INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team AxIS

Usage-centered design, analysis and improvement of information systems

Sophia Antipolis - Méditerranée, Paris - Rocquencourt

Theme : Knowledge and Data Representation and Management

Table of contents

1. Team **1**
2. Overall Objectives **1**
2.1. Overall Objectives 1
2.2. Highlights 2
3. Application Domains **3**
3.1. Panorama - Sustainable Development for Smart Cities 3
3.2. Transportation Systems & Mobility 3
3.3. Tourism 4
3.4. e-participation in Environment and Health 5
4. Software **5**
4.1. Introduction 5
4.2. Data Mining and Web Usage Mining 6
4.2.1. Clustering and Classification Toolbox 6
4.2.2. Clustering Methods for mining Sequential Patterns in Data Streams 6
4.2.3. AWLH for Pre-processing Web Logs 7
4.2.4. Two Methods for Extracting Sequential Patterns with Low Support 7
4.2.5. ATWUEDA for Analysing Evolving Web Usage Data 7
4.3. Information Retrieval 8
4.3.1. CBR*Tools for Managing and Reusing Past Experiences based on Historical Data 8
4.3.2. Broadway*Tools for Building Recommender Systems on the Web 8
4.4. Other Software 8
4.4.1. K-MADe: Kernel of Model for Activity Description Environment 8
4.4.2. CLF - Computer Language Factory 8
4.4.3. BibAdmin 9
5. New Results **9**
5.1. Introduction 9
5.2. Clustering and Optimal Segmentation of Curves 9
5.3. Block clustering and Web Content Data Mining 10
5.4. Using Multiple Dissimilarity Data Tables for Documents Categorization 10
5.5. Fuzzy Clustering on Multiple Dissimilarity Matrices 11
5.6. Anti-Bouncing Model for Usage Data Streams 11
5.7. Discovering Informative Feature Set over High-dimensions 11
5.8. Discovering Evolution Patterns from Satellite Image Time Series 12
5.9. User Oriented Expert Finding 12
5.10. Using Web Graph Structure for Person Name Disambiguation 12
5.11. Latent Variable Models for Expert Finding 13
5.12. Living Lab Landscape of Research Streams: towards User Co-creation 13
5.13. Future Internet Domain Landscape 13
5.14. Comparison of Usability Methods: Inspection vs. User Testing 14
5.15. Formal Design and Validation of the Dialogue in Interactive Software 15
6. Contracts and Grants with Industry **16**
6.1. Orange Labs CRE (2009-2010) 16
6.2. Cassette Voyage (2009-2010) 16
7. Other Grants and Activities **17**
7.1. Regional Initiatives 17
7.1.1. ICT Usage Lab: Activities & projects, Terriotries, Campus STIC Usages 17
7.1.2. PACA CPER TELIUS: FocusLab Platform 19
7.1.3. PREDIT ADEME: TicTac (jan 2010- jan 2012) 19
7.1.4. PacaLabs: ECOFFICES Project (aug 2010- sep 2011) 20

7.1.5. PacaLabs: HOTEL-REF-PACA Project (dec 2010- dec 2011) 20
7.1.6. Color Inria Action: Open Data - CDISOD (2010) 20
7.1.7. Collaborations with SMEs 21
7.2. National Initiatives 21
7.2.1. Introduction 21
7.2.2. ANR: MIDAS Project (2008-2010) - extended to June 2011 21
7.2.3. ANR: MyCitizSpace Project(2008-2010)- extended to June 2011 22
7.2.4. ANR: PIMI Project (end 2010- end 2013) 22
7.2.5. Web 2.0: Scar Project (nov 2009- nov 2011) 23
7.2.6. Web 2.0: 2.0 Process Project (nov.2009-nov 2011) 23
7.2.7. ICT Usage Lab, member of the French Network of Living labs 23
7.2.8. National Inria projects 24
7.2.9. Other Collaborations 24
7.3. European Initiatives 24
7.3.1. Introduction 24
7.3.2. ICT Usage Lab: LLSS2010 Summer School and EnOLL association 25
7.3.3. FP7 ICT objective 1.3: ELLIOT STREP 25
7.3.4. FP7 ICT objective 1.6: FIREBALL Coordination Support Action 26
7.3.5. EU-Asian IDEAS Project (2010-2012) 26
7.3.6. COST TwinTide Action 27
7.3.7. Others initiatives 27
7.3.8. Other Collaborations 28
7.4. International Initiatives 28
7.4.1. Introduction 28
7.4.2. Brazil: FACEPE-INRIA 28
7.4.3. Morocco : the WRUM Project 28
7.4.4. Tunisia: STIC program 29
7.4.5. USA: Explorer Program 29
7.4.6. Participation to Standards in Ergonomics 29
7.5. Animation of the Scientific Community 30
7.5.1. Introduction 30
7.5.2. Reviewing Activities 30
7.5.3. Organization of Conferences / Workshops 32
7.5.4. University Teaching 32
7.5.5. Ph.D. Thesis 32
7.5.6. Internships 33
7.5.7. Activities of General Interest 33
8. Bibliography **34**

1. Team

Research Scientists

Brigitte Trousse [CR1 Inria, Team Leader]
Yves Lechevallier [DR2 Inria, Vice-Team Leader]
Thierry Despeyroux [CR1 Inria, part-time 70%]
Florent Masseglia [CR1 Inria, HdR]
Bernard Senach [CR1 Inria]
Dominique Scapin [DR2 Inria]

External Collaborator

Marc Csernel [Associate Professor, Univ. Paris IX Dauphine until August, Inria Delegation since september]

Technical Staff

Pascal Marie-Dessoude [myCitySpace and Process 2.0 Contracts]
Nicolas Béchet [Addictrip Contract since September]
Marc Pallot [Fireball Contract, 20 % part-time from September]
Guillaume Pilot [Tic Tac Contract, from September 14th]
Mylène Leitzelman [Elliot Contract, 40 % part-time from September 23rd]
Carole Goffart [Ecoffices Contract, from October 11th]
Nelly Bessega [Mymed Contract & Contract Management Assistant, part-time from November 8th]
Anne-Laure Negri [Scar Contract, 80 % part-time from December 1st]
Florian Bonacina [Scar and Hotel-ref-Paca Contracts, from December 1st]

PhD Students

Malika Charrad [PhD, INS & ENSI Tunisia & CNAM Paris, until june, visit in March]
Elena Smirnova [Funded by CORDIS, Univ. Nice Sophia Antipolis (UNS-STIC)]
Chongsheng Zhang [Funded by ANR, ANR Midas, Univ. Nice Sophia Antipolis (UNS-STIC)]

Visiting Scientists

Hajer Zghal [Associate Professor, Univ. de la Manouba, Tunisie, January 11th - 14th]
Henda Ben Ghezala [Professor, Univ. de la Manouba, Tunisie, January 11th - 14th]
F.A.T De Carvalho [Professor, Federal Univ. of Pernambuco, Brazil, January 17 - 23 and July 2 - 31]
Carlos Eduardo Ribeiro de Mello [PhD, Federal Univ. of Pernambuco, Brazil, until March 31th]
Amar Rebbouh [Associate Professor, USTHB, Algeria, April 6th - 16th]
Sergio Queiroz [Prof., Federal Univ. of Pernambuco, Brazil, July 2 - 31]

Administrative Assistants

Stephanie Chaix [Temporary staff, part-time January]
Rosanne Bonin [Temporary staff, part-time between February and August]
Stephanie Aubin [TR Inria, part-time between September and November]
Florence Barbara [AI Inria, part-time 50 % since December]
Anais Cassino [AI Inria, part-time until end of November]
Delphine Robage [Temporary staff, part-time from December 1st]

Others

Amine Louati [STIC program, Master Internship, ENSI, Tunisie, between March and August]
Maurice Yared [Midas Contract, Internship, July-September]

2. Overall Objectives

2.1. Overall Objectives

AxIS is carrying out research in the area of Information and Knowledge Systems (ISs) with a special interest in evolving large ISs such as Web based-information Systems. Our ultimate goal is to contribute to user-driven open innovation as a way to foster innovation, to improve the overall quality of ISs, to support designers during the design process and to ensure ease of use to end users.

We are convinced that to reach this goal, according to the constant evolution of actual and future ISs, it is necessary to involve the users in the design process and to empower them, so that they can become co-designers. This is a new way to anticipate the usage and its analysis and also to consider maintenance very early in the design process.

To achieve such a research, we have set up in July 2003 a multidisciplinary team that involves people from different computer sciences domains (Artificial Intelligence, Data Mining & Analysis, Software Engineering, Document Management from 2004) and at the end of 2005 from Ergonomics (Human Sciences), all of them focusing on information systems. Our goal is of course to improve **efficiency of machine learning and data mining methods** but also to improve the **quality of results**. The originality of AxIS project-team is to adopt a **cognitive and inter-disciplinary approach for the whole KDD[1] process** and for each step (preprocessing, data mining, interpretation).

To address this challenge, relying on our scientific foundations (see our 2007 activity report, Section Scientific Foundations), we had a first 4 years steps dedicated to the design of methodological and technical building blocks for IS mining (usage, content and structure).

Our researches are organised to support the disruptive process of continuous innovation.

In this continuous process: design is never ended and relies on very short test-adapt-test cycles where users are co-designers: they can contribute to design before/after market launch as ideas providers, as participants in test beds or field experimentations or even as solution providers when they are given the convenient tools.

To support this process, a large collection of tools and methods are needed and numerous efforts have already been engaged at european level to provide infrastructures for experimentations (for instance the Future Internet Research & Experimentation (FIRE) initiative launched in summer 2008), tools for creativity or sharing (Laboranova, CoSpaces, etc.).

In this context, our team focuses its effort on the technical and methodological environment needed to extract meaning from the huge amount of data issued from large and distributed information systems.

