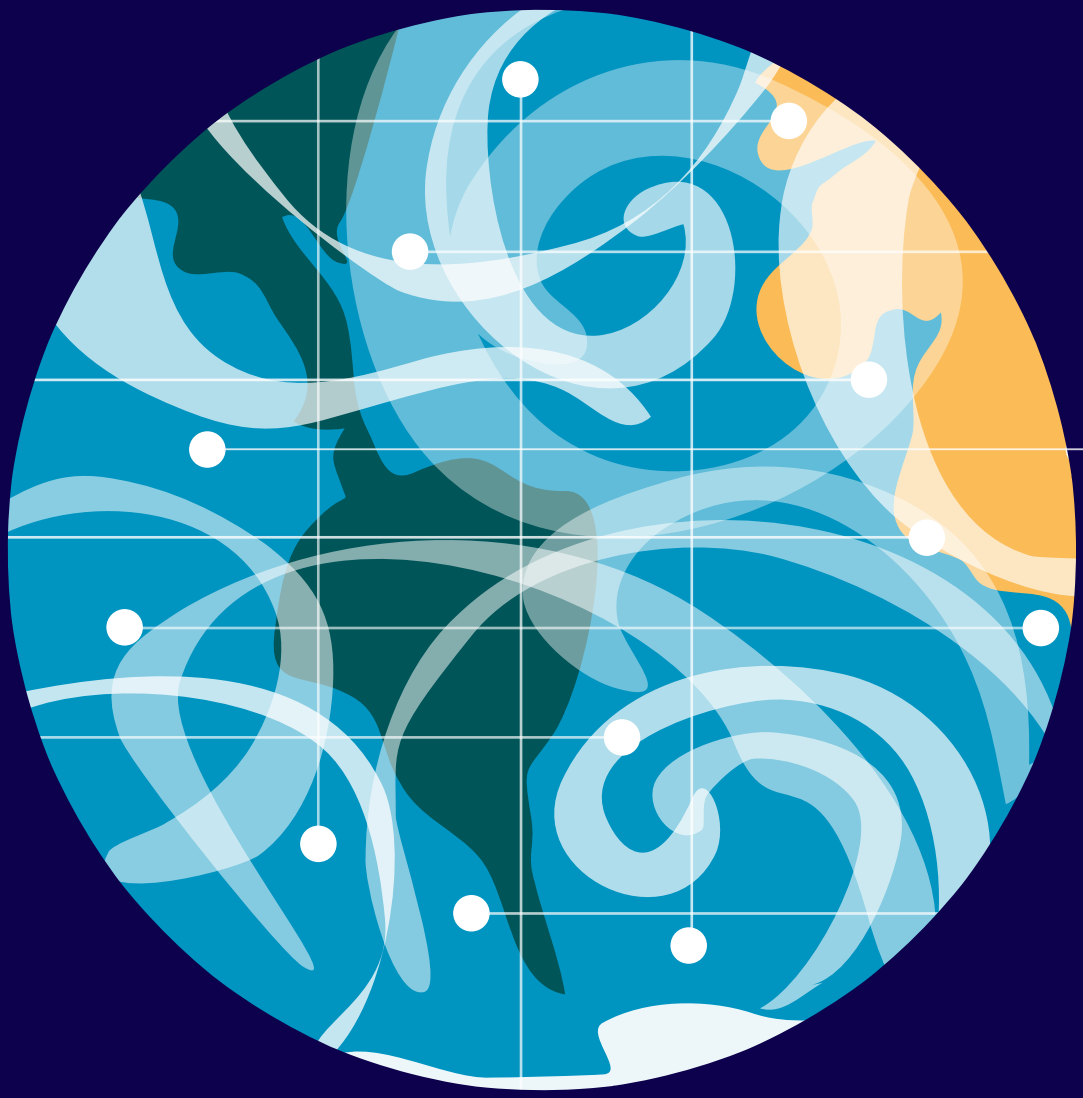
MICROSOFT RESEARCH-FAPESP VIRTUAL INSTITUTE FOR IT RESEARCH



WORLD CLASS RESEARCH WITH REAL LIFE IMPACT

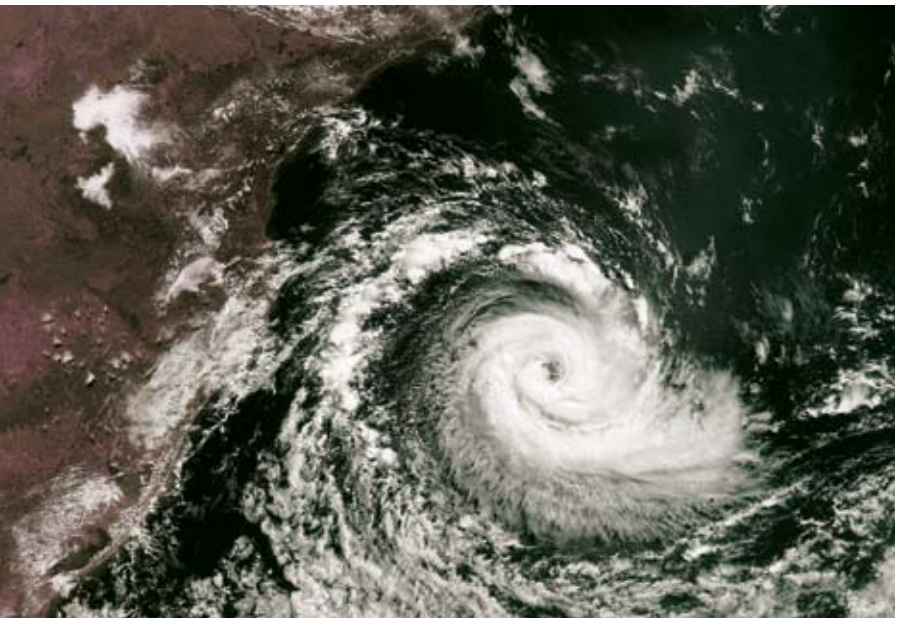
INFORMATION TECHNOLOGY TOWARDS SOCIAL PROBLEMS

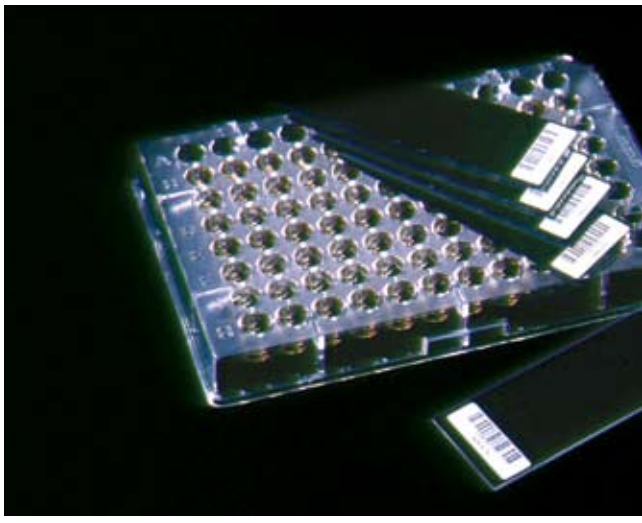
The Microsoft Research-FAPESP Virtual Institute for IT Research supports high-quality fundamental research in information and communication technologies that relates to social and economic development challenges in the region

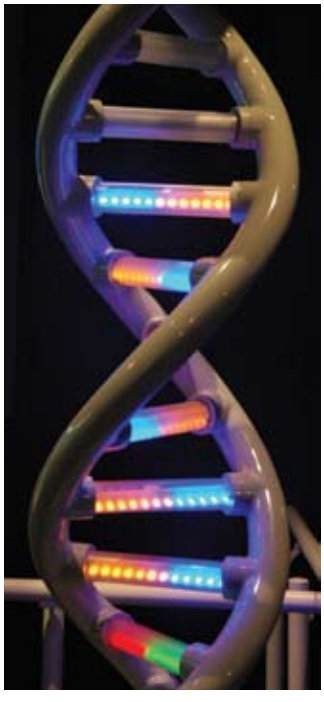
The Institute is a result of the Research and Collaboration Agreement between the State of São Paulo Research Foundation (FAPESP) and Microsoft Research, signed in December 2006. Within this Agreement, the two institutions, contributed a total of US$ 1.5 million to support research projects led by researchers in universities and research institutes in the State of São Paulo, Brazil.

Proposals are encouraged in the area of Computer Science and Engineering, which includes ideas stemming from collaborations between researchers in these areas and colleagues from others areas of knowledge such as Education, Healthcare/Wellbeing, Energy, the various disciplines covered by the Environmental Sciences, Biodiversity and the Social Sciences in general.

The Microsoft Research-FAPESP Virtual Institute for IT Research selected through Calls for Proposals.





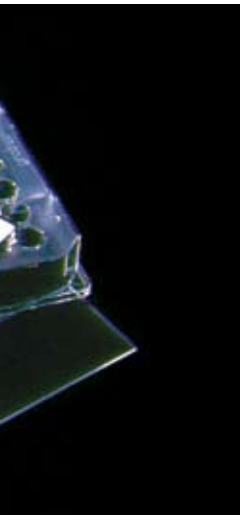


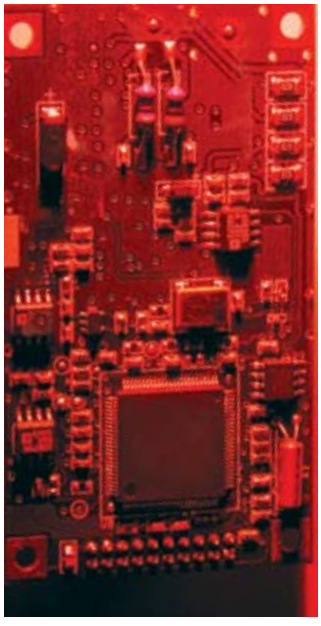
THE INSTITUTE´S OBJECTIVE

There has been growing recognition that the power, tools and techniques found in computing and information technology can be applied to create research breakthroughs and insight into key issues for individuals, society and the sustainability of the planet. The goal of the MSR-FAPESP Virtual Institute is to encourage multidisciplinary research involving those key areas and Computer Science.

The Microsoft Research-FAPESP Virtual Institute for IT Research looks for fundamental, high impact research proposals undertaken in order to gain knowledge and understanding towards critical problems in the areas above mentioned.

The impact and potential contribution of the research projects in this context is significant since it focuses on complex and high priority problems that demand innovative solutions.





