, S, B Farmer, R Moore, H Keurentjes and M Caron. 1998. Le contrôle laitier du troisième millionaire. Pages 113-135 *in* Symposium sur les bovins laitiers. Conseil des productions animales du Québec inc., Ste-Hyacinthe, Québec.

Borchersen, S. 1996. The Federation of Danish A.I. Societies. Pages in 34-37 Annual Report, National Committee on Danish Cattle Husbandry.

Brick, B, and J Henry. 1999. Enterprise Portals: Not Just Information Delivery. Intelligent Enterprise 2(16): 24-27.

Bunge, M A. 1979. Ontology: A world of Systems (vol. 2). Dordrecht. D. Reidel.

Burnside, E. 1993. Genetic Resources. Pages 205 -237 *in* Animal Production in Canada, Martin, J, RJ Hudson and BA Young, ed. University of Alberta, Faculty of Extension, Edmonton, Alberta.

Burch, J G, F R Strater, and G Grudnitski. 1983. Information Systems: Theory and Practice, 3rd Ed. John Wiley & Sons, New York, NY.

Canadian Dairy Data Dictionary, The. 1989. Canadian Dairy Information Exchange, Brantford Ontario, Canada.

Christensen, D A and M I Fehr. 1993. The Dairy Industry. Pages 75-92 *in* Animal Production in Canada, J Martin, R J Hudson and B A Young, eds. University of Alberta, Faculty of Extension, Edmonton, Alberta.

Churchman, C W. 1968. The Systems Approach. Delta Books, New York, NY.

Codd, E F. The Relational Model for Database Management. Addison-Wesley Pub. Co. Inc., Reading, MA. 1990.

Covill, Randal Jorde. 1998. Implementing Extranets: The Internet as a Virtual Private Network. Digit Press, Boston.

Cue, R C, and J D Fletcher. 1998. Predicting Milk Supply. Pages 11-15 *in* 1998 Annual Research Report, Department of Animal Science, McGill University.

Craig, I. 1995. Blackboard Systems. Alex Pub. Corp. Norwood, NJ, USA.

Crowe, A S and J P Mutch. 1990. Assessing the Migration and Transformation of Pesticides in the Subsurface: The Role of Expert Systems. Water Poll. Res. J. Canada, 25(3): 293-323.

Daft, R L and R H Lengel. 1986. Organizational Information Requirements, Media Richness and Structural Design. Management Science, 32(5): 554-571.

Davis, J R. 1986. Possibilities for Expert Systems in Agriculture. Acta Horticulturae, 175: 157-161.

Davis, J C, G A Parmer, B R Smith, and E H Simpson III. 1994. Providing Internet connectivity for remote Alabama offices. Pages 718-721 *in* Proc. 5th Int. Confr. Comp. Agric., Orlando, FL.

Dertouzos, M L. 1995. Communications, Computers and Networks. Pages 22-29 *in* Scientific American, special issue: The computer in the 21st century, Rennie, J, M Press and JT Rogers, eds.

Dhar, V, and M Jarke. 1993. On Modeling Processes. Decision Support Systems 9: 39-49.

Dhar, V, and R Stein. 1997. Intelligent Decision Support Methods: The Science of Knowledge Work. Prentice Hall, Upper Saddle River, NJ, USA.

Dickson, G W, J A Senn and N L Chervany. 1977. Research in Management Information Systems: The Minnesota Experiments. Management Science, 23(9): 913-923.

Dijkhuizen, A A, R B M Huirne, A W Jalvingh, and J W Visscher. 1996. Towards Integrated Information Systems for On-Farm Decision Support. Pages 54-60 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Dill, D D. 1990. University/Industry Research Collaborations: An Analysis of Interorganizational Relationships. R&D Management 20(2): 123-129.

Dindorp, U, M Flensted-Jensen, and A P Ravn. 1996. A standard format for data exchange - SFD. Pages 368-373 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Dutta, S, B Wierenga, and A Dalebout. 1997. Designing Management Systems Using an Integrative Perspective. Communications of the ACM 40(6): 70-79.

Egeberg R. D., and D. G. Rice. 1994. ExtNet: extension information delivery in North Dakota. Pages 702-706 *in* Proc. 5th Int. Confr. on Comp. in Agric. Orlando, FL.

Englemore, R S, A J Morgan and H P Nii. Introduction. Pages 1-22 *in* Blackboard Systems, *eds* R. Englemore and T. Morgan. 1988. Addison-Wesley, Wokingham, England.

