Information system for horses

Darja ČOP*, Janja URANKAR, Milena KOVAČ

Animal Science Department, Biotechnical Faculty, University of Ljubljana, Domžale, Slovenia

Abstract

The aim of our study was to develop information system (IS) for horses in Slovenia. The computer was rarely used for data storage, therefore the paper documents represented the existing IS. Development of the new IS based on system analysis that assures normalized database structure. Model file, where the database structure is described, enables the connection between the real data flow in horse breeding and computer programs. IS is divided into modules (independent units) that enable optional selection according to the necessity of the users. Modules cover identification and registration, female reproduction, young, growth (genes analyses, exterior traits, working ability), and module base that links all modules together and managing the database. Optional selection of modules gives the opportunity to cover all horse breed specifics and offer enough space to introduce new modules if needed. Due to these aspects the IS satisfies horse societies herds, as well the international organizations. Transition from the old to the new IS is possible by inserting data using screen forms that at the same time enable accessibility to the stored data in the database. IS allows data exchange among participants in electronic form, too. The current IS mostly comprehends identification and pedigree data and enables review over population. Zootechnical certificates can be issued. In the future the other modules will be filled with the data collected in current population.

Key words: horses, information system, system analysis, database structure

1 Introduction

The oldest Stud-book in Slovenia was founded for Lipizzaner horses. The stud farm Lipica was established in the year 1580 and was under the property of the Vienna Court until 1918. Nowadays, the Stud-book archive exists in libraries as state property. Pedigrees from the last years of 19th century till now are kept in Lipica. For other horse breeds, Stud-books were established and are kept at the Veterinary Faculty in Ljubljana.

The main emphasis in data collection is given to the pedigree data and less to other information. Because of small herds, breeders believe they could remember other data about their horses in their mind or at most, they write on a sheet of paper. Only part of information, especially working ability and exterior data, are appended to the zootechnical certificate. Reproduction data were kept mainly in paper forms. Thus, reproduction traits were analysed with a lot of effort and very rarely. Some breeders started to use the computer for ad-hoc applications, but not for data manipulation.The information systems in use do not assure data consistency and analysis of horse characteristics.

Nowadays, the existing information system (IS) in paper forms or as ad-hoc application is not appropriate anymore because a fast and correct information have a high value in horse breeding as well. Legislation aspects as well as breeder interest require efficient IS that enable traceability of animal, good herd management and genetic improvement.

The aim of our study was to develop computer supported information system for horses in Slovenia. We will avoid more detailed description of system analysis and will focus mainly on applications.

2 APIIS - Adaptable Platform Independent Information System

International project named APIIS (Adaptable Platform Independent Information System) started with the aim to develop framework of IS that would be independent of species, production system, recording system, country and would satisfy needs of small farms as well as larger computer centers performing genetic evaluation (Groeneveld, 1999, 2004). Slovenia was involved in the project from the beginning with the belief that the methods and procedures of data recording do not differ crucially (Drobnič and Kovač, 1993) in spite every species have its specific features.

The important aim in the APIIS is to set up efficient IS at low costs. The open source packages were chosen, thus, IS can be developed and implemented without any software costs. In order to make the information system flexible, the development was running on operation system UNIX/LINUX applying relational database PostgreSQL which is using standard structured query language (SQL). Perl with its packages is used for programming for common (Groeneveld, 1999) as well as specific applications.

3 System analysis

The first step in development of any information system is system analysis. When the generic database structure was set up for pigs, cattle, rabbits, chicken (Kovač et al., 2000). Besides generic structure, they also implemented the specific steps for system analyses of Yordanova and Groeneveld (2000). It was tested on Slovenian horses and upgraded by Čop et al. (2003) in order to assure normalized database structure. Current status of data flow was analysed and compared with the purposed general structure. The applications were mainly defined from the literature. All information were arranged into the process diagram which showed data flow as well as data usage. It also allow to define time, expected frequency, place and the user. Furthermore, entity relationship diagrams were used to eliminate dependencies and ensure normalised database. Structure of the database is described in the model file which serves as data dictionary and connects the data flow in horse breeding with the database.

