Weblabs of Brazilian Bees

Marcelo S J Ferreira

University of São Paulo, Agricultural Automation Laboratory – Polytechnic School
São Paulo, Brazil, 05508-900
msucci@usp.br

Teresa C Giannini
University of São Paulo . Bees Laboratory– Biosciences Institute
São Paulo, Brazil, 05508-900
giannini@usp.br

Pedro L P Corrêa
Antônio M Saraiva
University of São Paulo, Agricultural Automation Laboratory – Polytechnic School
São Paulo, Brazil, 05508-900
{pedro.correa , antonio.saraiva@poli.usp.br}

Abstract

This paper presents an information system on stingless bees (Meliponini) of Brazil. Due to their environmental importance, immediate attention is necessary to increase the knowledge about their biological role, particularly in pollination, conservation and their breeding and preservation. An information system has been developed to organize data and produce knowledge. A database is a central part of this system, storing data collected automatically by data loggers and by a weather station in experiments on thermoregulation and flight activity of colonies. By using WebLabs, experiments can be conducted via a high speed network where video, sound and other data from colonies are available to a remote researcher. Weblabs are accessed through a Portal, according to an architecture based in the SOA (Service Oriented Architecture) paradigm. Data and metadata collected from experiments are recorded in EML (Ecological Metadata Language) format allowing integration with other systems like SEEK(Science Environment for Ecological Knowledge)-Ecogrid. The system architecture is presented and its application is discussed.

Keywords: Biodiversity Informatics, EML, Weblabs.

1 Introduction

Performing experiments in laboratories is essential in many research activities. Nevertheless, costs related to the acquisition and set up of the infrastructure can be an obstacle to many institutions. An approach to face this problem is the use of simulators, which are computer programs designed to replicate, within some extension, the behaviour observed in the real world. Simulation, however, is limited to the complexity of the system to be simulated, and can be used only in a experiment that could be mathematically modelled *a priori*. This limits the application of this method in more complexes experiments, such as experimentation using living organisms.

A alternative is the use of laboratories to accomplish experiments remotely, through Internet. Those laboratories have been called Weblabs [1] where equipments of a laboratory are connected to a network – Internet, in a broader case – in such a way that experiment conditions can be modified and resulting data can be collected

Some advantages related to the use of weblabs are listed in [8], as follows:

- A better use of laboratory infra structure (available all time since the weblab was operational);
- A better organization of time to projects' accomplishment;
- To make possible to offer distance courses using the complete laboratory infra structure.

ViNCES (Virtual Network Center of Ecosystem Services) is a laboratory research consortium which focuses on ecosystem services, and is part of the Tidia-Kyatera project [18].The viability of a network of virtual laboratories as proposed by ViNCES demands a uniform access to services offered by each of these laboratories.

A web portal is a form of associating web sites that are different, but still related. When there are many weblabs in a same area of knowledge, a portal could be created to integrate these laboratories, providing an uniform and centralized access to their resources. The Service Oriented Architecture paradigm is an alternative to approach ViNCES proposals, since it defines a model to integrate portal and services providers [13].

2 The Infra-structure of the Bee Laboratory

Activities at the ViNCES project include pollination researches. These particular studies use stingless bees as the main pollinator. Stingless bees belong to Meliponini tribe,are eusocial and, so far, 192 species have been described [16]. They have an important role as pollinator of forest fragments and agricultural crops, and had been used in programs of sustainable development.

The information about bees has been stored in an information system that presents two kinds of database: static and dynamic data.

The first kind of data stores information related with stingless bee diversity [15]. These data are related with important plants used by bees as source of food (pollen, nectar, or both) or as nesting site with photos showing the inflorescence, the plant species appearance and the pollen grain. There are photos (300x400 pixels) that aim to characterize different species, showing details of females, males and queens. Other ones show the interior and the entrances of hive that are built by bees and present peculiarities related to different species. Additional information is available, such as places - geographical coordinates included - where the bees have been collected and climate data such as temperature and humidity at the moment of capture. All these data can be accessed through a "bee card" and a "plant card" .