Erman, L D, F Hayes-Roth, V R Lesser and D R Reddy. The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty. Pages 31-86 *in* Blackboard Systems, *eds* R Engelmore and T Morgan. 1988. Addison-Wesley, Wokingham, England.

Etgen, W M, R E James and P M Reaves, Eds. 1987. Dairy Cattle Feeding and Management. John Wiley & Sons, New York, U.S.A.

Fiez, E A. 1998. Flow of Records Within NCDHIP. U.S. National Dairy Database. HTTP://www.inform.umd.edu/EdRes/Topic/AgrEnv/ndd.

Fournier, A, R Lacroix and K M Wade. 1997. Ouvrir les horizons grâce aux groupes de discussion. Le producteur de lait Québécoise. 17(9): 38-41.

Fowler, S R, J L Novak, and E H Simpson III. 1994. Modem-accessible information services for agric. and ext. Page 617 *in* Proc. 5th Int. Conf. on Comp. in Agric., Orlando, FL.

Frederiks, P J M, A H M ter Hofstede and E Lippe. 1997. Information and Software Techhnology. A Unifying Framework for Conceptual Data Modelling Concepts. Information and Software Technology 39: 15-25.

Gaines, B. R., and J. H. Boose, *eds*. 1988. Knowledge Acquisition for Knowledge-Based Systems: Knowledge-Based Systems Vol. 1. Academic Press, Harcourt Brace Jovanovich, New York, USA.

Ganguly, S and M Doordewier. 1996. PROXIMAL: A Database System for the Efficient

Retrieval of Genetic Information. Computed. Biol. Med. 26(3): 199-207.

Gauthier, L, and R Kok. 1988. A Microcomputer-Based farm Management/Operating System. Canadian Agricultural Engineering, 30: 69-76.

Getz, R R, K S Harker, D M Ihle, S D Adams and E H Simpson, III. 1994. Regional Delivery of Agricultural Weather Information Using the Internet. Pages 712-717 *in* Proc 5^{th} Int. Confr. Comp in Agric, Orlando, Fl.

Goldsmith, P D, J-F Forest, R Lacroix, and K M Wade. 1996. A Framework for Information Cataloguing and Management Information System Design: an Economic Perspective. Pages 688-701 *in* Proc. 6^{th} Int. Conf. Comput. Agric., Cancun, MX.

Goodhue, D L, Wybo, M D and L J Kirsch. The Impact of Data Integration on the Costs and Benefits of Information Systems. 1992. Management Information Science Quarterly, 16(3): 293-311.

Graves, M., E. R. Bergeman, and C. B. Lawrence. 1996. A Graph Conceptual Model for Developing Human Genome Center Databases. Comput Biol Med. 26(3): 183-197.

Greer, J E, S Falk, K J Greer and M J Bentham. 1994. Explaining and Justifying Recommendations in an Agriculture Decision Support System. Computers and Electronics in Agriculture, 11: 195-214.

Groeneveld, E, and T Gutzmann. 1990. Strategic Considerations in the Choice of Basic Software and Hardware for Information Systems in Animal Production. Computers and Electronics in Agriculture, 5: 135-149.

Groeneveld, E, and P Lacher. 1992. An Open information System for the Swine Production and Marketing Industry: Its Scope, Topology and Telecommunications Strategy. Computers and Electronics in Agriculture, 7: 163-185.

Hackathorn, Richard D. Web Farming for the Data Warehouse: Exploiting Business Intelligence and Knowledge Managmenet. Morgan Kaufmann Publishers Inc., San Francisco, Ca. 1999.

Hall, G M. 1991. Strategy, Systems, and Integration: A Handbook for Information Managers. Tab Books: Professional and Reference Books. Pa, U.S.A.

Harrington, J L. Relational Database Design Clearly Explained. AP Professional, Chestnut Hill, MA. 1998.

Hart, P. E., and J. Graham, R. 1997. Query-Free Information Retrieval. IEEE Expert 12(5): 32-37.

Hochman, Z, R Barlow, JF Ayres and CJ Pearson. 1991. A Multiple Domain Expert System for Beef Cattle Farmers. Computers and Electronics in Agriculture, 6: 143-156.

Hochman, Z, H Hearnshaw, R Barlow, JF Ayres and CJ Pearson. 1995. X-Breed: a Multiple-domain Knowledge Based System Integrated Through a Blackboard Architecture. Agricultural Systems 48: 243-270.