ABOUT FAPESP AND MICROSOFT RESEARCH

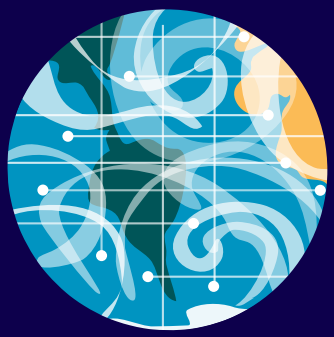
The State of São Paulo Research Foundation (FAPESP), founded en 1962, is one of the main funding agencies for scientific research in Brazil. Since then, it supports research in all areas of knowledge in the State of São Paulo.

Microsoft Research was launched in 1991 and its mission is to advance the state of the art in the areas in which they do research, rapidly transfer innovative technologies into Microsoft products and ensure that Microsoft products have a future. Through External Research, they also support research engagements with the academia throughout the world.

For more details about Microsoft Research-FAPESP Virtual Institute for IT Research and the Calls for Proposals access: www.fapesp.br/convenios/microsoft.

FAPESP Microsoft®

www.fapesp.br/convenios/microsoft



FAPESP

THE STATE OF SÃO PAULO
RESEARCH FOUNDATION

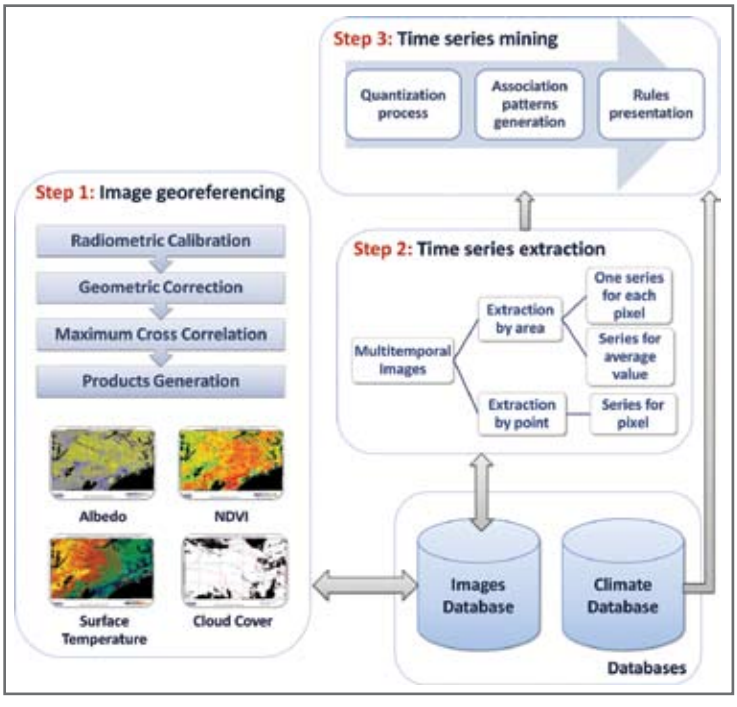
RUA PIO XI, 1500, ALTO DA LAPA
CEP: 05468-901 - SÃO PAULO - SP - BRASIL
FONE: 55.11. 3838-4000
www.fapesp.br

AGRODATAMINE: DEVELOPMENT OF ALGORITHMS AND METHODS OF DATA MINING TO SUPPORT RESEARCHES ON CLIMATE CHANGES RECORDING AGROMETEOROLOGY

Agma Juci Machado Traina

São Carlos Institute of Mathematics and Computer Sciences / University of São Paulo (USP)

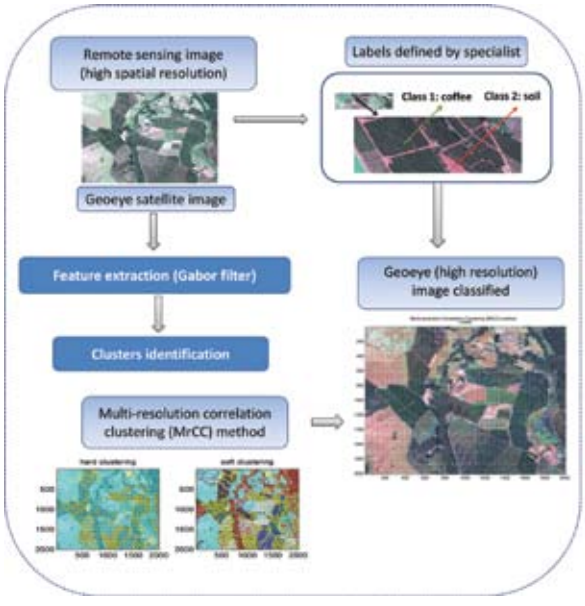
This research project aims at investigating and developing techniques and computational methods to evaluate complex data sets used by meteorologists and agro-meteorologists to assist in their research on the impacts of climate change in the Agriculture. Therefore, this project proposes: 1) developing new methods to analyze the outputs of regional models of climate change prediction; and 2) to compare them with real measurements collected by surface stations and satellites in order to assess the quality of the generated data and make allowances for calibrating the models. We intend to develop new methods to filter, analyze and extract association patterns between researchers from the Database and Images Group (GBDI) at ICMC/USP, the Embrapa Agriculture Informatics, the Cepagri (Unicamp), the Databases Groups at UFSCar, UFABC and CPTEC-Inpe.



MAIN RESULTS

1. A large database of the Agrodatamine project: The database was designed considering data requirements identified by the experts in the agrometeorology area members of the project. We have developed the conceptual and logical designs for the database schema, respectively using the Entity-Relationship Model and Relational Model. In order to build the database, we have chosen the DBMS (Databse Management System) PostGreSQL 8.4 (http://www.postgresql.org), as it is an open source and provides the resources to meet the project needs. This database is only for the project use, since the data are part of an agreement between Cepagri-Unicamp and CPTEC-Inpe.
2. The first version of the software SatimageExplorer: the software has been developed aimed at automatically generating time series from satellite image sequences. The specialist can define a region of interest (ROI) to be analyzed and the system provides the time series of the region regarding a given index or measurement. New indexes and measurements can be designed and introduced to the system in order to provide fast creation of new time series and their analysis. The system is a valuable tool to aid agrometeorologists to evaluate a crop development and climate changes. Versions for MS Windows as well as Linux Ubuntu are now available. It is already available as a beta test version.

PRODUCTS/PUBLICATIONS



JOURNALS

Vespa TG, Traina AJM, Traina Jr. C. Efficient bulk-loading on dynamic metric access methods. *Information Systems Journal (IS).* Elsevier, **35(5)**: 557-569. July 2010.

Vieira MR, Chino FJT, Traina Jr. C, Traina AJM. A visual framework to understand similarity queries and explore data in metric access methods. Special issue on Beyond Multimedia and XML Streams Querying and Mining. *International Journal of Business Intelligence and Data Mining (IJBIDM).* **5(4)**: 370-397. 2010.

Kaster D, Bugatti PH, Traina AJM, Traina Jr.C. FMI-SiR: A flexible and efficient module for similarity searching on oracle database. *Journal of Information and Data Management (JIDM)*. **1(2)**: 229-244. Sep 2010.

Romani LAS, Avila AMH, Zullo Jr. J, Traina Jr.C, Traina AJM. Mining relevant and extreme patterns on climate time series with CLIPSMiner. *Journal of Information and Data Management (JIDM)*. **1(3)**: 245-260, Sep 2010.

Traina Jr. C, Traina AJM, Leejay W, Faloutsos C. Fast feature selection using fractal dimension. *Journal of Information and Data Management (JIDM)*. **1(1)**: 3-16. June 2010.

Rodrigues Jr. JF, Tong H, Traina AJM, Faloutsos C. Large graph visualization with GMine. *Journal of Visual Languages & Computing*. Elsevier, 28 pags. (*accepted*)

INTERNATIONAL CONFERENCES

Cordeiro RLF, Guo F, Haverkamp DS, Horne JH, Hughes EK, Kim G, Traina AJM, Traina Jr. C, Faloutsos C. QMAS: querying, mining and summarizing multi-modal databases. *10th IEEE International Conference on Data Mining (ICDM´2010).* December 14-17, 2010. Sydney, Australia.

Bueno R, Ribeiro MX, Traina AJM, Traina Jr. C. Improving medical image retrieval through multi-descriptor similarity functions and association rules. *23rd International Symposium on Computer-Based Medical Systems (CBMS 2010),* pp. 309-314. October 12-15, 2010. Perth, Australia.

Porto Ferreira MR, Ponciano Silva M, Amo S, Pereira FSF, Traina AJM, Traina Jr. C, Chbeir R. Integrating user preference to similarity queries over medical images datasets. *23rd International Symposium on Computer-Based Medical Systems (CBMS 2010)*, pp. 486-491. October 12-15, 2010. Perth, Australia.

Silva SF, Brandoli B, Eler DM, Batista Neto JES, Traina AJM. Silhouette-based feature selection for classification of medical images. *23rd International Symposium on Computer-Based Medical Systems (CBMS 2010)*. pp 315-320. October 12-15, 2010. Perth, Australia.

NATIONAL CONFERENCES

Chino DYT, Romani LAS, Traina AJM. Extração de séries temporais de imagens de satélite para monitoramento de safras agrícolas e de dados climáticos. *CTIC - XXIX Concurso de Trabalhos de Iniciação Científica da SBC*, pp. 08 pags. 20-23 de julho de 2010, Belo Horizonte, MG. Second place in the context.

Nunes SA, Romani LAS, Avila AMH, Traina Jr. C, Sousa EPM, Traina AJM. Análise baseada em fractais para identificação de mudanças de tendências em múltiplas séries climáticas. *25 Brazilian Symposium on Databases*. Belo Horizonte, MG. *Proceedings of the XXV Brazilian Symposium on Databases - Short Paper Session, 2010*. pp. 65-72.

Vieira MR, Chino FJT, Traina Jr. C, Traina AJM. MAMView: a framework for visualization of metric trees. *Simpósio Brasileiro de Banco de Dados (SBBD 2010).* Belo Horizonte, MG. October 5-8, 2010. *16 – Demo Session, 2010*. pp. 1-6.

POSTERS

Traina AJM, Romani LAS, Cordeiro R, Sousa EPM, Ribeiro MX, Ávila AMH, Zullo Jr. J, Traina Jr. C. How to find relevant patterns in climate data: an efficient and effective framework to mine climate time series and remote sensing images. *2010 SIAM Annual Meeting (AN10).* Pittsburgh. July 12-16, 2010.