Through ViNCES project, photos of high resolution (2560x1920pixels) and videos (640x480pixels) that aims to offer more detailed morphological characteristic of different species and to show the behavior of bees inside the hives are been included in these cards (Figure 1).

Figure 1. One "bee card" accessed through ViNCES portal

The second kind of information is related to real time data acquired through a weather station located at the experimental place and a device that is monitoring a bee hive of *Melipona quadrisfaciata anthidioides*. This device includes a video camera and a microphone and aims to relate the weather conditions with the behaviors showed by bees inside the hive. The microphone aims to record the sounds produced because this seems to play an important role in the communication of these bees, since it is used for coding information about distance [10] and height of floral resources [12]. It will be included also two LON equipments (Local Operating Network). One of them will count the number of bees that enter and leave the hive during the day and the other will record the temperature and humidity inside the hive (Figure 2).

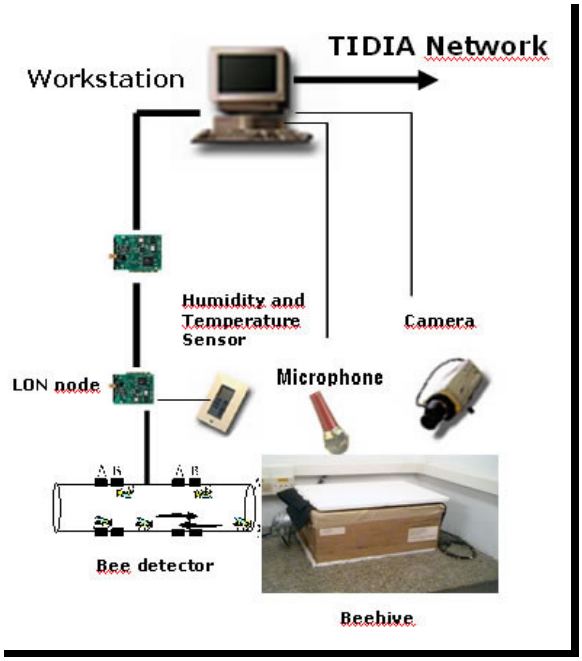

Figure 2. The monitoring system used at a hive of *Melipona quadrifasciata anthidiodes* by the ViNCES project

3. System architecture

Portal-Provider Weblab Architecture

The architecture includes three components: experiments, weblabs and portal. The experiments are executed through the weblabs. External users access the portal which communicates with weblabs. This relationship is represented in Figure 3.

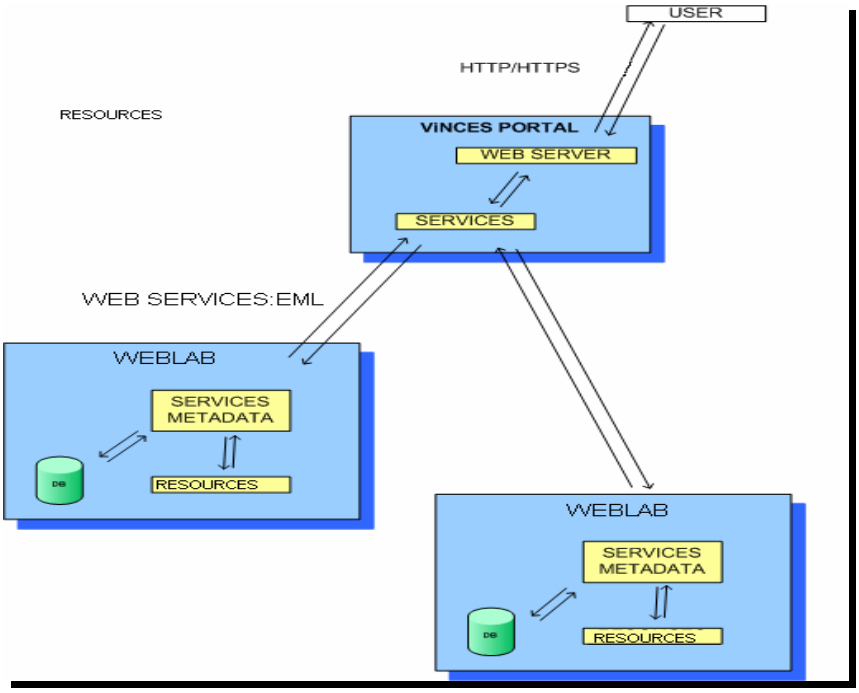

Figure 3. The integration of weblabs and portal

In the context of weblabs, a generic experiment would involve the following types of data:

- Data of the actors of the experiment (data about the specie(s) and specimens or group of specimens);
- Data about the conditions that they are submitted (ambient conditions, e.g. temperature);
- Data about the behavior of the actors in the conditions (e.g. video, sounds)

The fourth type of data refers to cadastre data of the experiment, among them, information of the responsible researcher, eventual descriptions and notations of the experiment. Data relative of the access control to the experiment also must be enclosed here.

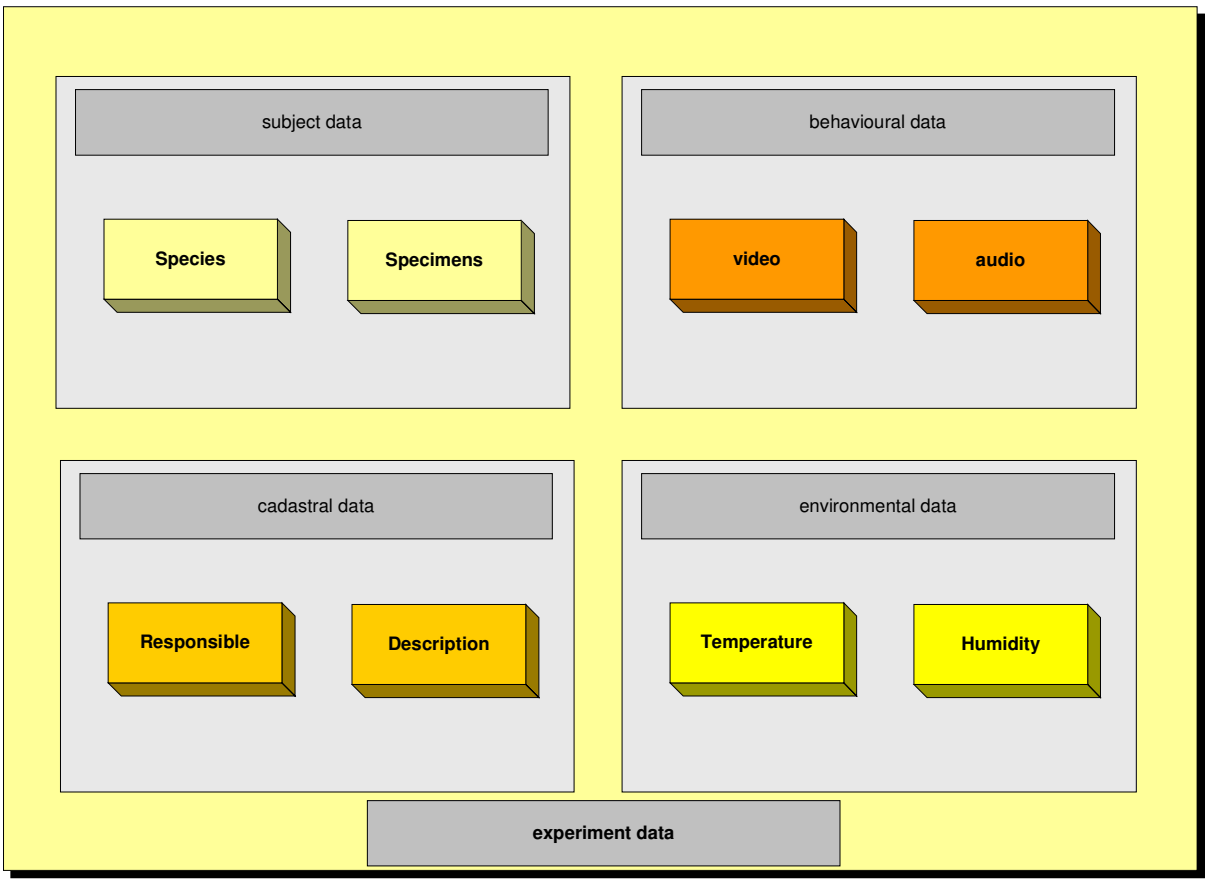

Figure 4 shows the model of data used to represent the experiments:

Each weblab includes one or more equipments, which are responsible for collecting data experiment (sensors) and also for setting desired environmental conditions (actuators) for the accomplishment of the experiences. Weblab also must make the communication with the portal when necessary, executing requested experiments from the portal. As these experiments are carry through, its data must be available for recovery from order made through the portal.

One of the non-functional requirements that must be considered in this operation refers to the access of the system and the data registered. Questions as "who can carry through the experiments" or "who has access to the data of an already registered experiment" are treated by control access lists provided by the responsible of weblab. More of this subject is described in the section

In this way, a data model for a weblab supposes:

- A database that lists available equipment for a weblab and its characteristics. A paradigm for this database is the MIB (Management Information Base), used in network management systems (Case et al, 1990). By analogy, this database is called WMIB (Weblab MIB);
- A control access list to the equipment.
- A database of current and previous data experiments for each weblab;

The accesses to data registered for the equipment and also its configurations are made by services which make possible operations similar to those performed by the commands *get* and *set* of SNMP (Simple Network Management Protocol) [6]. If requests originated from portal are in accordance with the permissions granted by the weblab manager, equipment can be set up and data collected and stored in the dabase for later access. Figure 5 synthesizes the model of data and considered services of a weblab.