Hochman, Z, P T Mears, R J Rarquharson, J H Hindmarsh and CJ Pearson. 1995. A Multiple-domain Knowledge Based System: Coordination and Integration of Agronomic and Economic Knowledge Bases with Databases and Bio-economic Models. Agricultural Systems, 48: 141-161.

Hogeveen, H, E N Nordhausen-Strasser, J F Schreinemakers and A Brand. 1991.

Development of an Integrated Knowledge-Based System for Management Support on Dairy Farms. J DAIRY Sci 74: 4377-4384.

Hogeveen, H, J F Schreinemakers, E N Nordhausen-Strasser. 1992. Application of Conditional Causality in an Integrated Knowledge-Based System for DAIRY Farms. AI Applications 6(3): 5-16.

Hogeveen, H, M A Varner, D S Bree, D E Dill, E N Nordhausen-Strasser, and A Brand. 1994. Knowledge Representation Methods for DAIRY Decision Support Systems. J DAIRY Sci 77: 3704-3715.

Huber, G P. 1983. Cognitive Style as a Basis For MIS and DSS Designs: Much ADO About Nothing? Management Science, 29(5): 567-579.

Huber, G P and R R McDaniel. 1986. The Decision-Making Paradigm of Organizational Design. Management Science, 32(5): 572-589.

Ivanovic, M K, and Z D Budimac. 1999. Application of Agent-Oriented Workflow in Agriculture. Pages 337 - 343 *in* Proc. 2nd European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment, Sept. 27-30, 1999. Bonn, Germany.

Jacobs, B E. Applied Database Logic Volume1: Fundamental Database Issues. Prentice-Hall Inc., Englewood Cliffs, NJ. 1985.

Kemp, G J L, J Dupont, and P M D Gray. 1996. Using the Functional Data Model to Integrate Distributed Biological Data Sources, Pages176-185 *in* Proceedings Eighth International Conference on Scientific and Statistical Database Management. June 18-20, 1996. IEEE Computer Society Press, Stockholm.

Kimball, R. 1999. Working in Web Time. Intelligent Enterprise 2(16): 28-32.

Innes, R J. 1994. Designing and implementing a low-cost county office network. Page 691-694 *in* Proc. 5th Int. Confr. Comp Agric. Orlando, FL.

Jofre-Giraudo, E, D H Streeter and W Lazarus. 1990. The Impact of Computer Information Systems on Dairy Farm Management Decisions. Agribusiness, 6(5): 463 - 474.

Jones, L R and S L Spahr. 1991. IDEA: Intelligent Data Retrieval in English for Agriculture. AI Applications 5(1): 56-66.

Kador, J. 1996. The Ultimate Middleware. Byte 1996(April): 79-83.

Kalter, R J, A L Skidmore, and C J Sniffen. 1992. Distributed Intelligence and Control: the New Approach to Dairy Farm Management. Pages 171-176 *in* Proc. 4th Int. Confr. Comp. Agric. Ext. Prog. Orlando, Fl.

Kok, R, and L Gauthier. 1986. Development of a Prototype Farm Information Management System. Computers and Electronics in Agriculture, 1: 125-141.

Kok, R, and R Lacroix. 1993. An Analytical Framework for the Design of Autonomous, Enclosed Agro-Ecosystems. Agricultural Systems, 43: 253-260.

Koh, C E, and H J Watson. 1998. Data management in executive information systems. Information & Management 33 (1998): 301-312.

Koorn, D S. 1996. Electronic Data Interchange to and From Dairy Farmers. Pages 40-44 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture.

Wageningen, the Netherlands.

Kroeze G H, C Lokhorst, H J van de Beek and F de Vries. 1996. Pages 611-616 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Lacroix, R, J Huijbers, R Tiemessn, D Lefebvre, D Marchand, and K M Wade. 1996. Adding value to Dairy Herd Improvement information through the use of Fuzzy Logic. Page 885-891 *in* Proc. 6th Int. Conf. Comput. Agric., Cancun, MX.

Lacroix, R, and K M Wade. 1996. Expected benefits from the use of the information superhighway in dairy herd improvement. Pages 175-178 in Proc. 30th Biennial Session of the International Committee for Animal Recording. Performance recording of animals: State of the Art. EAAP Publication No 87, 1996. Veldhoven, The Netherlands..

Lacroix, R and K M Wade. 1996. L'Autoroute de l'Information en Agriculture: Pas Pour Demain, Pour Aujourd'hui. Le Producteur de lait Québécoise. 17(1): 26-29.