Agma Juci Machado Traina

Instituto de Ciências Matemáticas e Computação
de São Carlos – Universidade de São Paulo (USP)

Av. do Trabalhador Sãocarlense, 400
Caixa Postal 668 – Centro
13560-950 – São Carlos, SP – Brazil

agma@icmc.usp.br
55.16.3373-9674

SINBIOTA 2.0 – BIOTA-FAPESP PROGRAM´S INFORMATION SYSTEM: PLANNING THE NEXT 10 YEARS

Carlos Alfredo Joly

Institute of Biology / State University of Campinas (Unicamp)

Translation of biodiversity data into standard metadata formats in the future SinBiota 2.0

For the last 10 years, SinBiota, the information system of the Biota-FAPESP Program, has served the community storing data, showing them in maps, and providing a common base for researchers in different areas to communicate and exchange biodiversity data. State administration has also used the system to support new laws and regulations. Now it is time to rethink the system, which has grown in an ad-hoc unstructured way, aiming at the support that will be needed for the next 10 years.

Usage of the system in certain conditions, such as for field work, for instance, were not initially anticipated by the original plan. In addition, system extensions to allow data exchange with modeling tools for species niches and climatic change scenarios, or integration with molecular data from GenBank, DNA barcoding, and other important data repositories were not initially envisaged.

MAIN RESULTS

The current project aims at the creation of a new specification of the SinBiota system that dares to include services and technologies on the verge of the research on information technology. These new services and technologies should guarantee the use and expansion of the system for the next ten years; support the availability of a larger amount of higher quality environmental data, oriented also to the educational and public administration sectors; provide more efficient sharing of data among BIOTA researchers; interoperate with international initiatives such as the Global Biodiversity Information Facility (GBIF); provide effective tools to assist researchers in finding relevant information amongst a large amount of environment data. The expected results at the end of this project are the composition of the specification (Reference Document) of the SinBiota 2.0 system and the implementation of a prototype that will replace the current system.

Deliverables produced in the first year of this research project:

1- Current SinBlota documentation (use cases, ERD document);

2- Biodiversity Information Systems studies (GBIF, ALA, and OBIS review report);

3- Technology Reviews (Cloud Computing, Database Scalability and Security, Microsoft Tools reports, Biodiversity Metadata, Modeling Tools Data Conversion, Multimedia Search, Social Network, Monitoring Networks, DNA Barcode BOLD System services);

4- Prototype Architecture;

5- Prototype Planning Documentation (project decisions);

6- Georeferential Maps Demo;

7- Data logging and Web statistics modules (prototype);

8- Taxonomy modeling (prototype);

9- System access and security features implementation report.

PRODUCTS/PUBLICATIONS

INTERNACIONAL CONFERENCES

Position paper "SinBiota 2.0: displaying biodiversity and environmental data from Brazil using bing maps". *Environmental Research Workshop*. July, 2010.

Invited paper "SinBiota 2.0: planning a new generation environmental information system". *Microsoft Research eScience Workshop.* October, 2010.

POSTERS

Poster "SinBiota 2.0: planning a new generation environmental information system". *Microsoft- FAPESP Environmental Science Workshop*. November, 2010.

Carlos Alfredo Joly

Instituto de Biologia
Universidade Estadual de Campinas (Unicamp)

Departamento de Biologia Vegetal
Caixa Postal 6109 – Barão Geraldo
13081-970 – Campinas, SP – Brazil

cjoly@unicamp.br
55.19.3521-6166 / 3521-6173

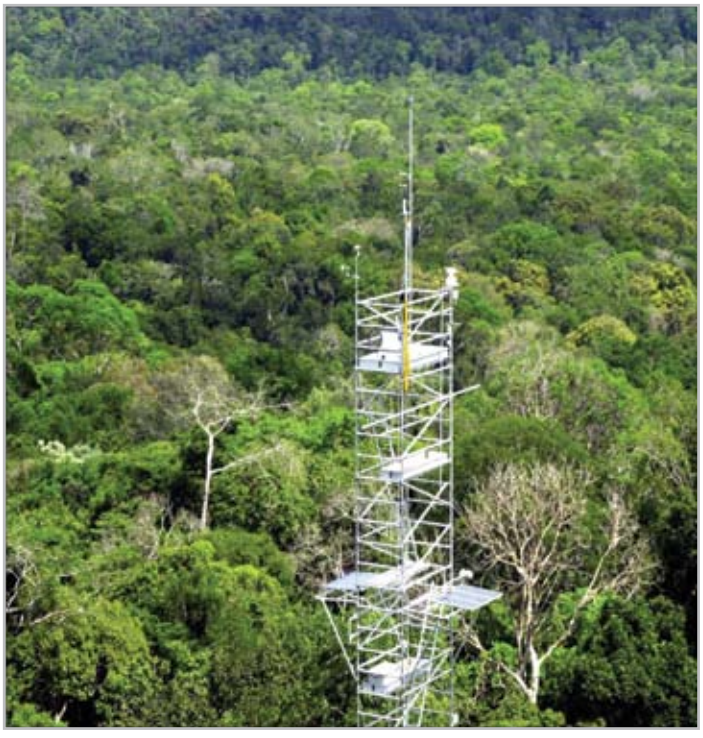
DEVELOPMENT AND APPLICATION OF NETWORK OF GEOSENSORS FOR ENVIRONMENTAL MONITORING

Celso von Randow

National Institute for Space Research (INPE)

This project proposes activities of research and technological development to test the use of a prototype of environmental sensors (geosensors) in a study site of tropical forest in Amazonia, Brazil, forming a sensor network with high spatial resolution, and to develop software tools for data quality control and integration. The main premise is that the geosensors should have relatively low cost, what enables the formation of monitoring netwoks with a large number of sensors spatially distributed.

The project is composed of three main components: 1) assembly and calibration of prototypes of geosensors of air temperature and humidity, with reproductive and reliable ceramic sensor elements that will adequately operate under the environmental conditions observed in the tropics; 2) development of software tools for management, quality control, visualization and integration of data collected in geosensor networks; and 3) realization of an experimental campaign, with the installation of the sensors in the Amazonian forest, with the objective of estimating the spatial variability of temperature and humidity complementing a flux tower in a complex terrain, to better describe the properties of the air flow and horizontal advection within and above the canopy near the tower.



Rain forest: image of the experimental area

MAIN RESULTS

Basic installed database/web server and development workstations with the following capabilities: a) The database can import and organize data collected from a network of sensors (since the sensor data for this project are not still available, it has been used data from a similar project to start building the database infrastructure); b) Basic reporting ability, including the ability to create time-series plots from each sensor/mote data; c) Basic data access through web services is under development and test.

Is has already started testing some analysis and visualization algorithms with the data. Since the sensor database that has been used contains only time series of measures of the temperature and humidity sensors, tools to visualize and highlight discrepancies in time series have been developed that should behave more or less the same.

Also under implementation is a generic selection and sampling tool that will create a subset of the dabatabe with the sensors´values selected accordingly to the users´criteria.

The next steps will be the creation of a more generic tool based on these concepts and its deployment through the project´s web page.

PRODUCTS/PUBLICATIONS

INTERNATIONAL CONFERENCES

Von Randow C, Santos RDC, Rocha HR. Network of environmental sensors in tropical rain forests. Talk presented at *Microsoft Research Faculty Summit 2010*. Redmond, WA. July 12-13, 2010.

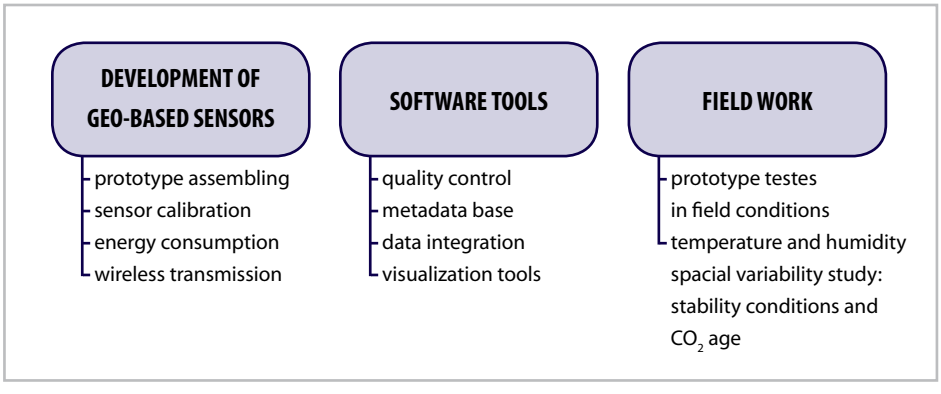
Santos RDC. Development and application of network of geosensors for environmental monitoring. Talk presented at *2010 Microsoft Research eScience Workshop.* Berkeley, CA. October 11-13, 2010.

Von Randow C, Santos RDC, Rocha HR. Network of environmental sensors in tropical rain forests. Paper accepted for presentation at *2010 AGU Fall Meeting.* San Francisco, CA. December 13-17, 2010.

NATIONAL/LOCAL CONFERENCES

Von Randow C. Tackling spatial variability in biosphere-atmosphere interaction studies. Talk presented at *Microsoft-FAPESP Environmental Science Workshop.* São Paulo, SP. November 11-12, 2010.

Rocha HR. The geosensor network experiment in the Atlantic forest. Talk presented at *Microsoft-FAPESP Environmental Science Workshop.* São Paulo, SP. November 11-12, 2010.



Project's main challenges diagram

Celso von Randow

Instituto Nacional de Pesquisas Espaciais (Inpe)
Ministério da Ciência e Tecnologia (MCT)

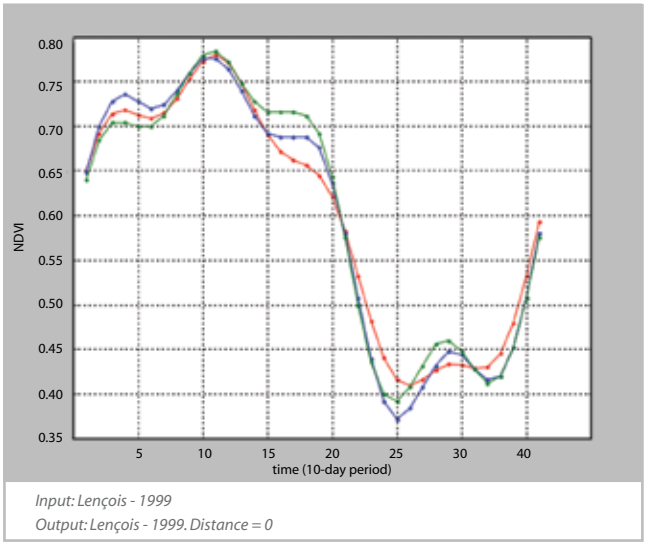
Rodovia Presidente Dutra, km 40
12630-000 – Cachoeira Paulista, SP – Brazil

celso.vonrandow@cptec.inpe.br
55.12.3186-9321 / 3186-9459

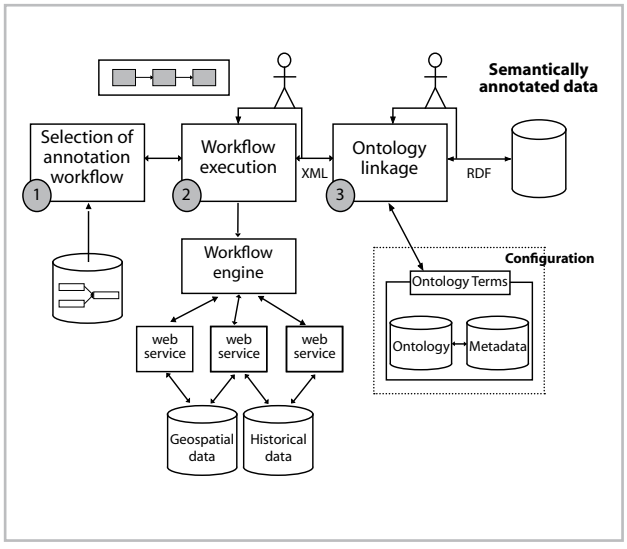
E-FARMS: A 2-WAY ROAD FROM SMALL FARMS INTO THE NETWORKED WORLD

Claudia Maria Bauzer Medeiros

Institute of Computer Science / State University of Campinas (Unicamp)



