The $\mathcal{IKBS}$-DARWIN Research & Learning System for Sharing Biodiversity Knowledge of Tropical Islands

N. Conruyt[1] and **Y. Geynet**[1] and **D. Grosser**[1] and **A. Senteni**[2] and **M. Santally**[2]

Abstract.

This paper describes an emergent collaborative project between the universities of Reunion (Institut de Recherches en Mathématiques et Informatique Appliquées - IREMIA) and Mauritius (Virtual Centre for Innovative Learning Technologies - VCILT) to make scientific research results in insular tropical environment available to a wider audience, create an awareness and contribute to a better protection of this environment. On the long term, the project also aims at developing a strong regional competency kernel through the use of technology as a tool to encourage the development, documentation and use of indigenous information and knowledge about tropical islands environment: evaluate what is available, how it can be preserved, strengthened, amplified or improved through the introduction of ICT and how it could participate, in the basket of globalisation, as our share for a common identity. We show how enhancement and dissemination of research results can be achieved by a combination of an upstream strategy based on $\mathcal{IKBS}$, a generic software for building knowledge bases on the internet in natural sciences, and a downstream one, based on a pedagogical approach borrowed from the DARWIN world. The final outcome expected for the overall $\mathcal{IKBS}$-DARWIN project is a transversal bilingual (French - English) research and educational portal, supporting enhancement online activities in several related domains at several levels of competency, from researchers and graduate students to undergraduate students, associations or communities of practice.

keywords : biodiversity, e-research and learning, knowledge building, enhancement and dissemination of research, insular tropical environment.

1 Introduction

The introduction of Teaching and Learning Technology (TLT) in Higher Education changes in a drastic manner the relationship between the two major processes on which relies the elaboration of knowledge in the university: Research on the one hand, Teaching & Learning on the other. By providing access to an endless variety of information and personal contacts, the internet introduces in traditional education frameworks, the need to narrow the gap between the production of knowledge, so far related to research, and its dissemination, so far related to learning. In the Information Society, Higher Education institutions build a reputation on their ability to respond to an increasing demand for true, reliable, illustrated, updated and easily accessible knowledge. Thus, the cycle gets shorter from research to teaching, while students are confronted to an ever-evolving, fluctuant, controversial information out of which meaning and knowledge can be built up mainly in contact with those who produce it. One such knowledge flow needs to be supported by tools and pedagogical scenarios in which students not only have access to information or expert knowledge, but also are actively involved in its elaboration at the same level as the experts themselves.

2 Sharing biodiversity knowledge of tropical islands

The issue of a biodiversity information system is to improve knowledge about beings living in tropical islands, in order to preserve and manage this terrestrial and marine patrimony. Citizens in the Mascarene archipelago (SW of Indian Ocean) must be aware of growing degradations in tropical forests and coral reefs, inherited from economic and anthropic pressures (tourism, urbanism, agricultural and fishing developments). In this context, Universities play a major part for the development and transmission of knowledge about tropical biodiversity, natural habitats, and customs. This situation calls for narrowing the gap between researchers, decision-makers and users.

Our objective is to work on shared diagnosis exercises for facilitating the appropriation of knowledge by the island people. Thus, the $\mathcal{IKBS}$-DARWIN portal contribution aims at bringing some expertise accessible for different target users wanting a long-standing development of their environment. More precisely, it makes scientific research on species in tropical islands available to a wider audience, in order to create a consciousness contributing to a better protection of these natural resources.

3 The $\mathcal{IKBS}$-DARWIN research & learning system

A regional cooperation project for the development of research & learning contents and facilities in the Indian Ocean region, is currently emerging between the universities of Reunion and Mauritius, through the IREMIA lab (Research Institute in Applied Maths and Computer Science in Reunion Island) and the VCILT (Virtual Centre for Innovative Learning Technologies) in Mauritius. To achieve this task between both universities, the chosen approach is twofold:

On the one hand, it means to offer methods and tools for the scientific enhancement and dissemination of research results. An upstream enhancement has been initiated five years ago as a e-research project relying on the Iterative Knowledge Base System $\mathcal{IKBS}$ [2].