Lacroix, R, K M Wade and A Fournier. 1996. Nouvelles en Vrac. Le Producteur de lait Québecoise. 17(4): 22-23.

Lacroix, R, K M Wade, R Kok and F Hayes. 1994. Functional Characteristics of a Decision Support System for Dairy Cattle Management. International Summer Meetings of the American Society of Agricultural Engineers, Kansas City, MO. Paper No. 94-3027 ASAE, 2950 Niles Rd., St. Joseph, MI 49085-9659, USA.

Lacroix, R, K M Wade, S Des Marchais and R Marchand. 1998. L'Élevage des Bovin

Laitiers et Internet; l'Information Génétique (Partie 1). Le Producteur de Lait Québecois, 18(10): 53-55.

Lacroix, R, K M Wade and S Des Marchais. 1998. L'Élevage des Bovin Laitiers et Internet. Partie 2: L'Offre de Produit Génétique. Le Producteur de lait Québecois, 18(11): 43-46.

Lee, Dong-Tsan. 1996. Semantic Data Modelling Using Linear Logic. Information Processing Letters 60: 19-27.

Lescourret, F, M Genest, J Barnouin, Michelle Chassagne, and B Faye. 1993. Data Modelling for Database Design in Production and Health Monitoring Systems for Dairy Herds. J Dairy Sci 76: 1053-1062.

Lescourret, F, P Dunaud, J Barnouin, Michelle Chassagne, and B Faye. 1994. Computerized representation of feeding regimes in monitored dairy herds. Computers and Electronics in Agriculture 10 (1994) 151-166.

Levy, A, C Knoblock, S Minton and W Cohen. 1998. The Information Manifold Approach to Data Integration. IEEE Intelligent Systems and Their Applications 12(5): 12-16.

Lokhorst, C and G H Kroeze. 1996. Farm Comparison By Means of Automatic Data Transfer. The Journal of University of Technology and Agriculture, Bydgoszcz. No 199-Agriculture 40: 225-230.

Manguerra, H B. 1997. Internet-based Decision Support Systems. Resource: Engineering and Technology for a Sustainable World, Vol 4(4):11.

Martin, R. 1995. Élever, Faire Élever ou Acheter? Le Producter de Lait Québécois, 16, 3:

22-25.

McFadden, Fred R and J A Hoffer. 1994. Modern Database Management. 4th ed. Benjamin / Cummings Publishing Company, Inc., Redwood City, CA.

McKendrick, I J, G Gettinby, Y Gu, A Peregrine, and C Revie. 1994. Hybrid information Systems for Agriculture: The Case of Cattle Trypanosomiasis in Africa. Outlook on Agriculture 23(4): 261-267.

McKinion, J M and H E Lemmon. 1985. Expert Systems for Agriculture. Computers and Electronics in Agriculture, 1: 31-40.

McQueen, RJ, SR Garner, CG Nevill-Manning, and IH Witten. 1995. Applying Machine Learning to Agricultural Data. Computers and Electronics in Agriculture 12: 275-293.

Microsoft. 1997. Visual Basic Guide to Data Access Objects, Crystal Reports for Visual Basic. Microsoft Corporation, Document No. DD93015-1296.

Misztal, I and T J Lawlor. 1999. Supply of Genetic Information – Amount, Format, and Frequency. Symposim: Information Technology for Animal Breeding Purposes. J Dairy Sci 82: 1052-1060.

Millikin, M D. 1994. DCE: Building the Distributed Future. Byte: Special Report--Distributed Computing, 19: 125-126,128,132,134.

Morimoto, T, K Hatou and Y Hashimoto. 1995. Intelligent Control for Plant Production System. Pages 139-144 *in* Artificial Intelligence in Agriculture, A J Udink ten Cate, (ed.), Oxford, Ox, UK: IFAC.

Nadeau, M. 1994. Remote Connections, in Byte: Special Report: Distributed Computing, 19: 197-198,200.

Nelson, J. 1994. Information Exchange Networks and Agricultural Development. Outlook on Agriculture 23(4): 243-247.

Newman, S, T Lynch and A A Plummer. 1999. Aspects of success and failure of decision support systems in agriculture. In: H A Abbass and M Towsey (Eds.) Application of Artificial Intelligence, Optimisation and Bayesian Methods in Agriculture. Pp. 17-30. Queensland University of Technology Press, Brisbane (ISBN 1 86435 463 1).