Screen copy from the eFarms site – co-evolution of NDVI time series, extracted from pixels of satellite images



Workflow-directed semantic annotation of geospatial digital sources

e-Farms was a multidisciplinary project conducted between Nov. 2007 and May 2010, combining research in computer science and agriculture. Its goal was to attack theoretical and practical problems involving sensor-based data management and wireless data communication in rural areas in Brazil. The project was centered on two needs of the Brazilian population: investigate low cost solutions for data communication in rural areas where data transmission is still very limited (due to cost and/or geographic factors); and provide models, methods and algorithms to support management, integration and analysis of sensor data, for decision support in crop management and agricultural planning. The research team was composed of computer scientists and researchers in agricultural sciences. The Cooxupe coffee cooperative, the largest of its kind in the world, with over 14.000 cooperating farms, was a partner of the project. Not only did it provide the basic case study, but also part of the software modules developed within the project was designed and validated within this large real context.

MAIN RESULTS

The project covered two kinds of activities: a) development and deployment of a wireless data communication network, from sensors to a central database server; and b) research on data structures, models, and algorithms applied to sensor and satellite data.

PROJECT SITE

www.lis.ic.unicamp.br/projects/efarms and
http://proj.lis.ic.unicamp.br/efarms

PRODUCTS/PUBLICATIONS

PUBLICATIONS

The main results of the project were reported in five journal papers in computer science, 15 conference papers, 13 of which in computer science and 2 in agricultural sciences. Five of the conference papers are in Brazilian conferences (in databases, image processing, and eScience), and 9 in international conferences (in databases, image processing, computer networks and GIS). These publications report theoretical and applied results in: a) databases (time series mining and correlations, workflow-driven annotation of geospatial data sources, query processing and real-time processing of sensor network data); b) satellite image processing (including new segmentation algorithms and automatic recognition of vegetation cover, for coffee crops, based on genetic programming); c) sensor network deployment and management; and d) methodologies for agricultural planning and crop productivity analysis from remote sensing data.

DEFENCES

The project was directly linked to the following defences: 1 MSc and 1 PhD in agricultural engineering; 3 MSc and 2 PhD in computer science. At the end of the project, additional graduate students were involved in research concerning project ramifications, namely 2 PhD in agricultural engineering; 3 PhD and 1 MSc in computer science.

TOOLS AND INFRASTRUCTURE

Part of the project consisted in the design and implementation of a multi-tier wireless sensor network, involving heterogeneous sensors, providing real time data that are relevant for crop monitoring – in particular, temperature, moisture and light. A first prototype of this network was deployed in Unicamp's experimental farm grounds. Sensor data are fed to the project's repositories, where they are cleaned, pre-processed, mined and analyzed by software tools. These tools implement new methodologies and algorithms for decision support in agriculture.

Claudia Maria Bauzer Medeiros

Instituto de Computação
Universidade Estadual de Campinas (Unicamp)
Av. Albert Einstein, 1251
Caixa Postal 6176
13084-971 – Campinas, SP – Brasil

cmbm@ic.unicamp.br
55.19.351215855

BORBOLETA: INTEGRATED MOBILE SYSTEM FOR HOME HEALTHCARE

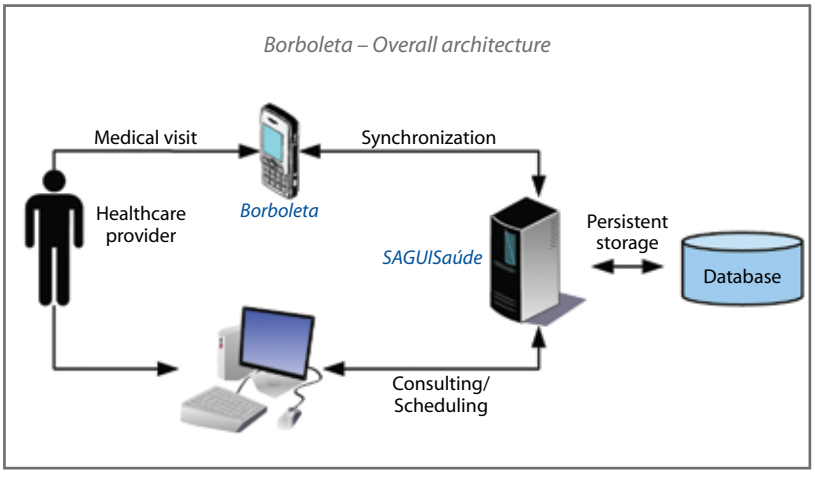
Fabio Kon

Department of Computer Science, Institute of Mathematics and Statistics / University of São Paulo (USP)

The Brazilian public health system takes care of the health of 140 million citizens. In this context, recent public programs targeted at the provision of preventive medicine at the homes of low income people have proven to be very effective in improving public health with a relatively low cost. However, in their current versions these programs have no support from Information Technology, leading to loss of agility and loss of valuable information. The goal of the Borboleta Project is to investigate novel software tools and IT methodologies to support public homecare programs, focusing on mobile applications for smartphones used by health professionals and a sophisticated multimedia database hosted on the Primary Healthcare Centers.

Website: http://ccsl.ime.usp.br/borboleta

SAGUISaúde open source system for management of Primary Healthcare Centers



MAIN RESULTS

CONFERENCE AND WORKSHOP PAPERS

Duarte GL, Correia R, Leal P, Domingues H, Kon F, Kon R, Ferreira JE. Borboleta and SaguiSaúde –open source mobile telehealth for public home healthcare. *In: Proceedings of the 8th International e-Health, Telemedicine and Health ICT Forum (Med-e-Tel)*. April 14-16, 2010. Luxembourg, Luxembourg.

Duarte G, Domingues H, Kon F, Ferreira JE. Heterogeneidade como solução para bancos de dados móveis em um sistema de saúde. Accepted for publication at the upcoming. *Workshop on Medical Informatics – SBC Brazil, 2009*.

Domingues H, Correia R, Kon F, Kon R, Ferreira JE. Análise e modelagem conceitual de um sistema de prontuário eletrônico para centros de saúde. *Workshop on Medical Informatics – SBC Brazil, 2008*.

Correia da Silva FS. 2008. Knowledge-based modality selection for information presentation in a mobile system for primary homecare. In Artificial Intelligence and Simulation of Behavior. *Workshop on Multimodal Output Generation*. Aberdee, UK.

Correia R, Kon F, Kon R. 2008. Borboleta: a mobile telehealth system for primary homecare. In *ACM Symposium on Applied Computing*. Fortaleza, Brazil.

Correia R, Kon Fabio, Conceição AF, Kon R. Sistema Móvel Multimídia de Código Aberto para Atenção Primária de Saúde com Visitas Domiciliares. *Workshop on Free Software – FISL´2009*.

POSTERS AND UNDERGRADUATE RESEARCH WORK

Conceição AF, Pereira RL, Rezende JVP, Silva BNM, Correia RJP, Domingues HH, Kon R, Kon F. Projeto Borboleta; Ferramentas móveis e multimídia para atenção básica domiciliar. *Congresso Brasileiro de Informática em Saúde – CBIS*. Campos do Jordão, Brasil. Novembro, 2008.

Kon R, Silva NSC, Lanferini GM, Faria SC. Desenvolvimento de modelo conceitual de sistema de informação para atenção primária à saúde. *Congresso Brasileiro de Informática em Saúde – CBIS*. Campos do Jordão, Brasil. Novembro, 2008.

Correia R, Kon F, Kon R, Ferreira JE. 2007. Um sistema de software livre para gerenciamento de centros de saúde. In *III Simpósio de Iniciação Científica e Pós-Graduação do IME-USP*. São Paulo, Brazil.

SPIN-OFF PROJECT

Sponsoring Agency: FINEP-MCT, Brazil
Title: Family Health Program – Mobile
Value: R$ 796.845,28
Coordinators: Arlindo da Conceição (Unifesp) and Vladimir Moreira (Infomobile)

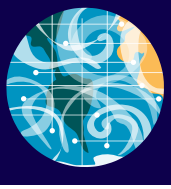
Borboleta open source system for homecare health provision using smartphones

Fabio Kon

Instituto de Matemática e Estatística
Universidade de São Paulo (USP)
Departamento de Ciência da Computação