On the other hand, a downstream enhancement intends to provide pedagogical approaches borrowing from "The Darwin World - Le monde de Darwin", a large e-learning project initiated a few years ago in Quebec, through the Cyberscol School net [6], to make these

[1] IREMIA, Université de la Réunion, 97715 Saint Denis, BP 7151, France. e-mail : ikbs@univ-reunion.fr
[2] VCILT, University of Mauritius

results available to an audience wider than the sole community of academic researchers. This side of the project interests undergraduate students or members of communities dedicated to the protection of the environment.

To contribute to the process aiming at the promotion and distribution of research work, IREMIA has been involved in the building up of an information system (IS) for helping decision-makers to manage insular tropical environments formed by the future natural Reunion reserves: $ETIC^3$ [10]. For example, the $\mathcal{IKBS}$ team is building up a coral biodiversity IS that takes into account two skill levels: knowledge and know-how. The former is made up of encyclopaedic information in books from different contextual data sources, bibliographic, taxonomic, bio-geographic, eco-biologic, photographic, etc. The latter includes the professional knowledge from the specialist, this intuitive experience being acquired during a whole-life practice in the field. IS are such research applications of a life science domain combining knowledge bases and databases.

These sort of IS will be taken by the VCILT as the kernel of an educational portal. The $\mathcal{IKBS}$ research and DARWIN pedagogical enhancements intend to disseminate a better knowledge on biodiversity, notably for corals and flora of the Mascarene archipelago (Reunion, Mauritius and Rodrigues islands). The combination is based on knowledge building approaches, in collaboration with content expert specialists, cogniticians and users.

Transfers of competence and technologies ($\mathcal{IKBS}$) between both institutions are targeted to allow the emergence of a virtual research & learning community whose activities can create, at the end of the day, an effective and efficient social awareness around those environmental issues, crucial for the Indian Ocean region.