Our researches are organised in three research topics:

1. **Topic 1 - Data Mining and IS mining**: Mining complex data and IS data, mainly temporal and spatial data, semantic Web mining (ontologies and Web mining) and semantics checking of an evolving IS[2]. Most of the effort is put into two problems related to mining temporal data: a) analysing the evolution of user behaviours and b) summarizing and mining data streams.
2. **Topic 2 - IS Mining based services for supporting Information Retrieval**: Mining collective usage data, mining social networks, community detection, expert finding, collaborative filtering based recommender systems for information retrieval, bookmark management, social networks based recommender systems, personalization, etc.
3. **Topic 3 - Pluridisciplinary Research for the development of the FocusLab platform in Living Labs**: Methods and tools based on a multidisciplinary approach (Social and Human Science and ICT) for the design and the evaluation of innovative services and for user-driven open innovation, Towards the Focus methodological and technical experimentation platform...

2.2. Highlights

- **Simon Régnier PhD prize**: Alzennyr Gomes Da Silva was the laureate of the Simon Régnier PhD thesis Prize [80] during the 27th conference of SFC[3] at Saint Denis (La Réunion). This prize was for her thesis defended on september 2009 and supervised by Y. Lechevallier (Inria) and Edwin Diday (University Paris IX dauphine). Title: Evolutive Data Analysis - application to Web Usage Data.

[1] **KDD: Knowledge Discovery From Databases**
[2] IS: Information System
[3] SFC:Francophon Society of Clustering

- **Finalist at Galileo GNSS Living Lab Prize**: B. Trousse and B. Senach (Inria) with G. Gallais and D. Emsellem from Vulog were one of the ten finalists of the european GNSS Living Lab Prize among 57 submissions.
- **COMPSTAT 2010**: Y. Lechevallier was vice-chair of the local organizing committee of the 19th International Conference on Computational Statistics (COMPSTAT) [53] in August 23-27 (CNAM, Paris) with more than **600 participants from 24 countries**. Site: http://www.compstat2010.fr/
- **LLSS 2010**: B. Trousse with K. Pawar (University of Nottingham) and R. Santoro (EsoceNet) was at the initiative of the First Living Lab Summer School School which was co-organised at the Cité des Sciences (Paris) by the two french living labs (ICT Usage Lab, Lutin Usage Lab), EsoceNet, universcience and Unbla. It was a great success with around **80 participants from 23 countries**. B. Trousse was vice-chair of the scientific committee and chair of the local organizing committee.
- **New grants on Topic 3**: Topic 3 has been developed thanks to six new grants where usage analysis and users'involvement are key issues: 1 PREDIT (TIC TAC), 2 PACALabs (Ecoffices, Hotel-Ref-Paca), 2 european contracts (Elliot, Fireball) and 1 ANR (PIM).
- **1 new contract with Industry (CRE Orange Labs** and new collaborations with numerous SMes are starting this year (Perferencement, Osmose, Wozaik, Cassette Voyage, VUlog etc.).

3. Application Domains

3.1. Panorama - Sustainable Development for Smart Cities

The project addresses applicative field which has the following features

a) requiring usage/data storage, preprocessing and analysis

- for designing, evaluating and improving huge evolving hypermedia information systems (mainly Web-based ISs), for which end-users are of primary concern,
- for a better understanding of service/product used with data mining techniques and knowledge management (for example in Transportation systems, Tourism, Security and Anomaly Detection, Internet of Objects, etc.),
- for social network analysis (for example in Web 2.0 applications, Business Intelligence, Sustainable Development, etc.)

and b) requiring user-driven innovation methods.

Even if our know how, methods and algorithms have a cross domain applicability, our team chooses to focus on **Sustainable Development for Smart Cities** with a particular stress put on three main domains:

- Transportation systems & Mobility (cf. 3.2),
- Tourism (cf. 3.3),
- e-participation in Environment and Health (cf. 3.4).

.

3.2. Transportation Systems & Mobility

Major recent evolutions in Intelligent Transportation Systems (ITS) are linked to rapid changes in communication technologies, such as ubiquitous computing, semantic web, contextual design. A strong emphasis is now put on mobility improvements. These improvements concern both the quality of traveller's information systems for trip planning, the ability to provide real time recommendations for changing transportation means according to traffic information, and the quality of embedded services in vehicles to provide enhanced navigation aids with contextualised and personalised information.

Web 2.0 technology plays now a role of growing importance, as it supports users feed-back which becomes a mean for improving quality of travelers information sytems. Exchange of information between users about delay, cancellation and others occurring events provides more accurate and precise information on the current state of the transportation system.
Let us cite various projects where AxIS was and is still involved:

- **Mining Mobility Data**, PREDIT (2004-2007): the MobiVIP project has been an opportunity to collaborate with local Institutions (*Communauté d'Agglomération de Sophia Antipolis - CASA*) and SMEs (VU Log) and apply AxIS know-how in mining spatial and temporal data issued from vehicules equipped with GPS and from the reservation server and in clustering trajectories (with semantic distances). Even if we didn't apply our know how in mining data streams in this project, this will be crucial in the future with more and more equipped vehicules with GPS.
- **Traveller's Information Systems** - evaluation of two Web sites (2007-2008):
  - the Envibus web site provides information about a bus network ; its evaluation was done by coupling ergonomic analysis and usage mining
  - the Otto&co web site support car-sharing our cooperation about car-sharing done in 2008 with an evaluation of the Otto&co site in the context of the action COLOR Cuscov is still lasting.
- **Advanced Transportation Systems - Multimodality**, PREDIT (2010-2012): the TIC TAC project (cf. 7.1.7) aims to optimize travel time by providing in an area with weak transportation services, a just in time on demand shuttle, based on real time information.

Moreover Axis has been requested to participate to the CLAIRE-SITI submission related to transportation and managed by G. Scemama (Inrets) for **Equipex** call . In case of acceptation, AxIS might contribute from 2012-2013 on **Mobility 2.0** for some methods and **Web 2.0 tools for capturing** and analysing **user feedback**.

3.3. Tourism

As tourism is a highly competitive domain, local tourism authorities have developed Web sites in order to promote their offer of services to citizens. Unfortunately, the way information is organised does not necessarily meet Internet users expectations and numerous improvements are necessary to enhance their understanding of visited sites. Thus, even if only for economical reasons, the quality and the diversity of tourism packages have to be improved, for example by highlighting cultural heritage.

Again to illustrate our role in such a domain, let us cite some past and current projects where AxIS is involved related mainly to **Semantic Web Mining**[4] and **Information Retrieval**.

- Researches have been carried out using log files from the city of Metz. This city was chosen because its Web site is in constant development and has been awarded several times, notably in 2003, 2004 and 2005 in the context of the Internet City label. The objective was to extract information about tourists behaviours from this site log files and to identify possible benefits in designing or **updating a tourism ontology** [103].
- **Providing Tourism Information linked to Transportation information**: AxIS has already studied recommender systems in order to provide users with personalised transportation information while looking for tourism information such as cultural information, leisure etc (cf. our recommender Be-TRIP (2006)) and [87].
- This year, the Pacalabs (call 2) proposal called HOTEL-REF-PACA was accepted. Its goal is **to better refer the web sites of hotels** from the region of TOURVAL in PACA (mainly *Vésubie* territory), with an approach based on a better understanding of usage from the internauts.
- We submitted a proposal called "Scientific Excellence in Tourism" (X2T) to Interreg Alcotra with CRT Tourism Paca, Tourism Offices of Digne and Aix en Provence and Provincia of Imperia (Italy).

[4] By Semantic Web Mining, we mean the mutual benefits between two communities *Semantic Web* and *Web Mining*. In our case, we exploit a) ontologies and semantic data for improving usage analysis, personnalised services, the quality of results of search engines and for checking the content of an IS and also b) we exploit usage data for updating ontologies.

3.4. e-participation in Environment and Health

Following the Rio Conference (1992) and the Agenda for the 21st Century, local territories are now directly concerned with the set up of actions for a sustainable development. In this frame, ICT tools have been supposed to be very efficient to re-engage people in the democratic process and to make decision-making more transparent, inclusive and accessible. So, sustainable development is closely associated with citizen participation. The emerging research field of e-democracy (so called Digital Democracy or eParticipation), concerned with the use of communications technologies such as the internet to enhance the democratic processes is now a very active field. Though still in its infancy, a lot of literature is already available (see for instance: http://itc.napier.ac.uk/ITC/publications.asp or http://www.demo-net.org/ for a global view of work in Europe) and numerous different topics are addressed in the field.

Below are some topics where AxIS was or is involved in:

- **Preprocessing and analysing collective usage data and social networks** from group discussions related to design process : in the CDISOD proposal (Citizen Driven Innovation Services based on Open data (cf. 7.1.6), citizens have to design innovative services based on public data. In the ELLIOT project (cf. 7.3.3), citizen will have fixed and mobile sensors and they will contribute to a better collect of environmental data and will define specific information services according to their needs.
- **Methods and tools for open innovation**: we submitted a proposal for supporting the design of innovative services by citizens from public data in collaboration with Fing (Marseille) and Ademe (Sophia Antipolis) for the Inria Color call for actions.

AxIS main topics relevant for these domains are: social network analysis, personalization and information retrieval, recommender systems, expert search, design and evaluation of methods and tools for open innovation, usage mining, mining data streams.

4. Software

4.1. Introduction

From its creation, AxIS has developed several softwares validated experimentally on various applications::

- Data Mining and Web Usage Mining
  - Clustering and Classification toolbox (cf. 4.2.1),
  - SMDS, SCDS and ICDS, three clustering softwares for mining sequential patterns in data streams (cf. 4.2.2).
  - AWLH (AxIS Web Log House) for preprocessing Web logs (cf. 4.2.3)
  - Cluster&Divide and Divide&Cluster, two methods for extracting Sequential Patterns with Low Support from Web logs (cf. 4.2.4)
  - ATWUEDA (Axis Tool for Web Usage Evolving Data Analysis) (cf. 4.2.5)
- Information Retrieval
  - CBR*Tools, an object-oriented platform for reusing experiences (requiring the management of historical data) (cf. 4.3.1),
  - Broadway*Tools for designing adaptive Web-based recommendation systems and collaborative information retrieval support (cf.4.3.2).
- Other softwares (cf. 4.4): a) K-MADe, Kernel of Model for Activity Description environment, b) CLF for generating efficient parsers and c) Bibadmin for the management of a collection of publications.