Rua do Matão, 1.010 – Cidade Universitária
05508-090 – São Paulo, SP – Brazil

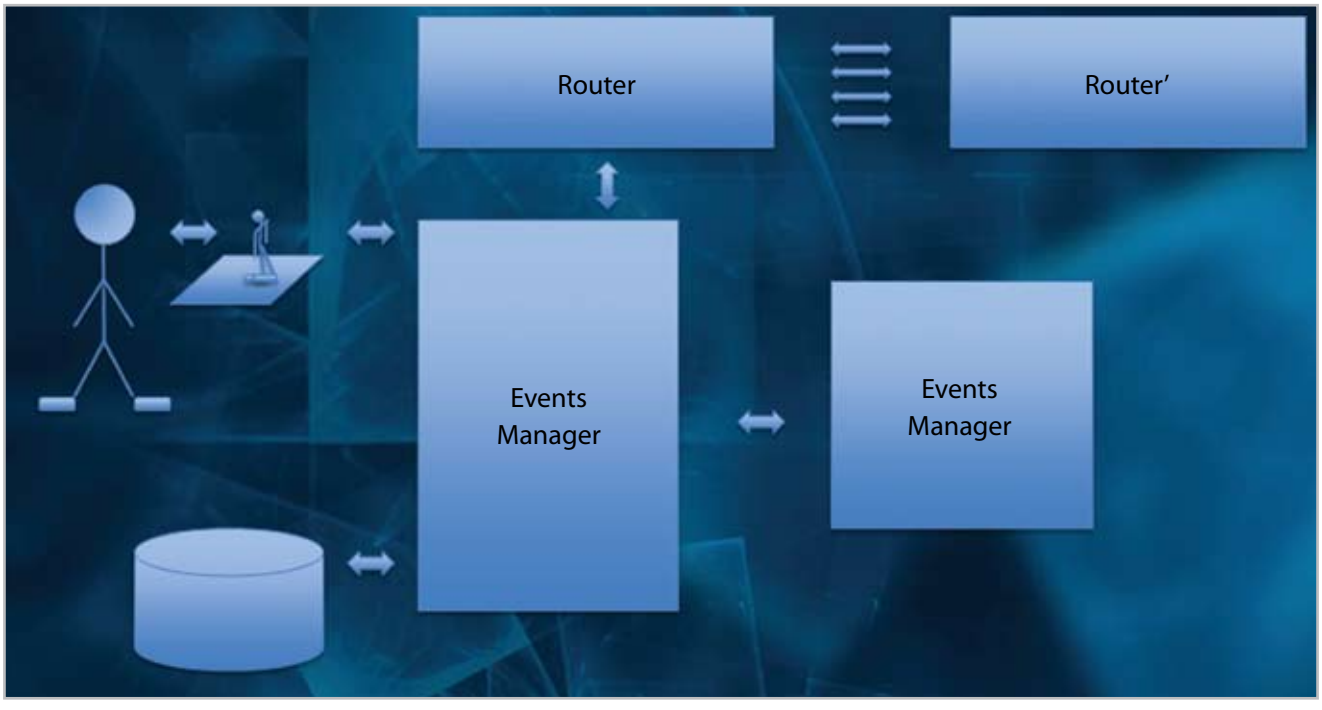
kon@ime.usp.br
55.11.3091-6498
55.11.3091-6135



JAMSESSION: A DECENTRALISED ARCHITECTURE FOR SPECIALISED VIRTUAL WORLDS AND THE WEB 3.0

Flávio Soares Corrêa da Silva

Institute of Mathematics and Statistics / University of São Paulo (USP)



JamSession – Virtual Worlds at your service

The technical goal of the project is the development of an architecture to build lightweight, specialised, decentralised virtual worlds. The design and implementation of specific worlds shall be based on specifications using a high level formal language, so that these activities become user friendly as well as formally verifiable. In order to illustrate,it will be develop some representative applications of this architecture. The broader ambition of the project is to enable entrepreneurs to build business initiatives based on virtual worlds using the JamSession architecture, based on minimised investments.

MAIN RESULTS

Several significant results have been recently obtained, directly related to the project goals. Some additional important results, however, shall be completed in the near future and featured as submitted papers to appropriate conferences and journals.

Two implemented prototypes have been implemented for the JamSession, both written in Prolog. Both prototypes can be downloaded from the project provisional webpage (http://lidetJamSession.wordpress.com). The latter prototype corresponds more closely to the up-to-date specification of JamSession. A series of papers have been written, some of which have already been published as technical reports. Downloads may be done from the project provisional webpage, as well as from the Lidet LAB webpage (http://lidet.wordpress.com).

PRODUCTS/PUBLICATIONS

Bressane Neto AF, Correa da Silva FS. 2009. Synthetic Characters with Personality and Emotion. *Intelligent Virtual Agents 2009. (poster)*

Araujo CJA, Correa da Silva FS. 2009. Governmental Virtual Institutions. Submitted to International Conference on Electronic Governance 2009. Colombia *(poster)*.

Correa da Silva FS. 2010. On the ethics of democratic access to web information. *Knowledge, Technology and Policy*, p. 97-107.

Correa da Silva FS. 2010. Interação Estado/Academia para inovação em governo eletrônico no Brasil. In: Cláudia do Socorro Ferreira Mesquita; Nazaré Lopes Breta. (Org.) *Panorama da Interoperabilidade no Brasil*. Brasília: MP-SLTI, p. 64-73.

JOURNALS

Knowledge-based interaction Protocols for intelligent interactive environments (*submitted*).

On the ethics of democratic access to web information (*Knowledge, Technoly and Policy*, 2010).

TimeSaver – Virtual worlds and active workflows to deliver friendly public services (*submitted*).

3-Nested Institutions – an organizational design pattern to optimize distributed workflows in electronic government (*submitted*).

NATIONAL/LOCAL CONFERENCES

On the construction of synthetic characters with personality and emotion (SBIA - Brazil, 2010).

POSTERS

Governmental virtual institutions (Icegov - Colombia, 2009).

Synthetic characters with personality and emotion (IVA - Netherlands, 2009).

TECHNICAL REPORTS

Please check http://lidet.wordpress.com for full list of technical reports.

PRESENTATIONS

Microsoft Research Latam Summit (Argentina, May 2009) – invited oral presentation.

Flávio Soares Corrêa da Silva

Instituto de Matemática e Estatística
Universidade de São Paulo (USP)
Rua do Matão, 1.010 – Cidade Universitária
05508-090 – São Paulo, SP – Brazil

fcs@ime.usp.br
55.11.3091-6135

AUTOMATED SCREENING FOR DIABETIC RETINOPATHIES: IT IN THE FIGHT AGAINST PREVENTABLE BLINDNESS

Jacques Wainer

Institute of Computer Science / State University of Campinas (Unicamp)

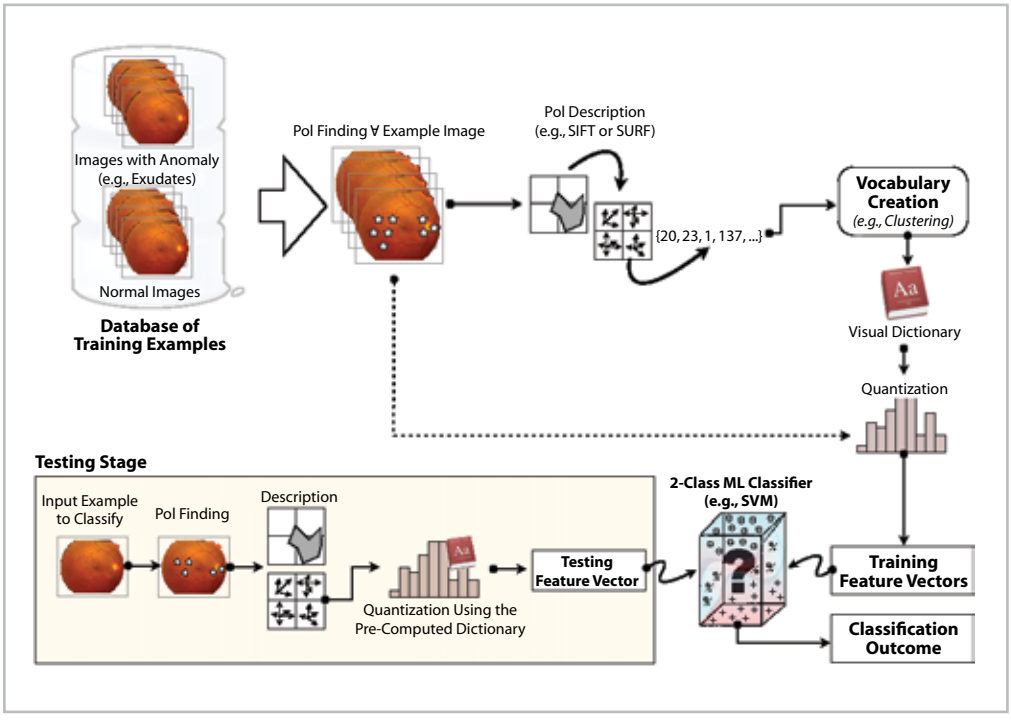


Figure 1

This study has the following objectives:

1) To develop a system for screening patients that will accurately detect cases of diabetic retinopathy through dilated eye examination – The system will be used as an automated retinopathy classification system, distinguishing between normal images (patients who do not require further attention) and images suggestive of retinopathy (patients who require the attention of a specialist). Therefore, specialists will not need to examine patients in whom no change in the *fundus oculi* is expected, allowing the specialists to do their job more efficiently. The expectation is that this system will produce few false-negative results (i.e., not detecting retinopathy when it is present), because an indication that there is no change in the *fundus oculi* would exclude the patient in question from receiving the attention of a specialist. The system should detect the following changes, all of which are suggestive of diabetic retinopathy: hemorrhage, exudates, vascular changes, scarring, hyperpigmentation, and hypopigmentation. The initial version of the system will not detect macular edema, which is also associated with diabetes.

2) To implement the system in a true teleophthalmology service, in which only images suggestive of disorder will be referred to specialists – This deployment is aimed at assessing the system "in production" and assessing the needs of infrastructure and organizational requirements for maintaining a diabetic retinopathy teleophthalmology service with automated support.

MAIN RESULTS

Detectors of the three most common abnormalities in diabetic retinopathy were developed. These detectors use a technique that is unusual in medical image processing: they detect "feature points" of the image (points at which there is textural discontinuity) and group the characteristics of those points into "visual words". Although this point characteristics/visual words technique (the so-called "bag-of-words" approach) has been used for image retrieval by similarity in other fields, it has not previously been applied to the classification of medical images. In comparison with the results obtained with state-of-the-art detectors, those achieved through the use of this technique were slightly less accurate for exudates etc., and comparable for microaneurysm and deep retinal hemorrhage. The advantage of this technique is that, in contrast with the state-of-the-art detectors, it does not appear to be specific to any one anomaly. The system can be "taught" to detect exudates and other anomalies with greater accuracy.