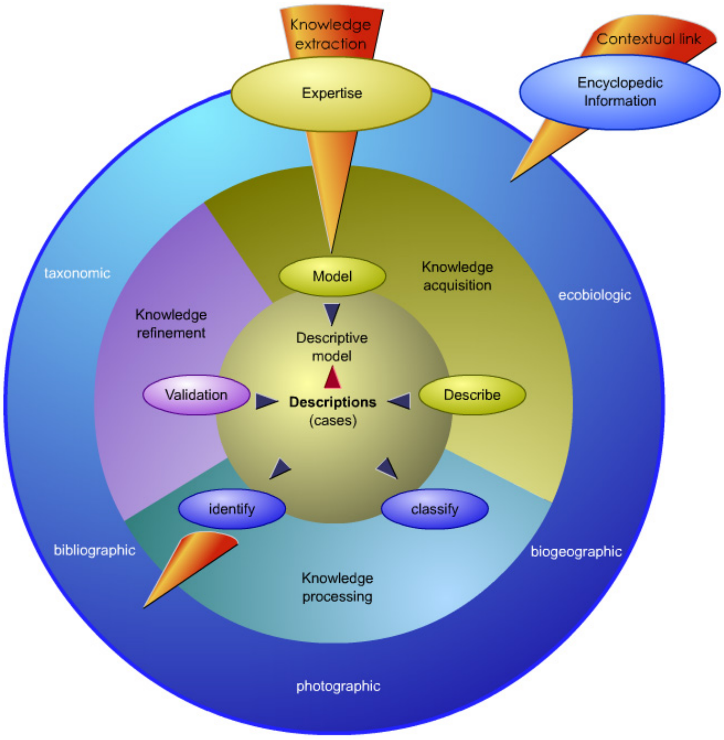

Figure 1. Knowledge management cycle of $\mathcal{IKBS}$

4 Upstream enhancement: the $\mathcal{IKBS}$ research approach

$\mathcal{IKBS}$ is a generic software workshop for building knowledge bases on the internet in natural sciences [2]. It was applied to Systematics, the scientific discipline that deals with listing, describing, naming, classifying and identifying living beings. For creating information systems on biodiversity that are accessible on the internet, we need to enhance and share the use of the systematician's expertise. For the identification task of a specimen by a non-specialist, a knowledge base is the kernel that delivers the know-how of experts in a given domain. $\mathcal{IKBS}$ manages complex knowledge (structured, variable, imprecise, noisy) with an object oriented representation formalism written in the Java language. It applies the experimental and inductive approach in biology, conjecture and test [5], with a natural process of knowledge management (Fig. 1) :

1. Acquisition of a descriptive model and related descriptions,
2. Processing of this knowledge for classification and identification, and
3. Experimentation, validation and refinement of cases and descriptive models (iteration).

For the knowledge acquisition part, the expert defines his domain at a taxonomic level with observable characters (Objects, Attributes, Values). They are structured in a description tree and the result is called a descriptive model. It is designed with the help of natural descriptive logics (decomposition, specialisation, contextual conditions, iteration, etc.). Then, a hypermedia questionnaire that matches the descriptive model is automatically built. Different users use it as an observation guide for collecting descriptions of observed specimens in a case base, and the expert references each description by its species' name.

In the knowledge processing stage, a decision tree is automatically learnt from the system by discriminating the pre-classified examples. Classification rules (diagnoses) may also be generated. Any user identifies new observations by following tree nodes from the root to leaves and answering questions. $\mathcal{IKBS}$ uses a **Case-Based Reasoning** [1] [8] approach to switch dynamically to alternative questions when an unknown answer is given for any particular question asked by the system: it then follows another decision sub-tree.

The knowledge refinement step uses an iterative process to evaluate the robustness of the previous descriptive model and descriptions. First, experimentation is led with a part of the case base. Results of classification and identification as well as the use of the questionnaire with the test set give the expert the possibility to detect inconsistencies in the case base (omissions or errors in descriptions) and descriptive model (misunderstood characters, bad illustrations). As applications, knowledge bases for several families of corals (Pocilloporidæ, Fungiidæ, Astrocœnidae, Siderastreidæ) of the Mascarene Archipelago were built [9]. The experimentation put stress on the difficulty for one expert to choose significant characters easy to observe and interpret. The meticulous choice of vocabulary, drawings and images is decisive for the robustness of the identification task. On the same way, covering the intra-specific variability by a sufficient number of cases is essential [3].

This is why $\mathcal{IKBS}$ aims to build knowledge bases collaboratively between experts. The objective is to draw up a thesaurus of terms and illustrations (i.e. the questionnaire) in order to reduce differences of interpretations that are the primary causes of misidentifications. As collections of specimens and experts are distributed around the world, we need to do remote systematics with very high-speed

	Pedagogical Sequence	Actions	Interacting Actors
1	Choose a family of corals or plants	TE[1] discusses with KE[2] and Domain Expert when available	TE - KE
2	Formalize a request for the development of a model of this family	Through the $\mathcal{IKBS}$-DARWIN educational portal	TE - KE
3	Engage in a research process about the family and related species	Field trip (gathering of samples, docs, photos, etc) + documentation	TE - KE
4	Use collected informations to develop and populate an $\mathcal{IKBS}$ model of the family	Work on the $\mathcal{IKBS}$ model editor to produce tentative model	TE - KE
5	Validation	Submit the model to the experts for validation – the final model is released on the web	TE - Recognised Domain Expert
6	Interpretation	Interpretation of specimens through an interaction with the $\mathcal{IKBS}$-model.	Add case to Knowledge Base End-Users, student public at large, etc
7	Maintenance and communication	Back to step 3	TE- KE Web Developer

Table 1. Applied to corals or endemic plants of the indian ocean region, the pedagogical sequence follows the seven detailed steps.