In the context of the FocusLab platform (CPER Télius 7.1.2), we are studying the architecture of the software part and as a starting point the integration of relevant AxIS softwares and different ways to fund human ressources for the development..

4.2. Data Mining and Web Usage Mining

4.2.1. Clustering and Classification Toolbox

Participants: Marc Csernel, Yves Lechevallier [co-correspondant], Brigitte Trousse [co-correspondant].

We developed and maintained a collection of clustering and classification software, written in C++ and/or Java:

- a Java library (Somlib) that provides efficient implementations of several SOM variants [78], [77], [100], [99], [104], especially those that can handle dissimilarity data (available on Inria's Gforge server http://gforge.inria.fr/projects/somlib/, developed by AxIS Rocquencourt and Brieuc Conan-Guez from Université de Metz.
- a functional Multi-Layer Perceptron library, called FNET, that implements in C++ supervised classification of functional data [95], [98], [97], [96] (developed by AxIS Rocquencourt).
- two partitioning clustering methods on the dissimilarity tables issued from a collaboration between AxIS Rocquencourt team and Recife University, Brazil: CDis and CCClust [83]. Both are written in C++ and use the "Symbolic Object Language" (SOL) developed for SODAS.
- two improved and standalone versions of SODAS modules, SCluster and DIVCLUS-T [74] (AxIS Rocquencourt).
- a Java implementation of the 2-3 AHC (developed by AxIS Sophia Antipolis). The software is available as a Java applet which runs the hierarchies visualization toolbox called HCT for Hierarchical Clustering Toolbox (see [75]).

A Web interface developed in C++ and running on our Apache internal Web server.is available for the following methods: SCluster, Div, Cdis, CCClust.
Previous versions of the above software have been integrated in the SODAS 2 Software [93] which was the result of the european project ASSO[5] (2001-2004). SODAS 2 software supports the analysis of multidimensional complex data (numerical and non numerical) coming from databases mainly in statistical offices and administration using Symbolic Data Analysis [69]. This software is registrated at APP. The latest executive version of the SODAS 2 software, with its user manual can be downloaded at http://www.info.fundp.ac.be/asso/sodaslink.htm. See 2009 AxIS annual report for more details of the main contributions of AxIS to SODAS [79], [105] which have been registered at APP.

4.2.2. Clustering Methods for mining Sequential Patterns in Data Streams

Participants: Maurice Yared, Florent Masseglia, Brigitte Trousse [correspondant], Yves Lechevallier.

As a result of Marascu's thesis (2007-2009) [91], a collection of softwares have been developed for knowledge discovery and security in data streams (cf. our 2009 annual report for more details on WOD, the outlier detection method and GEAR an implementation of the history management strategy).
Three **clustering methods for mining sequential patterns** (Java) in data streams have been developped in Java by A. Marascu during her thesis [91]. The softwares take batches of data in the format "Client-Date-Item" and provide clusters of sequences and their centroids in the form of an approximate sequential pattern calculated with an alignment technique.

- SMDS compares the sequences to each others with a complexity of $O(n2)$.
- SCDS is an improvement of SMDS, where the complexity is enhanced from $O(n2)$ to $O(n.m)$ with n the number of navigations and m the number of clusters.
- ICDS is a modification of SCDS. The principle is to keep the clusters' centroids from one batch to another.

[5] ASSO: Analysis System of Symbolic Official data

This year, the Java code of SMDS has been integrated in the MIDAS demonstrator [68].(cf. 7.2.2) and a C++ version [61] has been implemented for the CRE contract with Orange Labs with a visualisation module (in Java) (cf. 6.1). SMDS has been applied on data issued from mobile Orange portal.

4.2.3. AWLH for Pre-processing Web Logs

Participants: Yves Lechevallier [co-correspondant], Brigitte Trousse [co-correspondant].

AWLH is issued from AxISlogminer preprocessing software which implements the mult-site log preprocessing methodology developed by D. Tanasa in his thesis [15] for Web Usage Mining (WUM). In the context of the Eiffel project (2008-2009), we isolated and redesigned the core of AxISlogMiner preprocessing tool (we called it AWLH) composed of a set of tools for pre-processing web log files. AWLH can extract and structure log files from several Web servers using different input format. The web log files are cleaned as usually before to be used by data mining methods, as they contain many noisy entries (for example, robots bring a lot of noise in the analysis of user behaviour then it is important in this case to identify robot requests). The data are stored within a database whose model has been improved.

Now the current version of our Web log processing offers:

- Processing of several log files from several servers,
- Support of several input formats (CLF, ECLF, IIS, custom, ...);
- Incremental pre-processing;
- Java API to help integration of AWLH in external application.

For recording the click actions by a user in a real time, we developed in 2009 a tool based on an open source project called "OpenSymphony ClickStream" for capturing Web user actions. For capturing and structuring data issued from annotated documents inside discussion forums, an extended version of AWLH has been developed.

4.2.4. Two Methods for Extracting Sequential Patterns with Low Support

Participants: Brigitte Trousse [correspondant], Florent Masseglia.

Two methods for extracting sequential patterns with low support have been developed by D. Tanasa in his thesis [102] in collaboration with F. Masseglia and B. Trousse : **Cluster & Divide** [102] and **Divide & Discover** [13], [102].
See Chapter 3 of Tanasa's PhD document for more details on these two methods and on a framework for developing methods for extracting sequential patterns with low support.

4.2.5. ATWUEDA for Analysing Evolving Web Usage Data

Participant: Yves Lechevallier [correspondant].

ATWUEDA [82] for Web Usage Evolving Data Analysis was developed by A. Da Silva in her thesis [80]. It is available at INRIA's gforce website: http://gforge.inria.fr/projects/atwueda/. A. Da Silva presented part of her work in a working research group at CNAM-Paris [81].

This tool was developed in Java and uses the JRI library in order to allow the application of R functions in the Java environment. R is a programming language and software environment for statistical computing (http://www.r-project.org/. The ATWUEDA tools is able to read data from a cross table in a MySQL database, split the data according to the user specifications (in logical or temporal windows) and then apply the approach proposed in the Da Silva's thesis in order to detect changes in dynamic environment. The proposed approach characterizes the changes undergone by the usage groups (e.g. appearance, disappearance, fusion and split) at each timestamp. Graphics are generated for each analysed window, exhibiting statistics that characterizes changing points over time.

4.3. Information Retrieval

4.3.1. CBR*Tools for Managing and Reusing Past Experiences based on Historical Data

Participant: Brigitte Trousse [correspondant].

CBR*Tools [88], [89] is an object-oriented framework [90], [85] for Case-Based Reasoning which is specified with the UMT notation (Rational Rose) and written in Java. It offers a set of abstract classes to model the main concepts necessary to develop applications integrating case-based reasoning techniques: case, case base, index, measurements of similarity, reasoning control. It also offers a set of concrete classes which implements many traditional methods (closest neighbors indexing, Kd-tree indexing, neuronal approach based indexing, standards similarities measurements). CBR*Tools currently contains more than 240 classes divided in two main categories: the core package for basic functionality and the time package for the specific management of the behavioral situations. The programming of a new application is done by specialization of existing classes, objects aggregation or by using the parameters of the existing classes.

CBR*Tools addresses application fields where the re-use of cases indexed by behavioral situations is required. The CBR*Tools framework was evaluated via the design and the implementation of several applications such as Broadway-Web, Educaid, BeCKB, Broadway-Predict, e-behaviour and Be-TRIP.

CBR*Tools is concerned by two past contracts: EPIA and MobiVIP.

CBR*Tools will be available for research, teaching and academic purpose via the FOCUS platform. The user manual can be downloaded at the URL: http://www-sop.inria.fr/axis/cbrtools/manual/.

See also the web page http://www-sop.inria.fr/axis/cbrtools/manual/..

4.3.2. Broadway*Tools for Building Recommender Systems on the Web

Participant: Brigitte Trousse [correspondant].

Broadway*Tools is a toolbox supporting the creation of adaptive recommendation systems on the Web or in a Internet/Intranet information system. The toolbox offers different servers, including a server that computes recommendations based on the observation of the user sessions and on the re-use of user groups' former sessions. A recommender system created with Broadway*tools observes navigations of various users and gather the evaluations and annotations of those users, to draw up a list of relevant recommendations (Web documents, keywords, etc).

Based on Jaczynski'thesis [88], different recommender systems have been developed for supporting Web browsing, for supporting browsing inside a Web-based information system or for supporting query formulation in the context of a meta search engine.

4.4. Other Software

4.4.1. K-MADe: Kernel of Model for Activity Description Environment

Participant: Dominique Scapin [correspondant].

The **K-MADe tool** is intended for people wishing to describe, analyze and formalize the activities of human operators, of users, in environments (computerized or not), in real or simulated situation; in the field, or in the laboratory. Although all kinds of profiles of people are possible, this environment is particularly intended for ergonomics and HCI (Human Computer Interaction) specialists. It has been developed through collaboration between http://www.lisi.ensma.fr/ and INRIA. A new release has been delivered on november 1st 2010. It incorporates the findings from the work of Caffiau and al. (cf.5.15)). Its history, documentation and tool are available at: http://kmade.sourceforge.net/index.php.

4.4.2. CLF - Computer Language Factory

Participant: Thierry Despeyroux [correspondant].

CLF is a toolbox designed to ease the development of efficient parsers in Prolog. It currently contains a couple of tools. The first one uses Flex to perform lexical analysis and the second is an extension of Prolog DCGs [76], [94], [72] to perform syntactical analysis. It allows right recursion, take advantage of hash-coding of prolog clauses by modern prolog compilers and keep an automatic link to the source code to ease the development of tools as compilers with accurate error messages.

