"No caring without data sharing": an ontological approach to improve the data management system of a research animal facility in a public veterinary institute

L. Contalbrigo[1*], S. Borgo[2], G. Pozza[1] and L. Ravarotto[1]
1. Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro (PD), Italy
2. Laboratory for Applied Ontology (LOA), Institute of Cognitive Sciences and Technologies, CNR, Trento, Italy
* ***e-mail*** *lcontalbrigo@izsvenezie.it*

Introduction

Whether it is used for scientific or management purposes, data is the basis of every research study, production program or development plan in every type of organisation including research animal facilities and vivarium. Both reliability of collected data and the skill in selecting, recording, storing and in keeping them usable in time are essential in every decision making process. The need of daily managing a large amount of information, produced with several purposes, draws often research organizations to adopt many software systems which support the different phases of their working procedures e.g. software solutions for animal and material ordering, animal breeding and health management, environmental monitoring; application like LIMS (Laboratory Information Management System) for laboratory results or statistical analysis or even the use of Electronic Lab Notebooks. Frequently, these systems can not reliably interface each others, safely share data and uniformly integrate them; besides, their use often evolve within the organization without a consistent control plan. In a complex scenario, like a research animal facility or a vivarium, it is difficult to clearly understand the real requirements of the information management and consequently to build an efficient, shared and integrated data management system. Applied ontology may be an effective instrument to reach these aims. Applied ontology is the area resulting from merging ontological considerations and implementation techniques and is used in several domains (e.g. biomedicine, anatomy, genomics, information science). It focuses on basic categories, their hierarchy and their relations which are relevant to a specific scenario. Ontologies themselves are artifacts built to model a domain. Among these, formal ontologies are given as systems of axioms in a formal language[1], while computational ontologies are formalized in a language suitable for efficient computation[2]. Foundational ontologies (formal and/or computational) are ontologies focusing primarily on general and domain-indipendent categories[3]. Nowadays, the ontological techniques developed in the last twenty years are exploited by companies and public institution in many areas like data warehouse and intelligent interfaces as well as in knowledge modelling, organization and exploitation. The roles played by ontologies in biomedicine may be classified in three categories: knowledge management (indexing and retrieval of data and information, access to information, mapping among ontologies), data integration (exchange and interoperability); decision support and reasoning (data selection and aggregation, decision making process support)[4].

Aim

We aim to build a theoretical framework for the management of the research animal facility at the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe). The main goal is to identify which entities are involved in the different processes, which properties they have and which roles they play. This study helps to clarify the general perspective of the IZSVe research animal facility and will set the basis for the development of an integrated data management system.

Materials and Methods

Tools of the ontological analysis

The complex scenario of a research animal facility needs to be analyzed using ontological techniques coming from formal ontology to achieve the construction of a general framework in which all perspectives involved are included in a natural way. The foundational ontology DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*) is developed starting from commonsense considerations with a particular emphasis on social reality because categories introduced by DOLCE are thought as human cognitive artifacts. Note that DOLCE applies a multiplicative approach: entities are identified by clusters of properties so that different entities can share the same spatio-temporal region[5]. Our use of DOLCE focuses on the relationship among objects, events, qualities, descriptions and roles which are exploited to provide the right environment for data interpretation and their integration. DOLCE takes as fundamental categories those of objects, events, qualities and abstracts. Subcategories like those of descriptions and of roles are central for data interpretation and integration. Within this system, we introduce a notion of "knowledge object"[6] to formalize the interconnections between material entities, information and organization roles within the development of the IZSVe processes.

At last, we consider the notion of use context. This concept is of primary relevance and largely used in applications like mobile communication, instant planning or action detection[7]. Here we look at it to talk about functionality of a device in a given scenario. The notion of context helps to identify the relationships between the device, the user, the user's goal and the general conditions in the scenario[8].

The Scenario

The Istituto Zooprofilattico Sperimentale delle Venezie (IZSVE) is a public veterinary institute that conducts both laboratory controls and research activities in three main areas: animal health and welfare, food safety and environmental protection.

This institute has a conventional research animal facility and a bio-containment unit (BSL3) in which a number of both research and diagnostic protocols involving different species can run simultaneously (e.g. infectious disease studies, vaccine quality controls, biological tests). The conventional facility houses rodents (mice, rats, hamsters, guinea pigs and voles), rabbits and poultry whereas swine and ferrets are housed in the bio-containment unit. The facility is authorized by the Italian Ministry of Health and so it follows the law requirements of the D.Lgs n° 116/92. A large amount of information is produced everyday from each experimental activities and each procedure applied. Data are collected in four databases managed by the respective software and then in a number of paper registers sometimes turned into Excel files.

The databases included a LIMS called Izilab, where laboratory results of samples collected from animals enrolled in experimental procedures and diagnostic tests are recorded; a TAC Xenta system which controls, regulates and records environmental parameters e.g. temperature and pressures; a software solution for animals and materials ordering (NFS) and a specific software to record non-conformity, preventive and corrective actions as required by the quality management system of the facility. All information about internal procedures are codified by the quality management system and shared thanks to a file server. All other data are recorded in paper registers sometimes turned into Excel files. Therefore, data are fragmented and scattered: this prevents their organization to guarantee an effective support to decision making processes.