3.1 Data streams and modules in information system

The first step in system analysis is collecting data streams which describe or should describe existent IS. A data stream is defined as a group of data elements recorded at one time for an object of interest. Typically, a data stream finds its physical presentation in a paper form which is used for data recording (Groeneveld et al., 2004). The problem in Slovenian horse breeding is that classical diaries for data recording are not in use. Instead of them, some kind of reports exist, which contained many data recorded at earlier stages and should be already stored on a computer. In addition, there are a few fields for collecting new data or even repeating old data. Such data flows may cause data inconsistency. For the future, it is suggested to define separate data streams which contain only the data which occur at the specific breeding activity.

Data streams circumscribe all the events, like production process, testing, etc, in horse breeding. They specify an old information to be used in the process. They are important in the process of creating applications. During the process, new data arise. It is important to know how they can be used later in decision making. The data not in use are so called black hole and should be avoided. Data recorded must be well defined. The definition tells us how and when a trait is measured. Furthermore, the domain must be specified either by interval or a list of codes. Relationship among other traits are described by business rules necessary for data validation. To form the database appropriately, it is important to know the frequency as well as the sequence of appearances during the horse lifespan. By collecting documents we know which information are important, when the data are measured, how often they are collected and used.

In the presentation, we are focusing mainly on zootechnical measure in horses. The central part of the information system is composed of module Identification and registration. It covers information about stud-book data, ownership, location, buying, selling, and culling. It also allows renaming of the horse either to get additional identification (stud-book number, name) or is renamed by any reason. All the identifications are stored in the database and limited by opening and closing dates. The second module Reproduction collects information about female fertility like matings or inseminations, pregnancy checks, and parturitions or other outcomes. Male fertility is not included in this version. Nevertheless, it is easy to add a new module covering semen collection and its usage because it is covered in other species. The module Reproduction has relations with other modules, for example module Young, where we handle data about each individual foal. In addition, foal must be marked and registered which is recorded in module Identification and registration. Module Growth and meat is used for individual measuring growth performance as well as other traits obtained usually during growth period and/or at slaughtering. The data are often used for genetic evaluation and selection of breeding stock. In horses, we have exterior and conformation traits, working ability, genetic tests, and competition results. Some of the data may be collected on adult animal as well. The last module Base is not evident to the user. It is obligated and helps to run the database, links all modules together, manage the database, provide partners information, enable coding system etc. Most applications are hidden and are processed without user intervention.

Optional selection of modules gives the opportunity to cover specifics of other breed and offer enough space to introduce new modules if needed. In some breeds, sports results are different than in Lipizzaner and must be saved in an appropriate way. Genetic module with evaluation of breeding values will become important when the Growth module will be filled with the data. Some modules have to been planed to cover semen collection, health and veterinary actions, economic, ecology, nutrition. The IS satisfies breeders, breeding organisations, computer centers for genetic evaluation, as well as the international organizations.