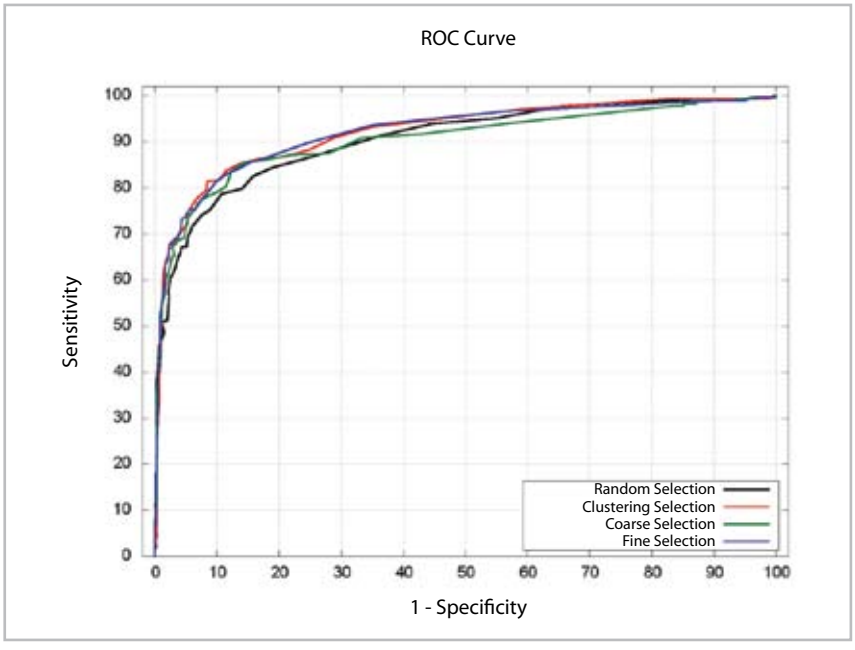


Figure 2

Using the three detectors in parallel, we can detect the three targeted abnormalities with a false-negative rate of less than 0.5% and a false-positive rate of approximately 20%. Unfortunately, these three detectors are not yet capable of detecting, at such rates, any other abnormality associated with diabetic retinopathy. Therefore, efforts are being made to develop detectors that can identify the next three leading anomalies.

In the first year of the project, we classified the 8,000 *fundus oculi* images initially available, and we are using those images in order to "train" and validate the detectors. Since September 2010, new images have been collected at a primary health center specializing in diabetes, where the imported retinal imaging system was made available for use, at an average rate of 20 per day. These new images have not yet been classified.

Jacques Wainer

Instituto de Computação
Universidade Estadual de Campinas (Unicamp)
Departamento de Sistemas de Informação

Caixa Postal 6176
13084-971 – Campinas, SP – Brazil

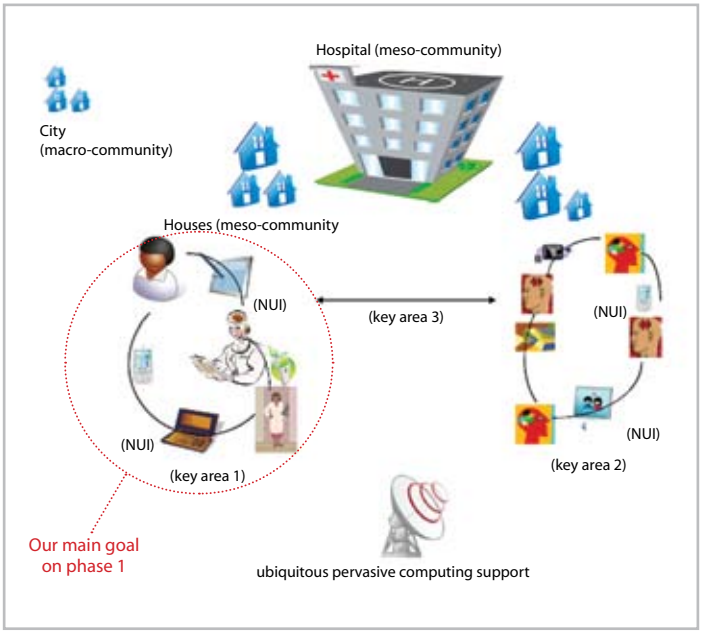
wainer@ic.unicamp.br
55.19.3521-5871

A CULTURALLY-SENSITIVE ENVIRONMENT FOR FLEXIBLE NATURAL INTERACTION TO SUPPORT TRANSITIONING FROM A CHRONIC CARE HOSPITAL CONTEXT

Junia Coutinho Anacleto

Center for Exact Sciences and Technology / Federal University of São Carlos (UFSCar)

This project, in a partnership with a special chronic care hospital for individuals with neurological and brain disorders, intends to focus on patients that are in the hospital selected for transitioning to live in society again. As changes in the therapeutic procedures for patients means changes in the working procedures for health professionals, this project looks mainly to investigate how Information and Communication Technology (ICT) can enhance natural interaction among health professionals that can lead to a more connected, tuned and active community, specially considering the nomadic nature of their work. Therefore, in the Hospital space and surrounds, it is necessary to investigate how a more natural, integrated and environmental ICT interactive system can provide better supportive tools and shared devices for executing those tasks. The main goal is to support them on the challenge of monitoring and communicating with patients, potentially allowing a smoother transition process for the patients from the Hospital to the community, the main concern and also mandatory by law in Brazil. Our research will investigate three kew areas to assist with their socialization, aiming to integrate institutionalized people into their home and the society: 1) supporting the community of health professionals to have new information channels to establish a support workflow and framework amongst themselves; 2) supporting the patients' community within their environment to allow them to establish communal skills and participate in their community; 3) support communication mechanisms between health professionals and the patient community based on natural interaction paradigms to coordinate monitoring and assistance appropriate to the level of support required by the patients to establish a sustainable model for community transitioning responsive to the dynamic nature of social inclusion. In each area, we are considering the health professional team and patients skills, culture, natural and flexible (adaptive and adaptable) interaction mechanisms and adopted procedures as well as how the health professionals and patients respond to ICTs as requirements to be addressed during the project. We anticipate the Natural User Interfaces we expect to design as tools for the health professional and patients will be useful for any group that has socialization deficits. As well, these technologies may also support patients' inclusion as they will provide access to the same social structures using the same devices as the non-disabled community.



Key areas for cultural and flexible interaction supporting transition with NUI

Junia Coutinho Anacleto

Laboratório de Interação Avançada (LIA)
Centro de Ciências Exatas e Tecnologia
Universidade Federal de São Carlos (UFSCar)
Rodovia Washington Luís, km 235
13565-905 – São Carlos, SP – Brazil

junia@dc.ufscar.br
55.16.3351-8615 e 55.16.8158-7626

E-PHENOLOGY: THE APPLICATION OF NEW TECHNOLOGIES TO MONITOR PLANT PHENOLOGY AND TRACK CLIMATE CHANGES IN THE TROPICS

Leonor Patrícia Cerdeira Morellato

Rio Claro Institute of Biosciences / São Paulo State University (Unesp)

The e-phenology is a multidisciplinary project combining research in Computer Science and Phenology. Its goal is to attack theoretical and practical problems involving the use of new technologies for remote phenological observation aiming to detect local environmental changes. It is geared towards three objectives: a) use of new technologies of environmental monitoring based on remote phenology monitoring systems; b) creation of a protocol for a Brazilian long term phenology monitoring program and for the integration across disciplines, advancing our knowledge of seasonal responses within tropics to climate change; and c) provide models, methods and algorithms to support management, integration and analysis of data of remote phenology systems. The research team is composed of computer scientists and researchers in phenology.

Leonor Patrícia Cerdeira Morellato

Instituto de Biosciências de Rio Claro
Universidade Estadual Paulista (Unesp)
Av. 24A, 1515
13506-900 – Rio Claro, SP – Brazil

pmorella@rc.unesp.br
55.19.3526-4205

E-CIDADANIA: SYSTEMS AND METHODS FOR THE CONSTITUTION OF A CULTURE MEDIATED BY INFORMATION AND COMMUNICATION TECHNOLOGY

Maria Cecilia Calani Baranauskas

Institute of Computer Science / State University of Campinas (Unicamp)

In the Brazilian society, we face a situation characterized by enormous differences with regard to socio-economics, culture, geographical region as well as access to technology and knowledge. The big challenge of Computer Science to change this reality lies in the search for methods and systems design that provide access and make sense to the users' community, thus supporting the formation of a digital culture that respects the diversity in our society. This research investigates and proposes solutions for interaction models and interfaces for the diversity of users and competencies that constitute the scenario of the digitally excluded people in our society. In general terms, the project proposed aims at studying and proposing solutions to the challenges of interaction and user interface design for systems in the context of citizenship practice. To reach this goal, the research team develops joint actions with a partner institution (network Jovem.com and communities around it) to conduct interaction and interface design of a pilot system to be implemented in the target community. Making available systems that make sense to and are accessible by the citizen – and thus also bring benefits for society as a whole – requires a socio-technical vision of the problem. Therefore, for this research we propose to use the frame of reference of Organizational Semiotics articulated with principles of Universal Design or Design for All. The contribution of this project lies in the advance of research in the area of (inclusive and universal) interfaces tailorable to users' diversity of competencies as well as in the creation of applications that actually make sense to the target community (i.e. constitution of inclusive social networks and e-citizenship).

Overview of e-Cidadania Project activities in a time-line:
http://www.nied.unicamp.br/ecidadania

MAIN RESULTS

RECENT PAPERS ACCEPTED FOR PUBLICATION IN SCIENTIFIC JOURNALS

Bonacin R, Baranauskas MCC, Liu K, Sun Lily. 2009. Norms-based simulation for personalized service provision. *Semiotica* (Berlin). **175**: 403-428.

Bonacin R, Mello AM, Simoni CAC, Baranauskas MCC. 2009. Accessibility and interoperability in e-Gov Systems: outlining an inclusive development process universal access in the information society. **2009**: 1-34.

Miranda LC, Hornung H, Baranauskas MCC. 2010. Adjustable interactive rings for iDTV. *IEEE Transactions on Consumer Electronics*. **56**: 1988-1996.

Hornung H, Baranauskas MCC. Towards a design rationale for inclusive eGov services. Accepted for publication in the *International Journal of Electronic Government Research*, 201x.

Neris VPA, Almeida LD, Miranda LC, Hayashi EC, Baranauskas MCC. Collective construction of meaning and system for an inclusive social network. Accepted for publication in the *International Journal of Information Systems and Social Change*. ISSN: 1941-868X.

Romani R, Baranauskas MCC. GWIDO – Games with interaction design objective. Accepted for publication in the *IADIS International Journal on WWW/Internet*.

SOME RECENT PAPERS IN INTERNATIONAL CONFERENCES PROCEEDINGS

Fortuna F, Bonacin R, Baranauskas MCC. 2010. A framework for flexibility at the interface: joining ajax technology and semiotics. In: *Proceedings of the 12th International Conference on Enterprise Information Systems*. Lisboa, Portugal: SciTePress Science and Technology Publications. **5**: 30-37.

Reis JC, Bonacin R, Baranauskas MCC. 2010. A semiotic-based qpproach to the design of web ontologies. In: *Proceedings of the 12th IFIP WG 8.1 International Conference on Informatics and Semiotics in Organisations ICISO 2010*, Reading, UK SciTePress Science and Technology Publications. 60-67.

Reis JC, Baranauskas MCC, Bonacin R. 2010. New perspectives for search in social networks – a challenge for inclusion. In: *Proceedings of the 12th International Conference on Enterprise Information Systems (ICEIS 2010)*, Funchal Portugal: SciTePress Science and Technology Publications. **5**: 53-62.