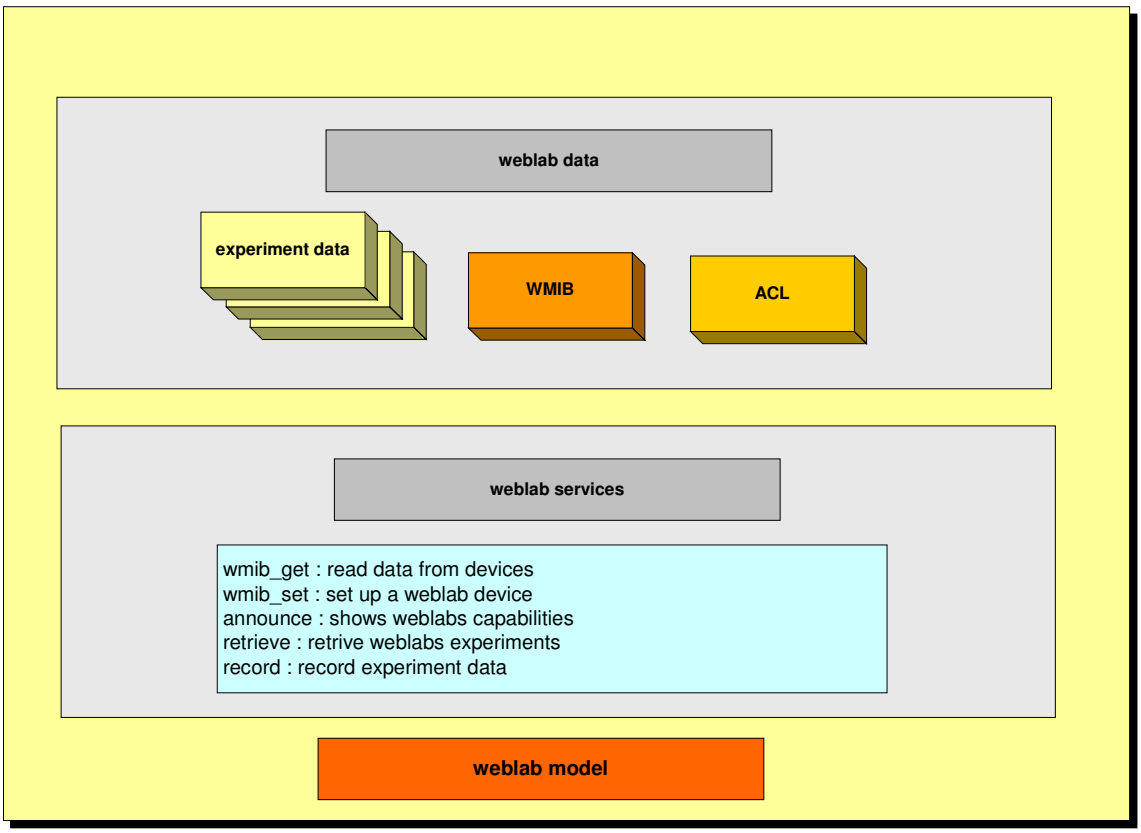

Figure 5. Weblab data model

As mentioned before, a list of equipment and correspondent variables should be available. Table 1 shows example parameters considered in a webeelab WMIB.

Table 1. Example Weblab MIB extract

Instrument	Thermometer	Meteorological station					Counter	Counter
Variable	Internal temperature	External temperatura	Humidity	Luminosity	Time	Date	Bee entries	Bee exits
Unity	°C	°C	%	LUX	hh:mm:ss	dd/mm/aa	individuals	individuals
Collect interval	1 min				15m	daily		