This toolbox has been used to produce a parser for XML. It has also been used to produce the specification formalism SeXML. The generated parsers have been intensively used in our team to parse and analyze XML files, mainly related to our research applied to the Inria annual activity reports.

A complete documentation is available in [84].

4.4.3. BibAdmin

Participant: Brigitte Trousse [correspondant].

BibAdmin developed by S. Chelcea (ex-PhD student) is a publication management tool corresponding to a collection of PHP/MySQL scripts for bibliographic (Bibtex) management over the Web. Publications are stored in a MySQL database and can be added/edited/modified via a Web interface. It is specially designed for research teams to easily manage their publications or references and to make their results more visible. Users can build different private/public bibliographies which can be then used to compile LaTeX documents. BibAdmin is made available since the end of 2005 under the GNU GPL license on INRIA's GForge server.

5. New Results

5.1. Introduction

This year we obtained new results in our three research topics:

1. **Topic 1 - Data Mining and IS mining** (cf. the first seven sections): Two thesis (Charrad and Chongsheng), clustering (Web content data clustering, document categorization, fuzzy clustering, clustering of curves), anti-bouncing model for usage data streams, summarising and mining data streams.

   Let us note that some past works on this topic described in previous annual reports were published in 2010: a) 2007 report in section 6.3.7 [19], b) 2009 report in section 5.3 [33], [46] and finally 2009 report in section 5.4 [50].

2. **Topic 2 - IS mining based services for supporting Information Retrieval**: our three results related to people search (user-oriented expert finding, person name disambiguation, models for expert finding) take place in the context of Elena Smirnova's thesis.

3. **Topic 3 - Pluridisciplinary Research for supporting the development of the FocusLab platform** (cf. the last four results):this pluridisciplinary research is dedicated to the design, tailoring and refinement of methodologies and tools for a better users' involvement in innovation processes. On going work concerns four topics: the Living Lab landscape, the convergence between research in Future Internet and Living labs domains for application in Smart Cities, the comparison of usability methods and a step towards the use of formal tools for design and validation of the dialogue in interactive software.

5.2. Clustering and Optimal Segmentation of Curves

Participant: Yves Lechevallier.

Functional Data Analysis is an extension of traditional data analysis to functional data. In this framework, each individual is described by one or several functions, rather than by a vector of R^n. This approach allows to take into account the regularity of the observed functions.

In 2010, we have continued our work on exploratory analysis algorithm for functional data in collaboration with F. Rossi, G. Hebrail from Telecom Paris Tech. Our method [21], [31] partitions a set of functions into K clusters and represents each cluster by a simple prototype (e.g., piecewise constant). The total number of segments in the prototypes, P, is chosen by the user and optimally distributed among the clusters via two dynamic programming algorithms. The main idea is to provide the analyst with a summary of the set with a manageable complexity.

We propose to merge the two approaches: we build a K-means like clustering of a set of functions in which each prototype is given by a simple function defined in a piecewise way. The input interval of each prototype is partitioned into sub-intervals on which the prototype assumes a simple form. Using dynamic programming, we obtain an optimal segmentation for each prototype while the number of segments used in each cluster is also optimally chosen with respect to a user specified total number segments. In the case of piecewise constant prototypes, a set of functions is summarized via $2P - K$ real values, where K is the number of prototypes and P the total number of segments used to represent the prototypes.

5.3. Block clustering and Web Content Data Mining

Participants: Malika Charrad, Yves Lechevallier.

Simultaneous clustering, usually designated by biclustering, co-clustering or block clustering, is an important technique in two way data analysis. The goal of simultaneous clustering is to find sub-matrices, which are subgroups of rows and subgroups of columns that exhibit a high correlation. Our aim is to analyze textual data of a web site. Our approach [27], [28] consists of three steps: Web pages classification, preprocessing of web pages content and block clustering. The first step consists in classifying web site pages into to major categories: auxiliary pages and content pages. In the second step, web pages content is preprocessed in order to select descriptors to represent each page in the web site. As a result, a matrix of web site pages and vectors of descriptors is constructed. In the last step, a simultaneous clustering is applied to rows and columns of this matrix to discover biclusters of pages and descriptors.

One of the major problems of simultaneous clustering algorithms, similarly to the simple clustering algorithms, is the need of specifying the optimal number of clusters. This problem has been subject of wide research. Numerous strategies have been proposed for finding the right number of clusters. However, these strategies can only be applied with one way clustering algorithms and there is a lack of approaches to find the best number of clusters in block clustering algorithms.

Ms Malika Charrad [17] defended her PhD in June 2010 at CNAM.

5.4. Using Multiple Dissimilarity Data Tables for Documents Categorization

Participants: Thierry Despeyroux, Yves Lechevallier.

In collaboration with F.A.T De Carvalho and Filipe M de Melo, we have developped a clustering algorithm that is able to partition objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. These matrices could have been generated using different sets of variables and a fixed dissimilarity function, using a fixed set of variables and different dissimilarity functions or using different sets of variables and dissimilarity functions. This method, which is based on the dynamic hard clustering algorithm for relational data, is designed to provide a partition and a prototype for each cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fit between clusters and their representatives. These relevance weights change at each algorithm iteration and are different from one cluster to another.

To illustrate the usefulness of the proposed clustering algorithm, we use it to categorize a document database. The document database is a collection of reports produced by every Inria research team in 2007. Research teams are grouped into scientific *themes* that do not correspond to an organizational structure (such as departments or divisions), but act as a virtual structure for the purpose of presentation, communication and evaluation. Choice of themes and team allocation are mostly related to strategic objectives and scientific

closeness between existing teams, but also take into account some geographical constraints, such as the desire for a theme to be representative of most Inria centers. Our aim is to compare the categorization given automatically by the clustering algorithm that we have developped with the *a priori* expert categorization given by INRIA.

To do that, we used the XML parser described in section 4.4.2 to parse and extract the proper parts of the reports, the treetagger program lemmatizes the extracted text, before giving this intermediate result to our clustering algorithm.

The comparison between the automatic clustering and the *a priori* expert categorization shows minor divergences that can be explained by political choices of Inria [32], [40].

5.5. Fuzzy Clustering on Multiple Dissimilarity Matrices

Participants: Marc Csernel, Yves Lechevallier.

This work is done in collaboration with F.A.T de Carvalho (University of Recife, Brazil) [41].

The goal of the fuzzy clustering algorithms is to partition objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. The aim is to obtain a collaborative role of the different dissimilarity matrices in order to obtain a final consensus partition. These matrices could have been obtained using different sets of variables and dissimilarity functions. These algorithms are designed to furnish a partition and a prototype for each fuzzy cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fitting between the fuzzy clusters and their representatives.

These relevance weights change at each algorithm iteration and can either be the same for all fuzzy clusters or different from one fuzzy cluster to another. Experiments with real-valued datasets from UCI machine learning repository as well as symbolic data sets show the usefulness of the proposed fuzzy clustering algorithms.

5.6. Anti-Bouncing Model for Usage Data Streams

Participants: Chongsheng Zhang, Florent Masseglia, Yves Lechevallier.

This work takes place in the context of Chongsheng Zhang's Ph.D thesis.
The bounce rate (BR) of a website is the percentage of visitors (or users) who hit a given page and do not visit any other page on that website. It is defined as $BR = (\frac{T_o}{T_v})$ with T_o the total number of visits viewing only one page and T_v the total number of visits. According to Wikipedia: *it essentially represents the percentage of initial visitors to a site who "bounce" away to a different site, rather than continue on to other pages within the same site*. Bounce rate is very important for usage analysis and most commercial websites would like to lower it. We claim that, in some cases, the observed bounce rate is higher than the real one, because of the data stream model.

In [39], [49] we introduce the ABS (Anti-Bouncing Stream) model, a new model relying on a novel point of view. The goal of ABS is to maintain a reliable representation of the recent data in the stream and to avoid breaking down the navigations. We show that with ABS, we observe lower bounce rates.

5.7. Discovering Informative Feature Set over High-dimensions

Participants: Chongsheng Zhang, Florent Masseglia.

This work takes place in the context of Chongsheng Zhang's Ph.D thesis. Part of it is funded by MIDAS (Mining Data Streams), an ANR project (cf. 7.2.2).

In this work, we tackle the problem of informative feature set selection over unlabeled high-dimensional data. Differing from frequent pattern mining, which counts the frequencies of the patterns when the features appear together in the transactions, informative feature set selection has to take into account many other existing cases. For instance, when the features did not appear together, e.g., some of the features appeared in a transaction but other features in the feature set did not. Selecting the most informative feature set having size k in high-dimensional data is a difficult problem. The difficulties are on two aspects: first, there are many candidate sets with k features, and for each candidate we have to count the probability for every existing case; second, high-dimensional data make it even more difficult as we have massive candidates to check. To tackle the problem, we propose a heuristic theory to reduce the candidate features for informative feature set to a quite small subset. In addition, we build a forward selection algorithm to discover the most informative feature set using the carefully selected features. Moreover, we make a data structure to promptly compute the entropy of the features and introduce a pruning strategy at each forward extension so as to minimize the candidates to evaluate. This work hasn't been published yet, but our experiments on real-world data sets demonstrate the efficiency and effectiveness of the heuristic theory.

5.8. Discovering Evolution Patterns from Satellite Image Time Series

Participant: Florent Masseglia.

Satellite Image Time Series (SITS) provide us with precious information on land cover evolution. By studying these series of images we can understand the changes of specific areas but also discover global phenomena that spread over larger areas. However, discovering these evolution patterns implies to consider two main challenges, related to the characteristics of SITS and the domain's constraints. First, satellite images associate multiple measures with a single pixel (the radiometric levels of different wavelengths corresponding to infra-red, red, etc.). Second, these evolution patterns spread over very long periods and they may have different start time and end time depending on the region. Furthermore, the non evolving regions, which are majority and dominate over the evolving ones, challenge the discovery of these patterns. In [35], [47] we propose a SITS mining framework that allows for discovering these patterns despite these constraints and characteristics. Our proposal is inspired from sequential pattern mining and provides a relevant visualisation principle. This work has been accepted to the EGC 2011 conference (january 2011).