Results and Discussion

We provide a diagram of the elements involved in the scenario and their classification with respect to DOLCE categories (Figure 1).

Figure 1. *Diagram of the entities involved in the scenario*

Data Generation and Management. We can distinguish two subcategories of data. One includes information expected by experimental protocols, called "research data", the other one is called "care data" and composed by sets of information which are not directly functional to the research project e.g. environmental data, management data, animal health and behavioural data, enrichment program validation, animal training report. Nowadays, in the IZSVe research animal facility, the two subcategories of data follow different paths and are not efficiently shared. This is a substantial limit in IZSVe knowledge management because care data should be effectively integrated with research ones since they can influence final interpretation of experimental results. A third subcategory of data is represented by quality assurance data; it may be seen as a part of research data and in our framework the software used to record this data subcategory could be used to integrate critical care information (e.g. environmental unsuitable data, internal audit report, failed laboratory results of the health monitoring program, etc.) with research one. However the information flow is not enough coherent and usable to answer both practical reasons and the new law requirements (Directive 2010/63/EU), for example number of animals enrolled in each experiment and much information about experimental procedures are recorded in paper working sheets and sometimes they are translated in Excel files so that data can not be easily shared into the organization and in case of inspection they can not be quickly usable.

Knowledge object description. We argue that the heart of our ontological design is the "the laboratory animal" (LA). In this project we are trying to extend the notion of knowledge object[6] to include also emotional and ethical aspects to model the animal in its integrity. The knowledge object is a combination of elements coming from disjoints categories as the result of the interdependencies among three elements at a certain time (t): a material entity (M), an information entity (D) and an agent role (A). In our scenario, M is any physical matter coming from the knowledge object (e.g. blood samples, swabs, organs and tissues), D collects all categories of data sourcing from the knowledge object, a sort of tuple in a relational database in a given time (t); A are internal roles in the IZSVe research animal facility and are played by the IZSVe staff involved in research procedures, quality assurance and LA care.

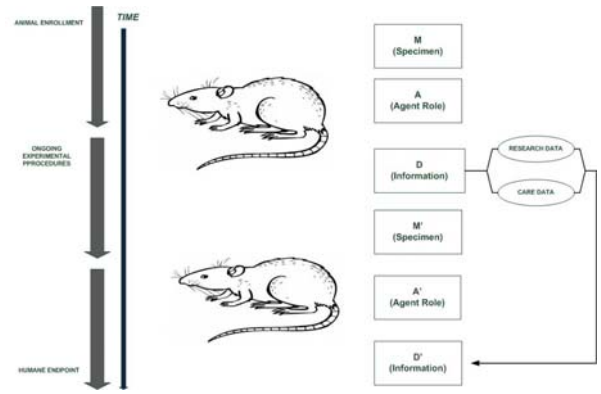

Figure 2. *The Knowledge Object in our scenario*

A use context for effective data sharing. Generally speaking, data received by a user can lead to multiple effects: the user revises its understanding of a situation according to these data and data drive its decision making process. In our scenario agent and users of information are sometimes the same. A correct distinction and classification of the user according to its goals and perspectives is mandatory to distribute data in an effective way and with different granularity. The use of data coordination among peers and mapping tables as suggested by Arenas *et al.*[9] may be useful to obtain data sharing and data integration among the organisation.

Conclusion

The ontological framework, sketched in this study, helps to relate the data about the knowledge object with the data characterizing the requirements for the research. This result, enriched with a notion of use context, is obtained by identifying the "knowledge object" as the focus of the IZSVe research activity and by intertwining the research and the facility procedures to set the basis of the development of a possible effective integrated data management system with a distribution of data to the users with different granularity.

References

1. N. Guarino (1998). Formal ontology in information system. In Proceedings of the Second International Conference on Formal Ontology in Information System, 1998, pp. 3-15. IOS Press.
2. T. R. Gruber (1993). Model formulation as a problem-solving task: computer-assisted engineering modeling. In Knowledge Acquisition as Modeling, 1993, pp. 105-127, John Wiley & Sons, Inc. New York, NY, USA.
3. C. Keet (2011). The use of foundational ontologies in ontology development: an empirical assessment. In The Semantic Web: Research and Applications, LNCS 6643, pp. 321-335.
4. O. Bodenreider (2008). Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearb Med Inform, pp. 67-79.
5. S. Borgo, C. Masolo, (2009). Foundational choices in DOLCE. In Handbook on ontologies (Second Edition), Springer Verlag, pp.361-382.
6. G. Pozza, S. Borgo, L. Ravarotto. From data to knowledge objects, ontological considerations with inputs from the public health domain. In Proceedings of the 10thEuropean Conference on knowledge management, pp. 635-641.
7. T. Grill, M. Tscheligi (2011). Towards a multi-perspectival approach of describing context. In CONTEXT 2011, LNAI 6967, pp. 115-118.
8. R. Mizoguchi, Y. Kitamura, S. Borgo (2012). Towards a unified definition of function. In International Conference on Formal Ontology in information system. FAIA 239, IOS Press, pp. 103-116.
9. M. Arenas, V. Kantere, A. Kementsietsidis, I. Kiringa, R. J. Miller, J. Mylopoulos (2003). The Hyperion Project: from data integration to data coordination. ACM SIGMOD Record 32(3), pp. 53-58.