Reis JC, Bonacin R, Baranauskas MCC. 2010. Search informed by a semiotic approach in Social Network Services. In: *Proceedings of the 10th annual International Conference on New Technologies of Distributed Systems NOTERE 2010*. Tozeur. IEEE, 2010. 321-326.

Santana VF, Baranauskas MCC. 2010. Summarizing observational client-side data to reveal web usage patterns. In: *Proceedings of the 25th ACM Symposium On Applied Computing (ACM SAC)*, 2010. Stoughton, WI, USA. The Printing House. **II**: 1219-1223.

Hayashi ECS, Baranauskas MCC. Understanding meta-communication in an inclusive scenario. In: *Proceedings of the 25th ACM Symposium On Applied Computing (ACM SAC) 2010*. Stoughton, WI, USA. The Printing House. **II**: 1213-1228.

Neris VPA, Baranauskas MCC. 2010. User interface design Informed by affordances and norms concepts. In: *Proceedings of the 12th IFIP WG 8.1 International Conference on Informatics and Semiotics in Organizations*. SciTePress Science and Technology Publication. 133-140.

MSC DISSERTATIONS AND PHD THESES

Santana VF. Identificação de padrões de utilização da web mediada por tecnologias assistivas. MSc dissertation. Institute of Computing, Unicamp. April 8th, 2009 (Capes scholarship).

Hayashi ECS. Estudo e proposta de ferramentas de comunicação e expressão em redes sociais inclusivas online. MSc dissertation. Institute of Computing, Unicamp. March 12th, 2010 (Capes scholarship).

Fortuna FJ. Normas no desenvolvimento de ambientes web inclusivos e flexíveis. MSc dissertation. Institute of Computing, Unicamp. May 14th, 2010. (MSR-FAPESP scholarship)

Neris VPA. Estudo e proposta de um framework para o design de interfaces de usuário ajustáveis. Ph.D. Thesis. Institute of Computing, Unicamp. June 14th, 2010. (FAPESP scholarship).

Miranda LC. Artefatos e linguagens de interação com artefatos digitais contemporâneos: os anéis interativos ajustáveis para a televisão digital interativa. Ph.D. Thesis. Institute of Computing, Unicamp. August 27th, 2010. (CNPq scholarship).

Almeida LDA. Awareness do espaço de trabalho em ambientes colaborativos inclusivos na web. Ph.D. Thesis. Institute of Computing, Unicamp. February 28th, 2011. (FAPESP scholarship).

Maria Cecília Calani Baranauskas

Instituto de Computação
Universidade Estadual de Campinas (Unicamp)
Av. Albert Einstein, 1251
Caixa Postal 6176
13084-971 – Campinas, SP – Brasil

cecília@ic.unicamp.br
55.19.3521-5870

ENVIRONMENTAL MONITORING AND MODELING OF THE GENETIC POTENTIAL OF SUGARCANE CULTIVARS UNDER APPROPRIATE WATER AVAILABILITY IN THE SOIL

Regina Célia de Matos Pires

Agronomic Institute of Campinas (IAC) / SAA-SP

The need for biofuels production is a growing concern on modern society due to the sustainability of the activities associated with the growth of human population and the growing economic demand. The expansion of sugarcane cultivation has been taking place in both adequate and marginal lands. When cultivation takes place on the so-called marginal lands, this terminology is used due to the characteristics related to soil and/or climate. Thus, studies that may correlate the development of new genotypes in relation to climatic changes, cultivation under DEF and under appropriate conditions of water availability in the soil are very relevant. To understand the dynamics involved in the process of water transference in the soil-plant-atmosphere system and their interactions on the productive system, with different genotypes, it becomes necessary the use of sensors for monitoring the climate, soil water and their interaction with the development of plants. These evaluations enable the development a model of plant growth to make production estimates. Currently, irrigation in Brazil is mainly used for vinasse application. The areas irrigated systemically are still just few if compared to the potential for the use of this technique. The use of irrigation in sugarcane may bring great advance on the increase of fertilization efficiency for this crop, especially when drip fertirrigation is adopted. Studies about use of sensors for monitoring are important to establish cultivation strategies for sustainability. The dynamism of process involved in the experimentation, the interface with different areas of knowledge such as irrigation, climate, soil, water, wireless sensors networks and informatics will offer conditions for the application of the knowledge in the field conditions under different ecological areas and genotypes.

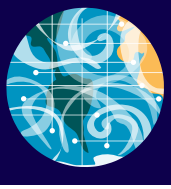
Sugarcane cultivars

Regina Célia de Matos Pires

Instituto Agronômico de Campinas (IAC)

Avenida Barão de Itapura, 1481
Caixa Postal 28
13012-970 – Campinas, SP – Brazil

rcmpires@iac.sp.gov.br
55.19.3202-1702

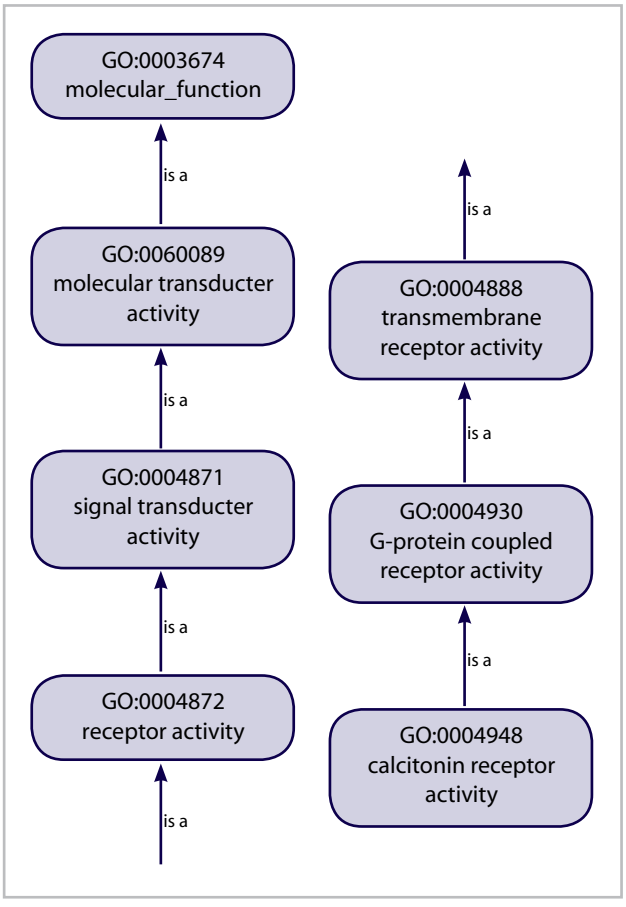


INFORMATION TECHNOLOGY APPLIED TO BIOENERGY GENOMICS: PROBABILISTIC ANNOTATION USING ARTIFICIAL INTELLIGENCE

Ricardo Zorzetto Nicoliello Vêncio

Ribeirão Preto School of Philosophy, Sciences and Literature / University of São Paulo (USP)

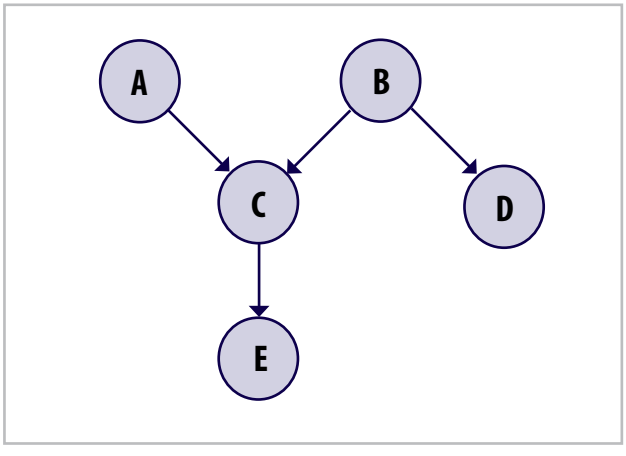
An alternative to the problem of fossil fuels depletion is the use of renewable energy. In Brazil, sugarcane (*Saccharum officinarum*) is used for years as alternative energy source and, therefore, Brazil has become a key player in alternative energy development. Our main aim is to develop methods and tools to attack some of the bioinformatics issues raised in sugarcane genomics research. In order to achieve this, we opt for Sifter (Engelhardt *et al.*, 2006), a powerful method based on Bayesian Networks. Our major aim was to establish a local implementation of the Sifter methodology for application in bioenergy related problems and following is to improve the original source code performance, potentially allowing it to be used in a genome-wide scale.



A given gene's molecular function in the Gene Ontology representation

MAIN RESULTS

We currently have completed scripts that allow full automation of the pipeline of Sifter methodology, with average performance gain of about 72.5% (quad core machine) and 67.7% (dual core) in relation to original scripts supplied with the software. To achieve this goal, we changed the originally proposed pipeline, and beyond that we added new functions to the scripts aiming user friendly software and a better detection performance, under evaluation. This new pipeline is designed to enable the analysis proposed by Sifter methodology in high-throughput analysis.



Example of Bayesian Network (BN). The nodes represent random variables and the directed edges represent statistical dependence relationships. This BN compactly represents the following probability distribution: $\mathbf{Pr}(A,B,C,D,E) = \mathbf{Pr}(E|C)\ \mathbf{Pr}(C|A,B)\ \mathbf{Pr}(D|B)\ \mathbf{Pr}(B)\ \mathbf{Pr}(A)$

PRODUCTS/PUBLICATIONS

POSTERS

Silva DCDA, Waldemarin RC, Vêncio RZN. Information technology applied to bioenergy genomics: probabilistic annotation using artificial intelligence. *X-Meeting 2010 – 6th International Conference of The Brazilian Association for Bioinformatics and Computational Biology*. November 15-18, 2010. Centro de Convenções e Artes da UFOP, Ouro Preto, Minas Gerais, Brazil.