Nidumolu, S R. 1995. Interorganizational Information Systems and the Structure and Climate of Seller-Buyer Relationships. Information & Management 28: 89-105.

Nicol, J R, C T Wilkes and FA Manola. 1993. Object Orientation in Heterogeneous Distributed Computing Systems. Computer, 26: 57-67.

Nord J H, and G D Nord. 1995. Executive information systems: A study and comparative analysis. Information & Management 29 (1995): 95-106.

O'Malley, C. 1995. The Wireless World: The Information Skyway is Coming Soon to a Wireless Device Near You, in *Popular Science*, Stepler, RL ed., 11: 56-62.

Olson, K E. 1991. Satellite Technology – A Farm Organization Perspective. J Dairy Sci, 75: 3261-2364.

OneSoft Corporation. 1998. The Extranet Solution. Pages 55-65 *in* Web-Weaving: intranets, extranets, and Strategic Alliances. Lloyd, P. and Boyle, P. Butterworth-Heinemann, Oxford, UK.

Papazoglou, M P and J Zeleznikow. 1991. The Next Generation of Information Systems – From Intelligence to Distribution and Cooperation. Pages 1-8 *in* The Next Generation of Information Systems: From Data to Knowledge. Papazoglou, M P and J Zeleznikow, ed. Springer-Verlag. From IJCAI-91 Workshops Sydney, Australia.

PATLQ. 1996. "Working Together." PATLQ News bulletin, Number 35 (May 1995.

Pedersen, J, G A Pedersen and U S Nielsen. 1996. Functional Traits in Danish Cattle Breeding. Pages 30 - 33 in Annual Report; National Committee on Danish Cattle Husbandry.

Pellerin, D, D Gilbert, J P Perrier and R Levallois. 1995. L'Élevage des Génissess, Combien ça Coûte? Le producter de lait Québécois, 16(3): 26, 27-30.

Pellerin, D, R Levallois, G St-Laurent, and J-P Perrier. 1994. LAIT-XPERT Vaches: An Expert System for Dairy Herd Management. J. of Dairy Sci., 77(8): 2308-2317.

Pietersma, D, R Lacroix, and K M Wade. 1998. A Framework for the Development of Computerized Management and Control Systems for Use in Dairy Farming. J Dairy Sci(81): 2962 - 2972.

Pinson, Suzanne D, Jorge Anacleto Louçã, Pavlos Moraitis. 1997. A Distributed Decision Support System for Strategic Planning. Decision Support System 20(1997): 35-51

Pountain, D. 1997. The Component Enterprise. Byte 97(May): 93-98.

Ramaprasad, A. 1987. Cognitive Process as a Basis for MIS and DSS Design. Management Science, 33(2): 139-148.

Revie, C W, I McKendrick, T Irwin, Y Gu, S. W J Reid, and G Gettinby. 1996. The Application of Hybrid Information Systems to Decision to Decision Support in the Veterinary and Agricultural Domains. Pages 48-57 *in* Proc. Society of Veterinary Epidemiology and Preventive Workshop, University of Strathclyde, Glasgow, March 27-29.

Ridson, P L. 1994. Internet's Role in the Technology Transfer Process. Proc. 5th Int. Confr. Comp. in Agric., Orlando, FL. 728-730.

Rimmington, G M, H S Hawkins, and J E Hoare. 1994. LandcareNET–promotion of lateral information transfer. Page 695-700 *in* Proc. 5th Int. Confr. Comp. Agric., Orlando, FL.

Robey, D. 1983. Cognitive Style and DSS Design: A Comment on Huber's Paper. Management Science, 29(5): 580-582.

Rowe, E. ed. 1997. Computer use is routine for them. Hoard's Dairyman, 1997(Jan): 50.

Rosskopf, K. 1999. Computer Use, Applications and Internet Use in Dairy Farming in New Zealand – the Leading Edge in the World? Pages 365-373 *in* Proceedings of 2nd European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment, Sept. 27-30, 1999. Bonn, Germany.

Rymer, J. R. 1996. The Muddle in the Middle. Byte 1996(April): 67-70.

Sargent, R, D Fuhrman, T Critchlow, T Di Sera, R Mecklenburg, G Lindstrom and P Cartwright. 1996. The Design and Implementation of a Database for Human Genome Research. Pages 220-225 *in* Proceedings Eighth International Conference on Scientific and Statistical Database Management, June 18-20, 1996. IEEE

Computer Society Press, Stockholm.