A more complete model can be used for weblabs when security requirements are considered. As presented in [7] access to weblabs and experiments data can be controlled by considering extra elements in order to provide AAA (Authentication, Authorization e Accountability) services. Figure 6 presents the weblab model when security requirements are considered.

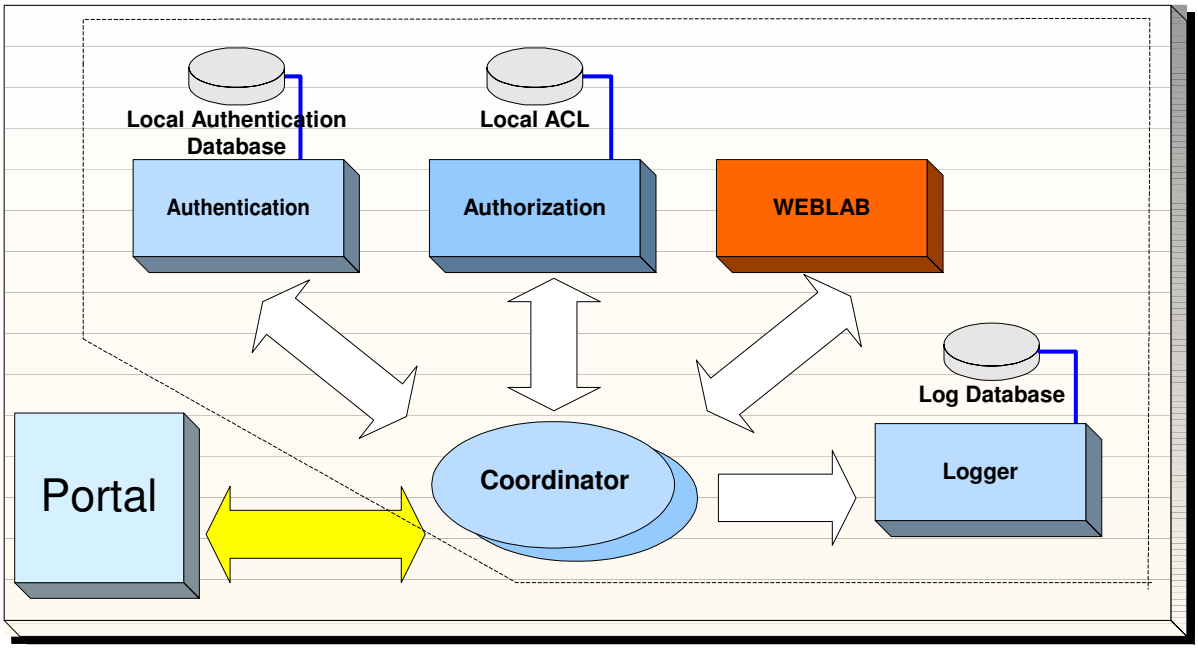

Figure 6. Weblab model considering security requirements. Adapted from Ferreira et al. (2006)