5.9. User Oriented Expert Finding

Participants: Elena Smirnova, Brigitte Trousse.

This work takes place in the context of Elena Smirnova's Ph.D thesis. supervised by B. Trousse (AxIS) and K. Avratchenkov (Maestro) and in collaboration with Krisztian Balog (ILPS group, University of Amsterdam).
The goal of expert search is to return a ranked list of people who are knowledgeable on a given topic. Several models have been proposed for expert finding, but so far, these have been focusing solely on returning the most knowledgeable persons as experts on a given topic [70] [101]. In this work we argue that in a real-world organizational setting, the notion of the best expert also depends on the actual user and his needs. We propose a user-oriented approach that balances between two factors that influence the user's choice: time to contact an expert and the knowledge value gained after. We use the distance between the user and an expert in a social network as an indication of contact time, and consider various social graphs, based on organizational hierarchy, geographical location, and collaboration, as well as the combination of these. We performed evaluation against a state-of-the-art baseline on the UvT Expert Collection using graded relevance judgements collected from real users, and demonstrated that the user-oriented approach significantly and substantially outperforms the baseline for all retrieval measures, namely, Mean Average Precision, Mean Reciprocal Rank, and precision at rank 5.

This work has been accepted to ECIR 2011 (only 20% of submissions were accepted)..

5.10. Using Web Graph Structure for Person Name Disambiguation

Participants: Elena Smirnova, Brigitte Trousse.

This work takes place in the context of Elena Smirnova's Ph.D thesis. supervised by B. Trousse (AxIS) and K. Avratchenkov (Maestro).
In the third edition of WePS campaign[6], we have undertaken the person name disambiguation problem referred to as a clustering task. Our aim was to make use of intrinsic link relationships among Web pages for name resolution in Web search results. To date, link structure has not been used for this purpose. However, Web graph can be a rich source of information about latent semantic similarity between pages. In our approach, we hypothesize that pages referring to one person should be linked through the Web graph structure, namely through topically related pages. Our clustering algorithm consists of two stages. In the first stage, we find topically related pages for each search result page using graph-based random walk method. After, in the next step, we cluster Web search result pages with common related pages. In the second stage, Web pages are further clustered using content-based clustering algorithm. The results of evaluation have showed that this algorithm can deliver competitive performance. The official performance ranking over WePS-3 participants showed that our algorithm took the second place (in F-0.5 measure) among 8 competitors within total 27 submitted runs. This work was presented at WePS [38] and also partially described in the context of a research report on Monte-Carlo methods [54].

5.11. Latent Variable Models for Expert Finding

Participants: Elena Smirnova, Brigitte Trousse.

This work takes place in the context of Elena Smirnova's Ph.D thesis. supervised by B. Trousse (AxIS) and K. Avratchenkov (Maestro).
The majority of existing expertise retrieval algorithms use document collection as a main source of evidence and analyze content with respect to persons associated with the documents [70]. In this setting, latent variable models can be directly applied to expert finding task because it implies existence of "hidden" experts who relate to "observed" documents. We propose an extension of latent Dirichlet allocation model [73] and apply it to expert finding problem. Our approach infers topics of expertise for each person and combines that with social relationships among persons. We assume that if a person has some expertise in the topic and so do persons linked to him, one have more confidence in that person's level of expertise. This also reflects an idea that person's expertise is related to other's through social relationships. Our model enables persons, topics and social relationships to be suitably coupled. Moreover, sparsity in topic representation of a document and person's expertise profile is easily achieved setting appropriate prior parameters.

5.12. Living Lab Landscape of Research Streams: towards User Co-creation

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach, Dominique Scapin.

Through the set up of the ICT Usage Lab and our involvement in several european initiatives, we strongly felt the need of a clear cartography of research streams in the Living Lab domain. To address this question, our priority was to launch a first state of the art about an ubiquitous notion, *user experience*.
From this review, progressively emerged a landscape that we organized through 4 main axes: granularity (individual/group), user'role (subject/actor), collaboration style (structured/unstructured), evaluation purpose (reliability/adoptability). This analysis is represented in the Figure 1.

From this tentative structure, it became easier to organize the different approaches of living labs, struggling between technology push and application pull. This tentative cartography of user centered innovation processes is a good start to plan further clarification. Such a landscape is used by finish colleagues for cartographying Living labs in Finland.

5.13. Future Internet Domain Landscape

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach, Mylène Leitzelman.

[6] WePS evaluation campaign

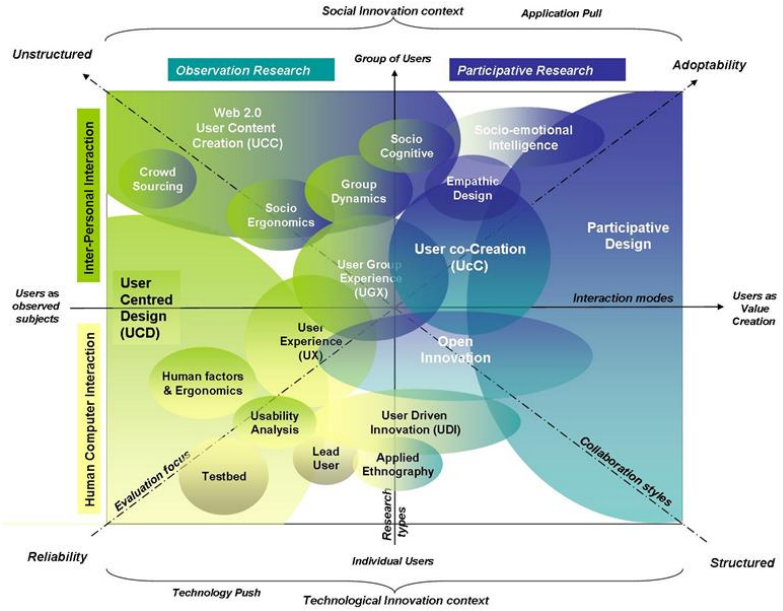

Figure 1. Living Lab Landscape of Research Streams

Several research domains are converging to support the development of innovation ecosystems. Smart cities are considered as an experimentation field for FI research and for user driven open innovation. The Fireball project (cf. 7.3.4) is to bring together three different constituencies: user driven open innovation, Future Internet, and Smart Cities. It aims at defining a roadmap, based on analysis of needs, opportunities and gaps, to benefit a wide scale implementation of the methodologies and concepts elaborated..
A first objective in the project was to get a clear view of the state of the art in each domain. During the review, progressively emerges a landscape that we organized along 4 main axes: wiring (wired/wireless), user's role (subject/convergence), Internet evolution approach (structured/unstructured), evaluation purpose (reliability/adoptability). This analysis is represented in the Figure 2.

As it is illustrated in Figure 2, a large variety of research have been engaged. If initial efforts in Future Internet research have been directed towards the goal of providing the technical infrastructure supporting the next network generation, a rising trend in this research field is to consider now a higher level layer, the layer of services.

5.14. Comparison of Usability Methods: Inspection vs. User Testing

Participant: Dominique Scapin.

In the domain of user interface usability, a very active research field aims at defining new evaluation methods and at improving existing ones. A study has been conducted to compare threes usability evaluation methods: User Testing (UT), Document-based Inspection (DI), and Expert Inspection (EI). In an experiment [18] based in the context of Virtual Environments (VEs) evaluation, twenty-nine individuals (10 end-users and 19 junior usability experts) participated during 1 hour each in the evaluation[2] of two VEs (a training VE and a 3D map). Quantitative results of the comparison show that:

- the effectiveness of UT and DI is significantly better than the effectiveness of EI.

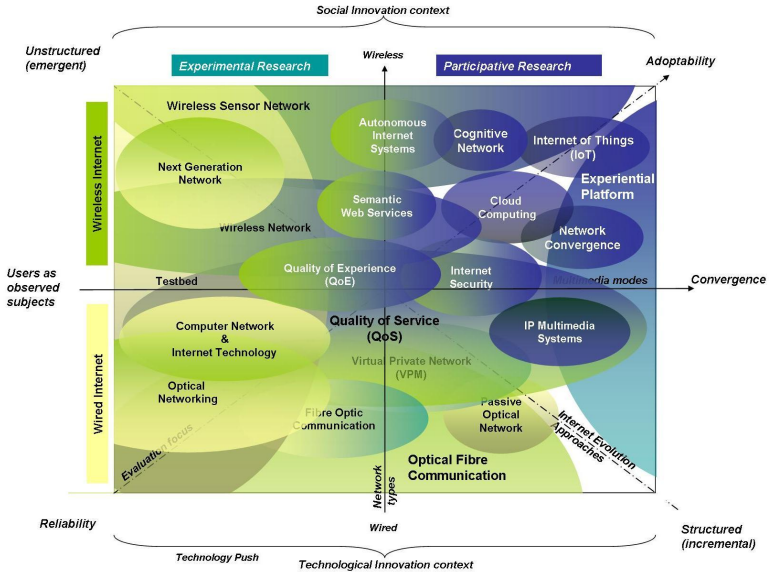

Figure 2. Future Internet domain Landscape

- DI- and UT-based diagnoses lead to more problem diversity than EI and thus have a better problem coverage.
- The identification impact of the whole set of usability problems is 60% for DI, 57% for UT, and only 36% for EI for both virtual environments.
- Reliability of UT and DI is significantly better than reliability of EI.

In addition, a qualitative analysis shows that:

- UT seems particularly efficient for the diagnosis of problems that require a particular state of interaction to be detectable.
- DI supports the identification of problems directly observable, often related to learnability and basic usability.

Also, the data obtained from these experiments have been analyzed further with a different perspective, through collaboration with M. Schmettow (University of Twente) and C. Bach (IRIT). The purpose here is to use a new statistical model to detect and measure heterogeneity bias in data during evaluation processes, what will improve efficiency of the process. This was published in a national conference [36]..