Schmoldt, D L and M Rauscher. 1994. A Knowledge Management Imperative and Six Supporting Technologies. Computers and Electronics in Agriculture, 10: 11-30.

Silver, G A, and M L Silver. 1994. Implementing communications networks. Page 396 *in* Data Communications for Business. Boyd and Fraser Publishing, Danvers, MA.

Smith, K D. 1994. The Indiana Extension Internet a Low Cost Full Service Solution. Pages 722-727 *in* Proc. 5th Int. Confr. Comp. in Agric., Orlando, FL.

Spahr, S L. 1993. New technologies and decision making in high producing herds. J. Dairy Sci. 76: 3269-3277.

Spahr, S L, L R Jones, and D E Dill. 1988. Expert Systems -- Their use in Dairy herd management. J. Dairy Sci. 71: 879-885.

Strasser, M. A Fuzzy Decision Support System For Dairy Cattle Culling. Masters of Science thesis. 1997. McGill University, Montreal, Canada.

Strasser, M, R Lacroix, R Kok, and K M Wade. 1997. An aggregated decision-support system for the recommendation of Dairy cattle culling. J. Dairy Sci. 80 (Suppl. 1):149.(Abstr.)

Strong, D M, Y W Lee, and R Y Wang. 1997. 10 Potholes in the Road to Information Quality. Computer. v. 30, Aug. 1997, p. 38-46.

Thompson-James, Margaret. 1999. Computer Use and Internet Use by Members of Rural Households. Rural and Small Town Canada Analysis, Vol. 1, No. 7 (May 1999), Statistics Canada. Catalogue no. 21-006-XIE.

Todd, P, and I Benbasat. 1992. The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. MIS Quarterly 16(3): 373-393.

Tomaszewski, M A. 1992. Using Advanced Computer Technologies to Increase Extension Effectiveness. J Dairy Sci 75: 3242-3245.

Tomaszewski, M A. 1993. Record-Keeping Systems and Control of Data Flow and Information Retrieval to Manage Large High Producing Herds. J Dairy Sci 76: 3188-3194.

Trustman, J and S Meshako. 1999. You Say You Want an Evolution. Intelligent Enterprise 2(16): 18-20.

Varner, M A and R A Cady. 1993. Dairy-L: An Electronic Information Exchange Network for Professionals Advising Dairy Producers. J Dairy Sci, 76: 2325-2331.

Varhol, P D. 1995. Three routes to OLAP (ESSbase from Arbor Software, Express from IRI Software, and Lightship from Pilot Software). Datamation 41(Aug. 15) 57-61.

Verstegen, J A A M, R B M Huirne, A A Dijkhuizen, and R P King. 1995. Quantifying Economic Benefits of Sow-Herd Management Information Systems Using Panel Data. Amer. J. Agr. Econ. 77(5): 387-396.

Voss D. 1994. Building the US Superhighway. Physics World, v 7 Sept: 43-47.

Wade, K. M., and R. Lacroix. 1997. The use of fuzzy logic membership values in ranking

Dairy cows for management purposes. J. Dairy Sci. 80 (Suppl. 1):232. (Abstr.)

Waksman, G., and M. Harkin. 1996. Green telematics: is Internet a breakthrough for farmers and extension workers? Pages 182-188 *in* Proc. 6th International Congress for Computer Technology in Agriculture on Information and Communication Technology Applications in Agriculture. Wageningen, the Netherlands.

Wiederhold, G. and M. Geneserth. 1997. The Conceptual Basis for Mediation Services. IEEE Expert 12(5): 38-47.

Winstead, J, and N Gershon. 1996. The World-Wide Web: A Global Source of Data and Information. Pages 21-28 *in* The Information Revolution: Impact on Science and Technology. Dubois, J.-E. and N. Gershon, ed. Springer-Verlag, Berlin, Germany.

Wybo, M D. 1993. An Assessment of the Validity of the Subunit Interdependence Construct in the Context of Information Systems. Working Paper, McGill University, Faculty of Management, Montreal, Quebec.

Young, D, and H J Watson. 1995. Determinates of EIS acceptance. Information & Management 29 (1995): 153-164.