Data collect in a experiment is stored in EML (Ecological Metadata Language) format. EML is a metadata standard based on prior work done by the Ecological Society of America and associated efforts [11]. EML is implemented as a series of XML document types that can be used in a modular and extensible manner to document ecological data. Each EML module is designed to describe one logical part of the total metadata that should be included with any ecological dataset.

4. General Discussion and Conclusions

The risks to agriculture and natural ecosystems involving reduction of pollinators were recognized by the Global Environmental Facility (GEF) and Food Agriculture Organization (FAO) of United Nations that approved the project *Conservation and Management of Pollinators for Sustainable Agriculture through an Ecosystem Approach* . The project presented four components: to develop one base of knowledge; to extend and promote best management practices; capacity building; and awareness, mainstreaming and results dissemination. To accomplish these objectives is necessary developing a new approach to deal with biological data.

The information technology plays an important role in the area of basic science. There is a lot of information about pollinators disperse in different kinds of formats and sources, including biological collections at museums or academicals institutions, unpublished sheets held by researchers, and published papers that are difficult to find. This information is spread and can't be accessed easily. It is necessary to catalogue, integrate and share it in order to perform new analyses and to acquire better results. There is an urgency to provide accurate suggestions related with best practices to local biodiversity issues, species and/or communities management and conservation and in order to do this is necessary a better base of knowledge.

Also, the field work could be beneficiated by information technology through the use of instrumentation and automation. The automatic acquisition of data made by devices developed to this end could provide the increasing of standardized data with better quality and accuracy, demanding less human effort. It could also stimulates the interaction between researchers groups through the accomplishment of integrated experiments or involving shared data, the union of efforts to build a common base of knowledge, and to spread the knowledge to different sectors of society [14].

Recent developments in information and communication technology, the availability of large-scale data resources, and the necessity of carrying out interdisciplinary and inter-regional research make the development of information systems concerned with data management, preservation, archiving, and dissemination a priority. Interoperable distributed information systems and the use of different data sources

that must be integrated or compared demand a common international approach to data and information management. Common protocols and standards must be developed cooperatively. Resources must be shared and long-term financial support for data providers and data management centers must be guaranteed. The success of initiatives such as GBIF and IABIN that promote cooperation and collaboration will most certainly help countries and institutions in managing and sharing their data and benefiting from data managed and shared by others [4, 5].

Nowadays in Brazil there are difficulties to be solved in the area of development and application of information systems to biological studies [3]. There is a gap between the developers of information technology and the research groups that deal with pollinators. This could be a result of a lack of information about information technology inside the community that deals with basic science. Few researchers have access to tools to data acquisition, to data control and analyzes, and modeling. But some tools are been suggested and involve systems and equipments to massive acquisition of data [9], databases used to perform meta-analyses [2] and modeling [17].

The system architecture suggested in this work has been developed and addresses some difficulties mentioned above. Initial tests indicates that it is flexible enough to be applied in other studies and work has been done in integrating completely the instrumentation network into this architecture, task that is simplified by the use of the SNMP paradigm, which isolates instrumentation complexity from the rest of the system. Another point that should be mentioned is that the choice of EML as format for data experiment allows integrations with other ecological initiatives, in special the Science Environment for Ecological Knowledge (SEEK), what could make these kind of data much more available for researchers worldwide.

Acknowledgments

The authors would like to thank *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP) for supporting ViNCES project, grant number 03/08134-4.

References

[1] Alamo,J.d.et al. 2002 .The MIT Microelectronics WebLab: a Web- Enabled Remote Laboratory for Microelectronic Device Characterization, World Congress on Networked Learning in a Global Environment.