Ricardo Zorzetto Nicoliello Vêncio

Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto
Universidade de São Paulo (USP)
Departamento de Computação e Matemática

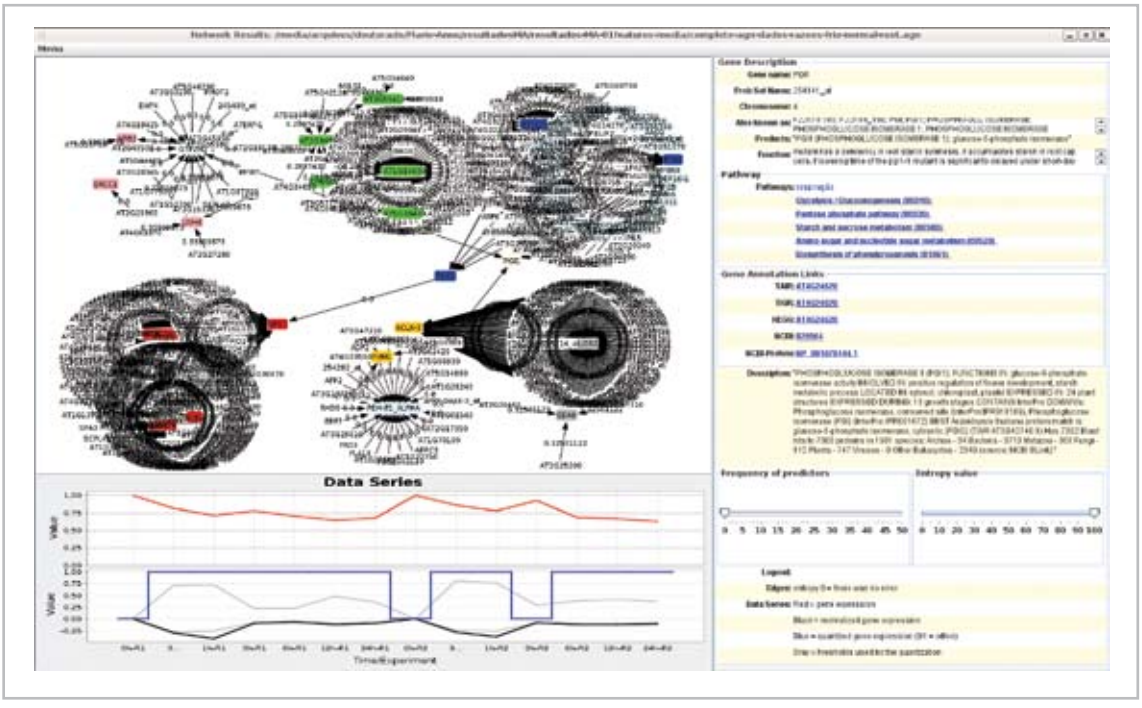
Av. Bandeirantes, 3.900 – Monte Alegre
14040-901 – Ribeirão Preto, SP – Brazil

rvencio@usp.br
55.16.3602-3718

DATA INTEGRATION IN SYSTEMS BIOLOGY: CHARACTERIZATION OF BIOLOGICAL PHENOMENA FROM STRUCTURAL AND FUNCTIONAL INFORMATION

Ronaldo Fumio Hashimoto

Institute of Mathematics and Statistics / University of São Paulo (USP)



One of the most challenging research problem of System Biology nowadays is the inference (or reverse-engineering) of gene regulatory networks (GRNs) from expression profiles. This research issue became important after the development of high-throughput technologies for extraction of gene expressions, such as DNA microarrays [74] or SAGE [84], and more recently RNA-Seq [86]. This problem regards on discover regulatory relationships between biological molecules in order to recover a complex network of interrelationships, which can reveal/describe not only diverse biological functions but also the dynamics of molecular activities. It is very important to understand how many biological processes happen and in most cases, how to prevent it from happening (diseases).

In the context of expression profiles, a big challenge that researchers need to face is the large number of variables or genes (thousands) for just a few experiments available (dozens). In order to infer relationships among those variables, it is needed a great effort in developing novel computational and statistical techniques that are able to alleviate the intrinsic error estimation committed in the presence of small number of samples with huge dimensionalities. In general, it is not possible to recover the GRNs very accurately. The main reasons for this are thee lack information about the biological organism, the high complexity of the networks and the intrinsic noise of the expression measurements. Thus, infer, analyse and compare the interrelationship between genes with precision, generating Gene Regulatory Networks (GRNs), is an open research problem.

Ronaldo Fumio Hashimoto

Instituto de Matemática e Estatística
Universidade de São Paulo (USP)
Rua do Matão, 1.010 – sala 203, bloco C
05508-090 – São Paulo, SP – Brazil

ronaldo@ime.usp.br
55.11.3091-6135

PORSIMPLES: SIMPLIFICATION OF PORTUGUESE TEXT FOR DIGITAL INCLUSION AND ACCESSIBILITY

Sandra Maria Aluísio

São Carlos Institute of Mathematics and Computer Sciences / University of São Paulo (USP)

Figure 1. Main page of Simplifica

The main goal of PorSimples was to develop Natural Language Processing (NLP) technologies related to Text Adaptation (TA) to promote digital inclusion and accessibility for people with low levels of literacy. There are two general different approaches for TA: Text Simplification and Text Elaboration. The first can be defined as any task that reduces the lexical or syntactic complexity of a text, while trying to preserve meaning and information, and can be subdivided into Lexical and Syntactic Simplification, Automatic Summarization, and other techniques. Text Elaboration aims at clarifying and explaining information and making connections explicit in a text, for example, providing definitions or synonyms for words known to only a few speakers of a language. The technologies developed in PorSimples are available by means of three systems aimed to distinct users: 1) an authoring system, called Simplifica, to help authors to produce simplified texts targeting people with low literacy levels (*figure 1*); 2) an assistive technology system, called Facilita, which explores the tasks of summarization and simplification to allow poor literate people to read Web content (*figure 2*), and 3) a web content adaptation tool, named Educational Facilita, for assisting low-literacy readers to perform detailed reading. It exhibits questions that clarify the semantic relations linking verbs to their arguments, highlights the associations amongst the main ideas of the texts and the named entities, and perform lexical elaboration. Currently, Educational Facilita only explores the NLP tasks of lexical elaboration and named entity labeling.

Website: http://caravelas.icmc.usp.br/wiki

MAIN RESULTS

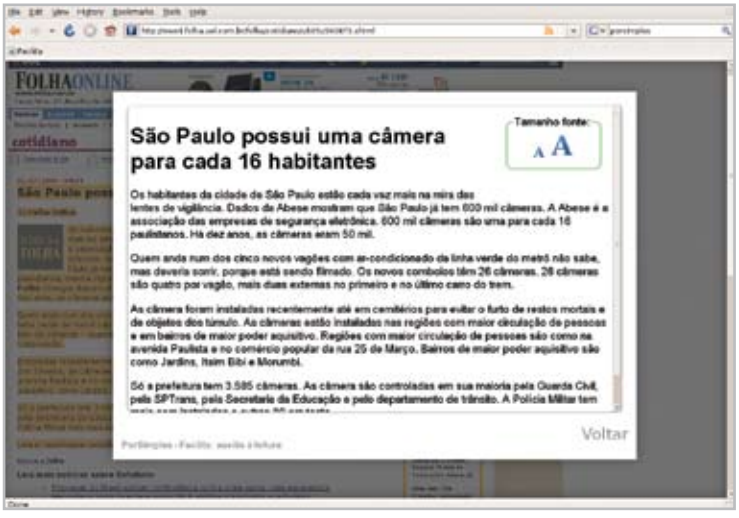


Figure 2. Facilita Demo in action: a summarized and simplified text (shown in front) of a text selected by the user from an online Brazilian newspaper (shown in back)

Watanabe WM, Cândido A, Amancio MA, Oliveira M, Pardo TAS, Fortes RPM, Aluísio SM. 2010. Adapting Web content for low-literacy readers by using lexical elaboration and named entities labeling. *New Review of Hypermedia and Multimedia*.**16**: 303-327.

Gasperin C, Maziero E, Aluísio, SM. 2010. Challenging choices for text simplification. In: *Proceedings of PROPOR 2010*, 40-50.

António Branco, Aldebaro Klautau, Renata Vieira, Vera Lúcia Strube de Lima (Eds.): Computational Processing of the Portuguese Language, 9th International Conference, PROPOR 2010, Porto Alegre, RS, Brazil. April 27-30, 2010. *Proceedings. Spring, 2010.* **6001**: 40-50.

Aluísio SM, Gasperin C. 2010. Fostering digital inclusion and accessibility: the PorSimples project for simplification of portuguese texts. *Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas*. New York: ACL, 2010. **1**: 46-53.

Aluísio SM, Specia L, Gasperin C, Scarton CE. Readability assessment for text simplification. In: NAACL 5th Workshop on Innovative Use of NLP for Building Educational Applications (BEA-2010), 2010, Los Angeles. *Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications*. New York: ACL, 2010. **1**: 1-9.

Scarton CE, Gasperin C, Aluísio SM. Revisiting the readability assessment of texts in portuguese. In: IBERAMIA 2010, 2010, Bahia Blanca. *Lecture Notes in Computer Science. Heidelberg: Springer, 2010.* **6433**: 306-315.

Watanabe WM, Candido Jr. A, Uzêda V, Fortes RPM, Pardo TAS, Aluísio SM. Facilita: reading assistance for low-literacy readers. In: *Proceedings of ACM SIGDOC 2009 - ACM International Conference on Design of Communication*, 2009, Bloomington, In. **1**: 29-36.

Scarton CE, Aluísio SM. 2010. Análise da inteligibilidade de textos via ferramentas de processamento de língua natural adaptando as métricas do Coh-Metrix para o português. *Linguamática* (Online journal about natural language processing of Iberic languages). **2(1)**: 45-61.

Candido Jr. A, Maziero E, Gasperin C, Pardo T, Specia L, Aluísio S. 2009. Supporting the adaptation of texts for poor literacy readers: a text simplification editor for brazilian portuguese. In: *Proceedings of the NAACL HLT Workshop on Innovative Use of NLP for Building Educational Applications*, 34-42.

Gasperin C, Specia L, Pereira T, Aluísio S. Learning when to simplify sentences for natural text simplification. In: *CSBC 2009 - XXIX Congresso da Sociedade Brasileira de Computação*, 2009, Bento Gonçalves, RS. ENIA 2009 – Encontro Nacional de Inteligência Artificial, 2009. 809-818.

PORSIMPLES IN NUMBERS

- Started in November 2007 and finished in April 2010
- Team:
  6 researchers/students supported by MSR-FAPESP
  11 other students joint the project
- Publications:
  28 papers (conferences and journal)
  6 Demos/Posters (short papers)
  12 Technical Reports
  1 submitted paper
- Research Collaborations: 13 senior researchers from
  Psycholinguistics
  Statistics
  Natural Language Processing
  Human-Computer Interaction
- Products:
  3 main systems
  6 types of text adaptation methods
  4 data resources
  3 supporting tools

Sandra Maria Aluísio

Instituto de Ciências Matemáticas e de Computação
de São Carlos – Universidade de São Paulo (USP)
Av. Trabalhador Sãocarlense, 400 – Centro
Caixa Postal 668
CEP 13560-970 – São Carlos, SP –Brasil

sandra@ icmc.usp.br
55.16.33739663