5.15. Formal Design and Validation of the Dialogue in Interactive Software

Participant: Dominique Scapin.

This work was carried out in collaboration with the University of Poitiers and ENSMA (S. Caffiau, P. Girard). It is a follow up on previous collaborative efforts that ended up with the creation of K-MAD and SUIDT. The main thrust of the work has been described in an international journal [20].

Task analysis is a critical step in the design process of interactive systems. The large set of task models available today may lead to the assumption that this step is well supported. However, very few task models are tool-supported. And in this latter category, few of them are based on a clear semantics. An assessment of the features that have been considered as essential in task modeling has been made by comparing different tool-supported methods, and evaluating the actual use of these features in K-MADe, a tool aimed at contributing to the incorporation of ergonomics into the design process of interactive systems through activity and task analysis. The analysis shows that :

- The originality of the K-MADe tool is to be based on a model whose expressive power lies on computable syntax while trying to be usable by every modelling knowledge designer.
- This facilitates task description and analysis, but also model query and the migration within software engineering models and software lifecycle steps.

Evaluation results demonstrate the usefulness of an increased expressive power for task models, and their acceptance by users. They also enlighten some weaknesses in the K-MAD method and suggest further improvements. Other papers have been published on the results obtained, focusing on learning, assessment, and process issues, such as, respectively:[23], [25], [26].

6. Contracts and Grants with Industry

6.1. Orange Labs CRE (2009-2010)

Participants: Yves Lechevallier, Florent Masseglia [resp.], Brigitte Trousse.

Orange and Inria have signed a CRE (Contrat de Rercherche Externalisée), where the goals are:

- To transfer a technology from AxIS to Orange.
- To add a new step devoted to specific needs of Orange.
- To perform experiments on Orange Data.

The AxIS technology involved in this project is the SMDS [92] method proposed in the Ph.D thesis of Alice Marascu. SMDS is an algorithm intending to extract sequential patterns from data streams. The data stream is processed by batches that accumulate the data for the same time period. Afterwards, each batch is processed in order to i) obtain clusters of data and ii) extract approximate sequential patterns that will summarize the clusters. Each such sequential pattern will be considered as a center of the cluster and will evolve from one batch to another.

The new step added by request of Orange is a second clustering process, applied to the approximate sequential patterns. A demonstrator was developed and applied on 3 months data provided by Orange coming from the mobile portal usage. These log files contain the request of users on the portal, made from their mobile devices. This project ended by July 2010, with the final deliverable containing the complete demonstrator [61].

6.2. Cassette Voyage (2009-2010)

Participants: Nicolas Béchet, Yves Lechevallier [resp.].

This work is done in collaboration with Marie-Aude Aufaure (Ecole Centrale de Paris, Business Intelligence Chair),
Addictrip.fr is a website created by the Cassette Voyage company. The goal of the Addictrip project is to produce a recommender system in the e-Tourism field. It focusses on a new travelers requirement: the city breaks. The assumption is to provide short breaks in major cities by offering stays in hotels at attractive prices. Addictrip collaborates with various partners as “booking” or “Venere” in order to propose a various choice of hotels.

Two main tasks in the contract AxIS-Cassette (OSEO-4997) started in september were needed to make these recommendations: computing of user profiles according to visited hotel web-pages (to realize this task, we used AWLH produced in the AxIS project) and computing hotel profiles based on the construction of an ontology of the hotel business field (We first manually build concepts of ontology by using the services of two representative partners. Then we proposed an approach to automatically expand ontology by adding services from other partners).

7. Other Grants and Activities

7.1. Regional Initiatives

Due to the bi-localization of the team, we are involved in two regions: PACA and Ile-de-France. We were this year very active in both regions,

In PACA: In 2008, the PACA Regional Council launched the Pacalabs initiative for supporting experimentations of innovative services (from SMEs) or exploratory ideas (from academic world) with real users. AxIS-Sophia is involved in the living lab ICT Usage (cf. 7.1.1) and in the CPER telius (cf. 7.1.2).
Five new projects started this year:

- A new PREDIT ADEME project TIC-TAC started in 2010 (Sophia Antipolis) (cf. 7.1.3),
- Two Pacalabs call 2 were accepted among our 3 proposals in 2009: ECOFFICES (cf. 7.1.4) and HOTEL-REF-PACA (cf. 7.1.5),
- One COLOR inria action on public data with ADEME, Fing and CETE - Médirannées,
- The INTERREG IV Alcotra proposal called MyMed in collaboration with the Inria Lognet team (leader: L. Liquori) has been accepted.

In Ile de France:

- A new Web 2.0 project "SCAR" started in 2010 with SME Wozaik (cf. 7.2.5),
- We submitted 2 FUI9 proposals (not accepted): NAVTI and Itransport GDS
- We have a lot of contacts in the context of Equipex related to “Saclay mobilité 2015”: due to our overload, we have postponed our implication in these projects.

7.1.1. ICT Usage Lab: Activities & projects, Terriotries, Campus STIC Usages

Participants: Brigitte Trousse, Bernard Senach.

Site: http://www.ictusagelab.fr
Activities and Projects
AxIS carried on with current founding partners (Orange Labs, INRIA, CSTB, UNSA) the setting up of the Living lab ICT Usage Lab (Sophia-Antipolis) who was labelled during the first wave by the European Network of Living Labs (ENoLL) in 2006 and was the first French living lab.

This year, at the regional level, the ICT Usage lab has driven its roots deeper in the territories through collaborations with SME, exchanges with local authorities and participation to events to improve its local visibility. The contacts established last year with SMEs and territories in order to answer to different calls for proposals were fruitfull.

- Two new Pacalabs call 2: Hotel Ref PACA (cf 7.1.5)) and ECOFFICES (cf. 7.1.4)
- a new Color action: CDISOD (Citizen Driven Innovation Services based on Open data) in order to initialize a collaboration with Fing and Ademe has been accepted (cf. 7.1.6).
- Bus230 project: We provided with Orange Labs (Sophia Antipolis) a technical assistance to Ademe for the collect of qualitative feed-back in a study aiming at the definition of a mobility label.
- As representative of the ICT Usage lab, B. Trousse was invited to participate to several meetings between Inria and Groupama (Innovation headquarter) to discuss the feasabiity of a joint Living lab focused on well being and supportive community living.
- During the Living lab Summer School (cf. 7.3.2), we met Pr. Harald von Kortzfleish from University of Koblenz-Landau and had another meeting in Sophia while he was invited teacher of "design thinking" at SKEMA. As he intended to set up a living lab in Germany, we planned to collaborate on European projects.

Communication: we put stress on communication this year through several actions:

- A beta version of the ICT Usage lab site is on line: http://www.ictusagelab.fr
- Short articles describing the ICT Usage lab are available on line (Inédit N°75, Lisa may 2010)
- Following the 1st Living Lab Summer Scool, post-proceedings are under progress and handbook about living labs is being written
- For the "Fete de la science" event (21-24 october 2010), the Living lab ICT Usage Lab was presented on two sites (Sophia Antipolis, Inria stand) and CSTB building and UNS (Nice, Mymed-Inria stand).

B. Trousse, as representative of the living lab ICT Usage Lab was invited to participate in a panel related to *Clusters and Living Labs* at Marseille, event organised by the Institute of the Méditerrannée.

Links with Territories in PACA

We had several meetings with representatives of the Urban Community of Nice Côte d'Azur for collaboration in different proposal:

- With CSTB, Gridpocket and the cluster Capenergies and NCA related to the proposal "Stories" Pilot Type B to the 4th call of the ICT PSP
- With Renault, Veolia Transport, University of Nice and SMEs for the project Lisem (Libre service Mobilité)
- With Symitam to envision a collaboration to the user interface specification of a future multimodal transportation system.

We took part in several meetings organized by the PACA Regional Council

- Workshop "Observatoire de l'Innovation en région", 13 october 2010, Marseille:B. Senach.
- Workshop "Ecosystème numérique", 18 october 2010, Marseille: B. Trousse

We took part in an innovation project in the City of Grasse "Innover Autrement", where QR-Code are available along routes in the city to provide touristic information.

We submitted in 2010 the INTEREG-Alcotra project "Excellence Technologique Touristique" with the Regional Council of Tourism of the PACA Region. the Tourism Office Of Aix, the Tourism Office of Digne-les-Bains and Provincia di Imperia (Italy, coordinator).

Campus STIC " Usages"

The ICT Usage lab leads researches related to the topic *Usages*, one of the four topics of the Campus Sophia-STIC http://sophia-stic.polytechnice.fr/page316.html. Cross Research related to the three others topics of the Campus STIC is addressed through several projects:

- Ubiquitous systems and networks: in the Elliot project (cf. 7.3.3) use of new services based on an IoT environment of sensors for environmental data will be studied.
- ICThealth: in the Elliot project (cf. 7.3.3) an IoT (Internet of Things) environment with sensors for environmental data will be installed. An expected application domain is health service for impaired people (asthmatic people).
- Sustainable development and environment: the Fireball project (cf. 7.3.4) is dedicated to Smart Cities with the goal of providing a roadmap toward successful innovation thanks to coordination and alignment of methodologies and approaches in the domains of Future Internet (FI) research and experimentation testbeds and user driven open innovation. See also the Ecoffices project (cf. 7.1.4).

In addition, the ICT Usage Lab has defined 3 main streams for its research activities on "Usages":

- Research for open innovation (ELLIOT, cf. 7.3.3 and CDISOD, cf. 7.1.6).
- Set up of methodology and software platform for usage analysis and support to innovative services (cf. MyMed platform (cf. 2) and FocusLab platform of the CPER Télius (cf. 7.1.2).
- Theory of usages: we began to address this subject few years ago through studies about confidence in carsharing and followed this year through a state of the art focused on user experience. This work has been used to set up a living lab landscape during LLSS 2010 (cf. 7.3.2).