3.2 Database structure

According to the APIIS database structure, the entities and their attributes are divided in three groups: mandatory, frequent and optional Kovač et al. (2000). The entities translate into tables in the database while attributes represent columns in each table. Mandatory entities are in modules IDENTIFICATION AND REGISTRATION and in module BASE. Identification and traceability must be assured independently of production system. Thus, tables ANIMAL and TRANSFER are mandatory in module IDENTIFICATION AND REGISTRATION. Entity ANIMAL serves information which are constant in animal lifespan (birth date, father, mother, origin, ...). In table TRANSFER, all renamings and animal transferring are stored (Groeneveld, 1999, 2004). Table TRANSFER has in APIIS special facility for an external numbering system. That means that each animal entering the IS get internal identification number which is used in all other tables. If animal has more than one active identification, than it has all external identifications saved together with the same internal identification in table TRANSFER. All external identifications are stored only in table TRANSFER which provide the users with a tool to communicate with the database while the internal numbering helps the programer more strait forward data manipulation.The same methodology is used in module BASE for table CODES, where all external codes are gathered together, arranged in classes (sex, breed, status ...), and altered into internal codes. Partners are detailed by names (PARTNER), addresses (ADDRESS) and their roles in system (JOB) as breeders, horse owners, veterinarians, breeding specialists, ... and maybe, even friends. The internal numbering makes IS more efficient, data are consistent and we avoid errors. However, the user does not realize the transformation because she/he uses her/his preferred external identifications to type the data into the system or to get applications running.

Mandatory entities and attributes are the same for each species, production system, recording system, country. Frequent entities and attributes could be specific for species. In horse breeding, the reproduction is similar to the other species. A mare is mated very often during breeding period and is pregnant longer than a sow or a cow. Nevertheless, these facts does not influence the database structure, although business rules must be regulated (Groeneveld and Kovač, 2000). Interval between matings could be less than 21 days, interval between the last mating and parturition will be more than nine month. Moreover, we expect only one or at most two foals which is, of course, less than in goats, sheep or pigs but close to cattle situation. Thus, we adopt the module REPRODUCTION in cattle and change the business rules adequately. Module GROWTH AND MEAT (tables EXTERIOR, MEASUREMENTS, TEST, GEN_SAMPLE, GEN_ANALYSE) differ more. It is still useful to apply mandatory structure of the module and add all the specific information.

Data structure must be normalized ensuring each data is stored only in one place. Although modules are declared to be independent units, any borders do not exist between them. Contrary, certain applications are using the data from many sources even from different modules. Normalized database structure does not allow the values extracted from other data to be stored into the database. For example, we should not store age at parturition but rather birth and parturition date.

3.3 Applications

The applications are grouped according to the purpose they are used. The first group of applications is used to extract important information from existing records or typing the data from documents using screen forms (Figure 1). Although the application is needed only at the time of transition to a new system it may be used more often when the data need some editing. In our case, the historic data in electronic form were rare, thus the data were inserted through screen forms which communicate with the database.

The regular forms are usually different from the initial forms and depend on the data streams and are used for inserting new data into the database. The data streams and tables are connected via object that validates the data and stores them into adequate tables. The same object may be used with forms as well as loading routines. Loading routines are set of programs to be used in any APIIS information system. Filling the database with the data is driven by the model file containing information about database, database structure (tables, attributes names), and business rules. The forms and loading routines must suit the real data flow in horse breeding and allow to insert only new data. Almost the same forms may be used to browse the data which are already stored. These are the simplest applications to see the data on the screen.

The IS for horses covers issuing zootechnical certificates prescribed by breeding organisation and zootechnical legislations. Data from the Stud-book and the lists of new horses were inserted by using using screen forms in figure 1. In transition time, some additional data are needed to identify pedigree animals, especially parents, marked by old identification systems. As soon as they do not appear as parents any more, the application can be simplified using only official identification. The system requires that the user take care only about the parents of the animal, the older generations appear automatically from ancestor records.

The other applications are still under development because the data are not collected yet. It must be decided what to do with historic data. The decision should depend on the data quality. If the data can be used it is worthwhile

to make an effort to insert data from the last years. On the other hand, the routine data streams must be defined for new data and bring them into practice. The applications tested on virtual data proved to be useless when applied to the real data. Monitoring basic statistics and phenotypes changes as well as genetic evaluation can start as soon as the data are accessible. Nowadays, we focus on the analysis of population structure and management tools of small population.