URL References

ABS Global. 1997. <http://www.absglobal.com>

Agriculture Canada Electronic Information System. 1997. <http://aceis.agr.ca>

Altagenetics. 1997. <www.altagenetics.com/dairy-menu.htm>

American Dairy Science Association. 1997. <http://www.adsa.uiuc.edu>

American Society of Animal Science. 1997. <http://www.asas.uiuc.edu>

Animal geneticists discussion group. 1997. <http://chuck.agsci.colostate.edu/>

Canadian Dairy Network. 1997. <http://www.cdn.ca>

Centre for the Genetic Improvement of Livestock. 1997. <http://www.aps.uoguelph.ca/cgil>

Centre d'Inseminasation Artificielle du Quebec, 1997. <www.ciaq.com/holstein.htm>

Dairy-L. 1997. dairy-l@umdd.umd.edu

Fiez, E. A. 1985. Flow of records within DCDHIP. <http://www.inform.umd.edu/EdRes/Topic/AgrEnv/ndd/dairy/FLOW_OF_RECORDS_WITHIN_NCDHIP.html

Gencor. 1997. <http://www.gencor.ca>

Genetic Evaluations Discussion Group. 1997. gedg@ansci.nw.uoguelph.ca

Holstein Association. 1997. <http://www.holsteinusa.com>

Holstein Canada. 1997. <http://www.holstein.ca>

LaiToile. 1997. http://infosys.AgrEnv.McGill.CA/laitoile/

Oklahoma State University: WWW Virtual Library. 1997. <http://www.ansi.okstate.edu/library/dairy>

Ontario Dairy Improvement Corp. 1997. <www.odic.com>

Programme d'Analyse des Troupeaux Laitiers du Québec. 1997. <http://www.patlq.com>

Select Sires. 1997. <www.selectsires.com>

USDA 1997. <http://cos.gdb.org/best/fedfund/usda/usda.form.html>

Appendix

This appendix lists the name of the individual fields appearing in the DAIRIE prototype and the corresponding names in the underlying database. The first five tables are fields that are found in the data selection form, the final table are fields that can only be selected when 'cow' level is selected from the data aggregation form.

Production information field names in DAIRIE and in the database.

DAIRIE field name	Data Source name
Milk	MilkTotal
Fat Pct	FatPct
Protein Pct	ProteinPct
SCC	SCC
DIM	DaysInMilk
Cum Milk	CumMilk
Cum Protein	CumFat
Cum Fat	CumProtein
305d Milk	Milk305Days
305d Protein	Fat305Days
305d Fat	Protein305Days
Urea	Urea
Cow Weight	BodyWeight
Prev Lactation's Num of Breedings	NumBreedingsPrevLactation
Gestation length	DaysInGestation
Milk Price	MilkPrice
Feed Cost	FeedCost
Total Milk Price	CumMilkPrice
Total Feed Cost	CumFeedCost

Lactation information field names in DAIRIE and the database.

DAIRIE field name	Data Source name
305-day Milk Yield	Milk305Days
305-day Fat Yield	Fat305Days
305-day Protein Yield	Protein305Days
DIM	DaysInMilk
Total Milk Produced	CumMilk
Total Fat Produced	CumFat
Total Protein Produced	CumProtein
Age at Calving (months)	AgeAtCalving
Body Weight	BodyWeight
Value of Milk Produced	CumMilkPrice
Cost of Feed Fed	CumFeedCost
Number of Breedings during this Lactation	NumBreedingsPrevLactation
Number of Breedings During Previous Lactation	NumberOfBreedings

Production EBV information field names in DAIRIE and the database.

DAIRIE field name	Data Source name
EBV milk	EBVMilk
EBV fat	EBVFat
EBV Fat Pct	EBVFatPct
EBV protein	EBVProtein
EBV Prot Pct	EBVProteinPct
Milk rank	MilkRank
Fat rank	FatRank
Protein rank	ProteinRank
Production Reliability	ProductionReliability

Classification information field names in DAIRIE and the database.

DAIRIE field name	Data Source name
Final Score	FinalScore
Fr/Cap	Frame/Capacity
Stature	Stature
Front End	HeightatFrontEnd
Size	Size
Chest Width	ChestWidth
Body Depth	BodyDepth
Loin Strength	LoinStrength
Rump	Rump
Pin Setting	PinSetting
PinWidth	PinWidth
Feet/Legs	FeetLegs
Foot Angle	FootAngle
Bone Quality	BoneQuality
Rear Leg Set	SetofRearLegs
Mammary System	MamSys
Udder Depth	UdderDepth
Udder Texture	UdderTexture
Median Suspensory	MedianSuspensory
Fore Udder	ForeUdder
Fore Attachment	ForeAttachment
Fore Teat Placement	ForeTeatPlacement
Fore Teat Length	ForeTeatLength
Rear Udder	RearUdder
Rear Attach Height	RearAttachmentHeight
Rear Teat Width	RearAttachmentWidth
Rear Teat Placement	RearTeatPlacement
Dairy Character	DairyCharacter
Dairy Form	DairyForm
Defective Traits	
Wry Face	WryFace
Undesirable Head	UndesirableHead
Weak Crops	WeakCrops
Weak Back	WeakBack
Lacks Style	LacksStyle