7.1.2. PACA CPER TELIUS: FocusLab Platform

Participants: Brigitte Trousse, Bernard Senach, Yves Lechevallier.

In 2007, in a framework agreement between french government and PACA territory (CPER Telius), AxIS proposed the creation of an experimentation platform FOCUS (Finding Out Collective Usage) renamed FocusLab. This proposal was accepted and in 2008 AxIS received funding for buying equipments useful for future experiments in the Region. The FOCUS platform is composed of three parts:

- Equipments to collect usage data
- Methods (good practices, handbooks)
- Software to collect, structure and process usage data

The planned software part of the FOCUS platform aims to support researchers, engineers or SMEs in evaluating their prototypes&services, interested in usage analysis, such as web usage, co-conception, user centered pilots in real life (scale 1) or experiments in laboratories contexts. Support offered to computer science community and industrial concerned by usage assessment, prototype validation and benchmarking will help them to improve practice in analyzing usage data.

To avoid ambiguities with a new research Inria team, we renamed our platform Focusab. This year it was enriched year with several technical equipment: 1 eye-tracker TOBII and 10 personal computers

A 3 days training session was held to master advanced eye tracking studies (participants: Inria, Unsa and Orange labs). A technical analysis of current AxIS software is currently in progress. It will lead to precise requirement for the portal giving access to our software.

7.1.3. PREDIT ADEME: TicTac (jan 2010- jan 2012)

Participants: Guillaume Pilot, Bernard Senach, Brigitte Trousse [resp.].

Requirements for sustainable development push innovation. If one wants people to use public transportation solution rather than individual vehicules, these collective solutions have to become very attractive.

Tictac project aims to provide an advanced travellers' information system in which real time information about waiting time at bus stop will be provided: users define their "favourite" and can call a vocal server which give them immediately the requested information. In a first step, a single bus line with in-vehicule GPS is concerned by the study but in the next step the information system will be extended to other means of transport (train, car pooling...) over a larger territory. Our data mining technologies are applied used to tracks of cell phone usage to understand how people use the service.
In this project, we study also new methods to collect users feed-back and improve information quality through Web 2.0 tools.
Partners: VuLog (project coordinator), INRIA, MHC Conseil.
Site: http://www.projettictac.org/

7.1.4. PacaLabs: ECOFFICES Project (aug 2010- sep 2011)

Participants: Carole Goffart, Bernard Senach, Brigitte Trousse [resp.].

ECOFFICES is an eco-challenge within an enterprise. Some offices are equipped with sensors and feed-back concerning energy consumption is provided. The goal of the project is to provoke behavioural changes. In this project our team is in charge of the evaluation phase: usage data concerning actions on actuators will be registered and employees behaviour will be tracked. The experimentation will consist in three successive stages: In the first stage data are registered during the usual work of the challengers. The user interface providing consumption feed-back is provided The second stage is the challenge phase where 3 teams are competing to reach the best economy level. In the last stage, after the challenge, data are registered to study the change of practices, if any. We provided one delivrable (D1.3) related to the experimentation protocol [65]
Partners: CASA, CSTB, Osmose (Project coordinator) and Inria
Site: http://www.ecoffices.com/

7.1.5. PacaLabs: HOTEL-REF-PACA Project (dec 2010- dec 2011)

Participants: Florian Bonacina, Bernard Senach, Brigitte Trousse [resp.].

This project, conducted with Perferencement, a SME specialized in web site referencing, aims at improving hinterland tourism. Experiments of different new referencing rules will be conducted with Web site visitors in order to study their effect on behavioural changes and touristic choices.
The experimentation will consist in three main stages:

- In the first stage data current referencing rules are studied and their efficiency estimated through eye-tracking experiments
- In the second stage potential new rules are explored and tested by users
- In the third stage, the selected new rules are tested and their efficiency evaluated through data mining and qualitative studies.

Hotel-keepers and tourists will be involved in the experimentation for validation of the new referencing rules.

The project was offically launched on December 7th at the Palais de la Méditerrannée, Nice.

Partners: Perferencement (projct coordinator), Inria and General Council of Alpes Maritimes (Territorial authority of the Riviera).

7.1.6. Color Inria Action: Open Data - CDISOD (2010)

Participants: Mylène Leitzelman, Bernard Senach, Brigitte Trousse [resp.].

During the year 2009, at a talk about Living lab, we met FING (Fondation Internet Nouvelle Generation, C. Nepote and J.-M Bourgogne), a French thinktank involved in open innovation projects and open access to public data. We decided then to collaborate on this subject and submitted this year with the collaboration of ADEME (G. Plassat) and CETE Méditerranée (P. Gendre), a color action (INRIA funding) named CDISOD (Citizen Driven Innovative Services based on Open Data). The goal of the project is to study how a collaborative platform can be used to support open innovation processes. The expected application domain was public data about a local transportation system but the access to the data was not available.A state of the art of open innovation tools is under finalisation.

7.1.7. Collaborations with SMEs

Participants: Yves Lechevallier, Bernard Senach, Brigitte Trousse.

We had this year many contacts with SMEs in Paca interested in usage analysis and IS evaluation.

- In order to submit the proposal "Stories" Pilot Type B to the 4th call of the ICT PSP we met with the cluster Capenergies and the SME Gridpocket.
- We had several meetings with SSII GFI to visit their Userlab and envision collaboration on innovation projects.
- After a first meeting at a "speed dating session" organized at INRIA Sophia Antipolis Méditerranée we submitted an APRF proposal "Smartdoc santé" with the SME Coexel (other partners CHU NICE, Mobeo).
- In the context of the proposal "Navti" (FUI10), we had contact (among other partners) with SMEs Senda and Moviken. The submission was not accepted but we keep contact with these SMEs for further collaboration.
- We have close relationship with AtmoPaca for the project Elliot (cf. 7.3.3).
- Memo project: We had several meetings with Eurecom and Kuantic to set up the proposal MEMO (ANR verso) which concerned communicating objects in mobility to avoid collision and improve security (not accepted).
- B. Senach and B. Trousse have pursued their contact with Orange Labs (Sophia Antipolis) for future collaborations in the context of the living lab *ICT Usage Lab* (cf. 7.1.1).

7.2. National Initiatives

7.2.1. Introduction

The following national intiatives are briefly described in this section.

- ANR: MIDAS Project (2008-2010) (cf. 7.2.2)
- ANR: MyCitizSpace Project (2008-2010) (cf. 7.2.3)
- ANR: PIMI Project (2010-2013) (cf. 7.2.4)
- Web 2.0: Scar (2009-2011) (cf. 7.2.5)
- Web 2.0: 2.0 Process (2009-2011) (cf. 7.2.6)
- Living Lab ICT Usage Lab, member of the French Network of Living Labs (cf. 7.2.7)
- National Inria Projects (cf. 7.2.8) and others collaborations (cf. 7.2.9)

7.2.2. ANR: MIDAS Project (2008-2010) - extended to June 2011

Participants: Yves Lechevallier, Florent Masseglia [resp.], Brigitte Trousse, Maurice Yared, Chongsheng Zhang.

The MIDAS project “Mining Data Streams”, granted by ANR, has started on January 2008 and will be completed by June 2011 (extended deadline).
The MIDAS project aims at studying, developing and demonstrating new methods for **summarizing data streams**. It tackles the following scientific challenges related to the construction of summaries:

- Summaries are built from infinite streams but must have a fixed or low increasing size;
- The construction of summaries must be incremental (done *on the fly*);
- The amount of CPU used to process each element of the streams must be compatible with the arrival rate of the new elements;
- The summaries must cover the whole stream and enable to build summaries of any past part of the history of a stream.

Last year, we have coordinated and participated in the final version of a delivrable written by all partners in their work, on the topic of "Summarizing data streams by means of patterns extraction" [86].
This year, we are authors of a delivrable made of a demonstrator on "Summarizing a data stream by means of SCDS" [68]. This demonstrator implements the SCDS method from the Ph.D. thesis of Alice Marascu [91].
Partners are Ceregmia, EDF, France Telecom R&D, LIRMM, Telecom ParisTech and Inria.
Site: http://midas.enst.fr/wakka.php?wiki=PublicationS

7.2.3. ANR: MyCitizSpace Project(2008-2010)- extended to June 2011

Participants: Pascal Marie-Dessoude, Dominique Scapin [resp.].

The ANR project MyCitizSpace started in January 2008 (36 months, contact). MyCitizSpace « Méthode et outils de conception basés sur une approche d'Ingénierie Dirigée par les Modèles (IDM) pour l'exécution distribuée des téléprocédures plastiques à espace de données personnelles sécurisé » aims at the design of a method and tools based on a Model-Driven Architecture for the distributed execution of plastic teleprocedures incorporating a secure personal data space). In short, it means studying, designing and evaluating software models, methods and tools that will support the design of e-procedures (part of e-government applications), with a user-centered focus. MyCitizSpace is jointly conducted by three research laboratories (INRIA Rocquencourt, LIG Grenoble, IRIT Toulouse), two companies specialized in teleprocedures, Model-Driven Architecture (Genigraph), security of the electronic documents lifecycle (Almetis), and two partners providing access to field applications in e-government (Région Midi-Pyrénées et Direction Régionale du Travail, de l'Emploi et de la Formation Professionnelle d'Ile-de-France). The research focuses on current developments in e-administration, particularly concerning the dematerialization (paperless environment); it concerns both the end users to facilitate their access to the administrative procedures, and the administration to streamline their processes. Overall, the goal is to aim for the most seamless possible electronic procedures between the administrations and the citizen.
In collaboration with IRIT, a study of future technologies allowing the electronic management of information or data "so-called personal" within an Individual Information Space has been conducted [60].
The research work resulted in a book chapter on e-government. [52] and two conference papers on interfaces plasticity [37], [48].
Partners: Genigraph (project coordinator), INRIA, IRIT, LIG (HCI team), Almetis, DRTEFP d'Ile de France, Région Midi-Pyrénées.

7.2.4. ANR: PIMI Project (end 2010- end 2013)

Participants: Pascal Marie-Dessoude, Dominique Scapin [resp.].