broadband networks. Each expert is responsible for modelling a family. Using videoconference, they share their interpretations of observations synchronously with specimens under the microscope and a camera connected to it. The product of such a tool is a collaborative knowledge base of a domain, that can evolve (by updating the knowledge) and be connected to distributed databases (bibliographic, photographic, geographic, taxonomic, etc. . .) that will yield information on species after the identification process of a new specimen. Moreover, the objective of $\mathcal{IKBS}$ is to offer more robust descriptive work in systematics in order to facilitate the identification of species by non specialists: we are faced with the problem of knowing, defining and reaching a consensus on: 1- what to describe (taxonomic level of descriptions), 2- how to represent descriptions (the choice of descriptive logics), and 3- what terms to use to create an illustrated thesaurus to be presented as a reference for specialists of a domain.

5 Downstream enhancement : the Darwin Pedagogical Approach

The pedagogical approach borrows from " Darwin's World – Le Monde de Darwin", a project initiated in Quebec a few years ago, at the university of Sherbrooke in the context of the Cyberscol educational network. with the objective to foster scientific thinking with hypermedia technology [6, 7].

"Darwin's World " pedagogical approach relies on three basic principles :

1. Science is a collective endeavour (scientific rigor emerges through exchanging between peer experts)
2. Science involves social commitment (scientists are socially responsible for the knowledge they produce, be it facts, theories or technologies)
3. Scientific progress stems from questions (scientists do not only look for answers, they crave for good questions as unique tools to probe and decipher the real world)

To apply these principles, the pedagogical sequence cycles through a set of seven steps (see table 1), involving different types of interactions from several kind of actors, human or technological.

[1] TE: Trainee-Expert
[2] KE: Knowledge Engineer

5.1 Step 1 : Choose a family of corals or plants

Based on the information present and expert availability for a particular family of plants or corals, the knowledge engineer with collaboration of an expert/trainee expert decide to work on the development of a model for that particular family.

5.2 Step 3 : Engage in a research process about the family and related species

Research helps to understand the aspects and characteristics of the chosen family. This is the analysis phase before the development of a model and it will concern mainly the expert and the knowledge engineer. Research teams may be organised to work on different aspects of the family. For example, a research team for a particular family of plant may be working on the medicinal aspects while another team may be working on the characteristics of its flowers. The teams must however work in collaboration to come up with a quality model.

5.3 Step 4 : Use collected informations to develop and populate an $\mathcal{IKBS}$ model of the family

Once the research phase is completed, a model is conceived for the family selected. An example will be a tree structure to highlight the model specifications or a specifications document. The model is populated with species of the family will full documentation of their characteristics. Relevant details for the model such as photographs, descriptive texts or animations are collected to provide better textual and visual representation of a particular model. A tentative model is then built using the $\mathcal{IKBS}$ model editor.

5.4 Step 5 : Validation

Contact other responsible persons (scientific coordinator/expert/linguistic moderator) by mail or any other means to discuss the scientific and linguistic validity of the model and the specifications document. These persons can make proposals to bring modifications to the model or its documentation. Once the model descriptions and specifications are validated by the responsible persons, it is entered in the $\mathcal{IKBS}$ system to be finally accessible on the Web. By entering the details at this stage in the $\mathcal{IKBS}$ system ensures that a quality model will be accessible. This process continues for other families of corals or plants and the same procedure may apply.

5.5 Step 6 : Interpretation

Once the model is online, it is important that there is enough information for users to use the system (Usable System), to understand terms that are not clear (i.e have access to a glossary), to get details on a particular family of plants and corals and interpret the results of the system – for instance successful identification of a species.

A part from identification, users will be able to browse through a populated database of plants and corals containing full details concerning habitats, cultural aspects, morphology etc of these species. Moreover, by having access to a wide variety of information that consist the pedagogical context of the system ($\mathcal{IKBS}$-DARWIN), the students or members of communities eventually place themselves at a "near-expert" level where they will, in turn be able to make the flow of knowledge a continuous and iterative process as they will become actively involved in its elaboration at the same level as the experts themselves. All these will be achieved only through a usable user-interface.