(Continued)

Not Well Sprung	NotWellSprung
Low Loin	LowLoin
Narrow Heart	NarrowHeart
Frail	Frail
Advanced Anus	AdvancedAnus
Recessed Tailhead	RecessedTailhead
High Tailhead	HighTailHead
Advanced Tailhead	AdvancedTailhead
Wry Tail	WryTail
Thurls too Far Back	ThurlsTooFarBack
Weak Pasterns	WeakPasterns
Crampy	Crampy
Toes Out Rear	ToesOutRear
Coarse Hocks	CoarseHocks
Open Toed	OpenToed
Shallow Heel	ShallowHeel
Undesirable Stance	UndesirableStance
Lacks Udder Shape	LacksUdderShape
Tilt	Tilt
Reverse Tilt	ReverseTilt
Bulgy	Bulgy
Heavy Fore	HeavyFore
Front Unbalanced	FrontUnbalanced
Front short	FrontShort
Front Not Plumb	FrontNotPlumb
Front Webbed	FrontWebbed
Front Blind	FrontBlind
Rear Unbalanced	RearUnbalanced
Rear Short	RearShort
Rear Not Plumb	RearNotPlumb
Rear Teats too Far Back	RearTeatsTooFarBack
Rear Webbed	RearWebbed
Rear Blind	RearBlind
Close Rib	CloseRib

Conformation EBV information field names in DAIRIE and the database.

DAIRIE field name	Data Source name
Conformation	EBVConformation
Fr/Cap	EBVFrameCapacity
Rump	EBVRump
Feet/Legs	EBVFeetLegs
F Udder	EBVForeUdder
R Udder	EBVRearUdder
Mam Sys	EBVMamSys
Dairy Character	EBVDairyCharacter
Stature	EBVStature
Size	EBVSize
Fr End	EBVFrontEnd
Chest Width	EBVChestWidth
Body Depth	EBVBodyDepth
Loin Strength	EBVLoinStrength
Pin Set	EBVPinSettingDesirability
Pin Width	EBVPinWidth
Foot Angle	EBVFootAngle
Bone Qlty	EBVBoneQuality
R Leg Set	EBVRearLegsDesirability
Uddr Depth	EBVUdderDepth
Uddr Txtr	EBVUdderTexture
Med Susp	EBVMedianSuspensory
Fore Attachment	EBVForeAttachment
Fore Teat Placement	EBVFrontTeatPlacement
Fore Teat Length	EBVFrontTeatLength
Rear Attach Height	EBVRearAttachmentHeight
Rear Attach Width	EBVRearAttachmentWidth
Rear Teat Placement	EBVRearTeatPlacement
Dairy Form	EBVDairyForm

Common cow variable information field names in the software and the database.

DAIRIE (GUI)appearance	Data Source name
Registration No	RegistrationNo
DHASID	DHASID
Sire ID	SireId
Dam ID	DamId
Maternal Grand Sire	MGSireId
Birth Date	BirthDate
Test Date	TestDate
Test Day Parity No	LactationNo
Test Day Calving Date	CalvingDate
Condition Code 1	CondCode1
Cond Code 1 Date	CC1Date
Condition Code 2	CondCode2
Cond Code 2 Date	CC2Date
Test Day Last Breeding Date	LastBreedingDate
Sire ID Last Breeding	LastServiceSireRegNum
CAR Code	ExplanationCode
Test Day Date Dry	DryDate
Projected Flag	ProjTDRealFlag
Off/On Lactation	FlagOffLact
Lactation Parity No	LactationNo
Lactation Calving Date	CalvingDate
Lactation Last Breeding Date	LastBreedingDate
Lactation Dry Date	DryDate
Lactation 305d Yield Flag	Flag305DayYield
Lactation Date of 305d Yield	Date305DayYield
Previous Lactation Breeding Date	CalvingDatePrevLactation
Previous Lactation Dry Date	DryDatePrevLactation
Proof Type	ProofType
Evaluation Date	EvaluationDate
Classification Date	ClassificationDate
Classification Round Number	TCRoundNo
HAC Client Code	HACClientID