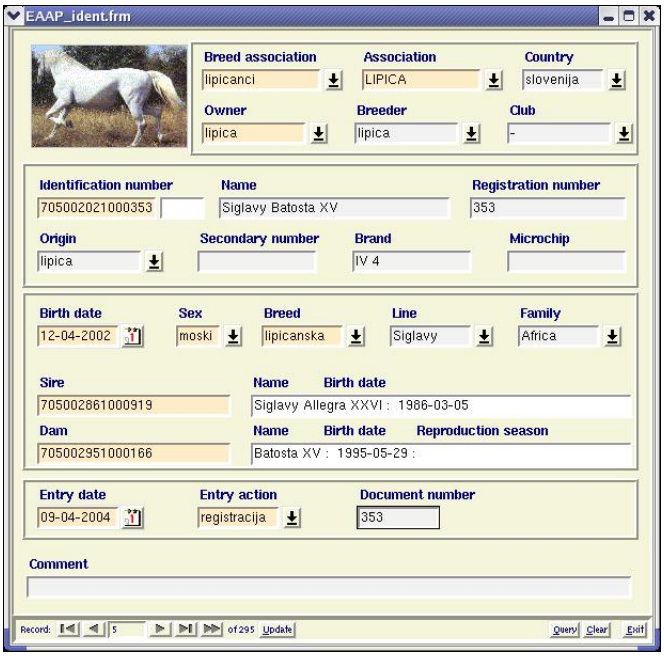

Figure 1: Screen form for inserting and accessing identification data in/from database

4 Conclusions

We are convinced that computer supported information system will be useful in horse breeding as it is in other species. The smaller is the population the more care must be dedicated to preserve or even extend the population. The stud-book keeping is the nuisance activity but it is not enough for successful horse breeding regardless of the purpose horses are used for. With computer assisted IS we can gain on data accuracy and consistency, shorten response time between event and action, have better overview over the population. The IS can be applied to other breeds as well.

References

Čop D., Kovač M., Brodnik A., Malovrh Š. 2003. System analysis adapted for livestock production. In: Proceedings of Genetika 2003. Ljubljana: Genetics Society of Slovenia, Bled, 2003-05-31/06-04, pp. 239–240, Bled, Slovenia.

Drobnič M., Kovač M. 1993. The genetic improvement module in an information system for animal production. Zb. Bioteh. Fak. Univ. Ljub. Kmet. (Zoot.), 62: 107–113.

Groeneveld E. 1999. Design of a protable platform independent pig information system: An internet based development project. In: Role and potential of IT, Intranet and Internet for advisory services, Second European Federation for Information Technology in Agriculture, Food and the Enviroment, Rickert U., Helbig R., Schiefer G. (Eds.), Bonn.

Groeneveld E. 2004. An adaptable platform independent information system in animal production: framework and generic database structure. Livest. Prod. Sci., 87: 1–12.

Groeneveld E., Kovač M. 2000. Implementation of non procedural business rules in APIIS. In: Vortragstagung der DGfZ. e.v. und der DGfT, Aus der Arbeit der Forschungsstatten fur Tierproduktion, p. A27, Kiel. Institut fur Tierzucht end Tierhaltung der C-A-Universitat.

Groeneveld E., Yordanova L., Hiemstra S.J. 2004. Organizational structure and information technological support of national gene banks. Livest. Prod. Sci. (article in press).

Kovač M., Groeneveld E., Malovrh Š. 2000. Comparison of data structure for various production systems in livestock and its application in APIIS. In: Vortragstagung der DGfZ. e.v. und der DGfT, Aus der Arbeit der Forschungsstatten fur Tierproduktion, p. A29, Kiel. Institut fur Tierzucht end Tierhaltung der C-A-Universitat.

Yordanova L., Groeneveld E. 2000. The implementation strategy of information system with existing data streams. In: Vortragstagung der DGfZ. e.v. und der DGfT, Aus der Arbeit der Forschungsstatten fur Tierproduktion, p. A28, Kiel. Institut fur Tierzucht end Tierhaltung der C-A-Universitat.