[2] Biesmeijer, J.C.; Slaa, J.; Castro, M.S.; Viana, B. F.; Kleinert, A. M. P. and Imperatriz-Fonseca, V.L. 2005. Connectance of Brazilian social bee-food plant networks is influenced by habitat, but not latitude, altitude or network size. Biota Neotropica 5 (1).

[3] Canhos, D.A.L., Canhos, V.P. & Souza, S. 2005. Coleções Biológicas e Sistemas de Informação. In: Diretrizes e Estratégias para a Modernização de Coleções Biológicas Brasileiras e a Consolidação de Sistemas Integrados de Informação sobre Biodiversidade. Centro de Gestão e Estudos Estratégicos (CGEE) & Ministério de Ciência e Tecnologia.

[4] Canhos, V.P.; Souza, S. and Canhos, D.A.L. 2003. Report. Building the Inter-American Biodiversity Information Network (IABIN). Sub-region 2 – Brazil. http://www.iabin-us.org/.

[5] Canhos, V.P.; Souza, S. and Canhos, D.A.L. 2004. Construindo a Rede Interamericana de Informação sobre Biodiversidade - IABIN, Sub-região 2 - Brasil, Fase II. Relatório para a Organização dos Estados Americanos. Julho. http://www.iabin.net/.

[6] Case,J.D;Fedor,M.S ; Schoff- Stall, M.L and Davin, J.R. 1990.A Simple Network Management Protocol (SNMP). IETF , Request for Comments,1157

[7] Ferreira, M.S.J; Corrêa, P.L.P aAnd Saraiva, A.M. 2006. A Security Architecture for Sharing Distributed Biodiversity Databases – 4th World Congress on Computers in Agriculture – WCCA.

[8] Garcia-Zubia, J. and del Moral, A. 2005. Suitability and Implementation of a WebLab in Engineering In: IEEE Conference on in Emerging Technologies and Factory Automation. ETFA 2: 49-56.

[9] Hilário, S. D. 2005. Atividade de vôo e termorregulação de Plebeia remota (Holmberg, 1903) (Hymenoptera, Apidae, Meliponini). IBUSP. S. Paulo, 124 p. Tese de Doutorado.

[10] Michener, C. D. 1974. The Social Behavior of the Bees. Cambridge, Belknap.

[11] Michener WK. 2000. Transforming data into information and knowledge. Pages 142–161 in Michener WK, Brunt JW, eds. Ecological Data: Design, Management and Processing. Malden (MA): Blackwell Science.

[12] Nieh, J. C. 2004. Recruitment communication in stingless bees (Hymenoptera, Apidae, Meliponini). Apidologie 35: 159-182.

[13] Papazoglou, M.P. and Georgakopoulos, D. 2003. Service-Oriented Computing In Communications of the ACM. 46 (10). Out 2003

[14] Saraiva, A.M. and Giannini, T.C. 2005. Tecnologia da Informação na Iniciativa Brasileira de Polinizadores: um estudo. Relatório. Produzido para a Iniciativa Brasileira dos Polinizadores com suporte da Organização das Nações Unidas para Alimentação e Agricultura (FAO/GEF).

[15] Saraiva, A.M.; Imperatriz-Fonseca, V.L.; Cunha, R.S. and Cartolano Junior, E. 2003. Webbee – a web based information network on bees. Revista de Engenharia de Computação e Sistemas Digitais. São Paulo 1: 77-86.

[16] Silveira, F. A.; Melo, G. A. R. and Almeida, E. A. B. 2002 (a). Abelhas Brasileiras. Belo Horizonte, Minas Gerais, 253 p.

[17] Siqueira, M.F. and Peterson, A.T. 2003. Consequences of global climate change for geographic distributions of cerrado tree species. Biota Neotropica, 3(2).

[18] TIDIA. Tecnologia da Informação no Desenvolvimento da Tecnologia Avançada.: http://www.tidia.fapesp.br/portal/I.projetos/kyaTera