5.6 Step 7 : Maintenance and communication

The Interpretation phase (Step 6) provides the infrastructure for the continuous flow of knowledge while the Maintenance phase (Step 7) is essentially for continuous development of the Web site/Technology and enhancement/editing of models and addition/identification of new species. This phase maintains an Iterative cycle of development for the system. This phase also consists of answering of electronic mails from users concerning species, the system itself as well as technical problems with web site etc.

6 Conclusion : domain and learning objects

For knowledge sharing between experts and learners, the $\mathcal{IKBS}$-DARWIN project aims at favouring the convergence of domain and learning objects. Professionals of the domain are able to define a descriptive model with $\mathcal{IKBS}$ that makes sense for the purpose of the domain. As an example here for biodiversity knowledge sharing, Systematics is the science that involves observing, describing, naming and classifying living organisms. In order to be useful for end-users, this ontology (also called metadata) must be shared with neophyte observers: these persons will have to use the above domain objects to identify a specimen and reach other useful information. The experts domain objects become learning objects when a negotiation on their interpretation in use is possible between practitioners: the DARWIN system suggests processes that enable learners to manipulate the knowledge components of conceptual network knowledge structures for purposes of description, classification, and identification of living organisms. The future of domain and learning objects is a continuous spiralling conversion of tacit knowledge (such as know-how and experience) into explicit knowledge that can be captured, shared with others, diffused within groups, and turned back into new tacit knowledge gained from learning by doing [4]. The performance of knowledge management in natural domains depends entirely on the possibility to develop a collective intelligence approach based on collaborative tools that permit the mediation between knowledge senders and receivers. On the long term, the project should develop a strong regional competency kernel through the use of technology as a tool to encourage the development, documentation and use of indigenous information and knowledge about insular tropical environment.

REFERENCES

[1] Aamodt A. and Plaza E., 'Case-based reasoning: Foundational issues, methodological variations, and system approaches', *AI Communications*, **1**(7), 39–59, (1994).

[2] Grosser D., *Construction itérative de bases de connaissances descriptives et classificatoires avec la plate-forme à objets IKBS : application à la systématique des coraux des Mascareignes*, Thèse de doctorat, Université de la Réunion, January 2002.

[3] Bailly G., Cros A., and Goulan V. Expérimentation des bases de connaissances sur les coraux des mascareignes. apport de stage de l'association Vie Océane pour le Conseil Régional de La Réunion, IREMIA, Université de La Réunion., 2001.

[4] Bailly G., Cros A., and Goulan V. The future of learning objects, the instructional use of learning objects. http://www.id2.usu.edu/Papers/KnowObjMentalModels/, 2002.

[5] Popper K.R., *La logique de la découverte scientifique*, Payot (Eds.) Press, Paris, 1973.

[6] Aubé M., 'Fostering scientific thinking with new technologies : a socio-cognitive approach', *J. A. Chambers (dir.), Selected Papers from the Eleventh International Conference on College Teaching and Learning. Center for the Advancement of Teaching and Learning.*, 1–8, (2000).

[7] Aubé M. and David R., "'darwin's world": Hypermedia technology to enhance scientific thinking in the kids', *Communication presented to ED-MEDIA 2000 : World Conference on Educational Multimedia, Hypermedia & Telecomunications*, (juin 2000).

[8] Conruyt N. and Grosser D., 'Managing complex knowledge in natural sciences', *Proc. 3rd International Conference on Case-Based Reasoning (ICCBR-99)*, 401–414, (1999).

[9] Conruyt N., Grosser D., Geynet Y., Faure G., Pichon M., Guillaume M., and Gravier-Bonnet N., 'Ikbs, a knowledge base management system for reengineering systematics: application to corals of the mascarene archipelago', *Proceedings of the 9th International Coral Reef Symposium*, (2000).

[10] Conruyt N., Courdier R., and Lun-Kwok-Sui T. Tic pour la valorisation et la diffusion des résultats de la recherche : le cas du programme pluriformations mer. Journal de la nature, Université de la Réunion, 2002.