The future Internet will bring a growing number of networked applications (services), devices and individual data (including private ones) to end-users. The important challenges are the organization of their access, and the guarantee of trust and privacy. The objectives of the PIMI project are the definition of a design environment and a deployment platform for Personal Information Management system (PIM). The future PIM must provide the end-user personal data access with services that are relevant to his needs. In order to take mobility into account, the PIM will be accessed both by mobile devices (smartphone) and Personal Computers. With the increasing number of services and associated data being accessible through Internet, the number and complexity of PIM will augment dramatically in the near future. This will require strong research investment in a number of topics, all contributing to the expected usability and accessibility of Individual Information Spaces for the end-user. Therefore, the main issues and the scientific locks are:

- Knowledge of the way users manage their information and services and how they can do it in the future.
- Knowledge of the perceived feeling of trust and security for the users, most importantly through the question of information sharing..
- Definition of a methodological approach to design and to tailor PIMs
- Design of secure means of data access and service usage.
- How to enable services to adapt to changes in user requirements and in the environment?

- How to make PIM user interface easy to use?
- Definition of a platform: technically, the goal is the integration of data and service composition, electronic trust semantics, security of data and service access, ergonomics of mobile PIM within a Model Driven Engineering approach..

Partners: Genigraph (project coordinator), LRI, IRIT, Institut Telecom, INRIA, Montimage, The Grand Duchy of Luxembourg.

7.2.5. Web 2.0: Scar Project (nov 2009- nov 2011)

Participants: Florian Bonacina, Anne-Laure Negri, Bernard Senach, Brigitte Trousse [resp.].

The goal of the Scar[7] projet is to provide to users of an advanced bookmarking system a recommender based on social navigation. With the technology available in the project, when visiting a web site, users can store "dynamic bookmarks", that is, they can select a part of the current web page and put it in their personal space. Later on, if the content of the selection is updated, the users will see the updated content on the selected aera. With this technology a personal web page can display a great variety of views presenting latest updated content.

For the recommender, the challenge comes from the dynamics of the bookmarks needing to consider the content of the selected area and to relate it to previous bookmarking. In this context, personalized recommendations will be calculated from past selection of the user, from bookmarks of users with a similar navigation profile. Some requirements for a user-centered design are made in [67].
Partners: Wozaïk (project coordinator), INRIA
Site: http://www.wozaik.com/welcome.php?invite=scar&try=3465&referer=scar

7.2.6. Web 2.0: 2.0 Process Project (nov.2009-nov 2011)

Participants: Pascal Marie-Dessoude, Dominique Scapin [resp.].

Process 2.0 proposes to consider the collaborative process as a composition of business services, "drawn" by end users through a collaborative design studio process that supports a multi-faceted modeling and integration of patterns. These processes leverage to deliver new business processes and creating real "social networks" that encourage industrial partnerships in providing these business services to professional communities. This approach incorporates a vision of "Software As A Service" and "Platform as a Service ", and allows a solution " from start to end " to users, freeing up investment and project management of deployments on their own infrastructure.
A review on tools for collaborative and end-user design was delivered [58] as well an ergonomics needs analysis [59].
Partners: Genigraph (project coordinator), INRIA, LIESP (via INSAVALOR), Petals link (ex EBM Websourcing), Région Midi-Pyrénées.
Site: http://genibeans.com/cgi-bin/twiki/view/ProcessTwoO

7.2.7. ICT Usage Lab, member of the French Network of Living labs

AxIS carried on with current founding partners (Orange Labs, INRIA, CSTB, UNS) the Living lab ICT Usage Lab (Sophia-Antipolis) activities at the national level through efforts towards the set up of the French Living lab network. After the 4th labelling wave launched by the ENoLL network, there are now 25 LL in France (14 new ones). The ICT Usage Lab organised in Paris the 3rd meeting of the French Living lab Network (17 participants, 12 living labs were represented). This meeting was an opportunity to define the roadmap for the development and structuring of the French living lab network.

[7] SCAR: Système Collaboratif et Adaptatif de Recommandations

7.2.8. National Inria projects

We had several contacts with on going or coming Inria projects:

- Inria DSI: a) Mobile jungle, contact with Julien Marboutin, b) DSI meeting with Bruno Sportisse for prospective about the Living Lab ICT Usage Lab and with Anne-Céline Lamballe for Transports domain.
- PAL Action (Action d'envergure), "Personal Assisted Living" (D. Daney) for collaboration envisioning human factors aspects and usage analysis. Short presentation of AxIS and the ICT Usage Lab was done by B. Trousse at the PAL workhop organised at Sophia Antipolis (december).

7.2.9. Other Collaborations

Academic Collaborations

- EGC association: We pursued our active participation in the EGC community. Y. Lechevallier participated at the meeting *Complexité due aux données multiples* of the EGC-FDC group (CNAM, Paris, 18 juin).
- Ecole Centrale de Paris: Yves Lechevallier collaborated with M-A. Aufaure related to our Tunisia STIC program [56] and the Addictrio contract (cf. 6.2)
- Telecom Paris Tech: we collaborated with Georges Hébrail and F. Rossi via the MIDAS project andor the BiLab laboratory
- AxIs is involved in the Business Intellience Laboratory (Bilab), a common laboratory between Telecom Paris Tech and EDF. Inria and France Telecom are the next members of this laboratory. We participated in Bilab Seminars (F. Masseglia and Y. Lechevallier).
- Paris Descartes: M. Csernel collaborated with Cazenade [30] and LePouliquen [44]
- CNAM: Y. Lechevallier collaborated with G. Saporta for the supervision of the Charrad's thesis [17] and with M-A. Aufaure in the context of the Addictrip project and the STIC France-Tunisia [56].
- University of Poitiers and ENSMA: D. Scapin collaborated with S. Caffiau and P. Girard [20], [23], [25], [26]
- IRIT (Toulouse): D. Scapin collaborated with C. Bach [18], [36] and M. Winckler [60], [52],
- LIG (Grenoble): D. Scapin collaborated with G. Calvary and A. Serna [52], [37], [48],
- University of Lille: F. Masseglia collaborated with F. PeztitJean and P. Gancarski [47].

Strategic Partnership between France Telecom - Inria; such a partnership (see RA2009) was in stand-by in 2010 at the national level. One concrete action for AxIS with Orange labs in this context was the CRE contract (cf.. 6.1)..

7.3. European Initiatives

7.3.1. Introduction

The following european initiatives are briefly described in this section:

- ICT Usage Lab: LLSS2010 Summer School and EnOLL association (cf. 7.3.2)
- FP7 ICT objective 1.3: ELLIOT STREP (cf. 7.3.3)
- FP7 ICT objective 1.6: FIREBALL Coordination Support Action (cf. 7.3.4)
- EU-Asian IDEAS Project (2010-2012) (cf. 7.3.5)
- COST TwinTide Action (cf. 7.3.6)
- Others initiatives (cf. 7.3.7) and collaborations (cf. 7.3.8)

7.3.2. ICT Usage Lab: LLSS2010 Summer School and EnOLL association

Participants: Bernard Senach, Brigitte Trousse, Bernard Senach, Brigitte Trousse.

the First European Living Labs Summer School

The First European Living Lab Summer School was held in Paris at the Cité des sciences et de l'industrie la Villette from August 25th to 27th, 2010 This summer school was proposed by the ami-community *Learning open innovation and living labs* created at the end of 2009. It was organised by two French living labs (ICT Usage Lab, Integrative Usage Lab), Esocenet and Unbla. The summer school was launched by Claudie Haigneré (President of universcience) and gathered more than 80 participants coming from 24 countries among which Canada, Mexico, Taïwan, Senegal and Singapore. The workshop have produced a landscape of practices and research knowledge and a roadmap for contributing to major societal issues innovation. Proceedings will soon be made available to the whole Living Lab Community. More additional information on the official web site: http://www-sop.inria.fr/llss2010

ENoLL, European Network of Living Labs

In january 25, 2010, an International Non for Profit Association under Belgian law (AISBL) was created from ENoLL and the ICT Usage Lab is "creating member" of the ENoLL association, CSTB, France Telecom - Orange Labs and University of Nice - Sophia Antipolis as co-funders of the ICT Usage Lab have accepted that INRIA was the legal entity representing the living lab.

B. Trousse (representative of the effective member ICT Usage Lab) was elected as member of the ENoLL administration board.

7.3.3. FP7 ICT objective 1.3: ELLIOT STREP

Participants: Mylène Leitzelman, Anne-Laure Negri, Bernard Senach, Brigitte Trousse [resp.].

The ELLIOT project (*Experiential Living Labs for the Internet of Things*) aims at developing an Internet Of Things (IOT) experiential platform where users/citizens are directly involved in co-creating, exploring and experimenting new ideas, concepts and technological artifacts related to IOT applications and services. Based on a three levels experiential model issued from previous European projects, the study will capitalize on existing practices of co-creation in IoT contexts . It will allow the exploration of the potential impact of IOT and of the Future Internet in the context of the Open User Centered Innovation paradigm followed in the Living Lab approach.
In this project AxIS is involved in all work packages with small contributions in the modeling phase (WP1) [62] the state of the art of user co-creation tools and techniques (WP3) [55] and dissemination (WP5). Main contributions, are in :

- WP2 where AxIS participates to the implementation of the Elliot experiential platform
- WP4 where AxIS is in charge of "Green Services" use cases which will take place in Antibes or in the Urban community of Nice Cote d'Azur (NCA) with collaboration of local territory authorities.

The Green Services use cases will involve three partners:

- VULOG, a private company which provides urban transportation services consisting of using on demand electrical vehicles,
- FING, a French thinktank involved in open innovation projects and open access to public data,
- AtmoPACA (a non-profit association in charge of pollution control)

In the Green Services use cases, existing data coming from sensors already installed in urban sites will be complemented by data provided by citizen. Indeed, participants to the experiment will have fixed andor mobile sensors (green watch, Arduino sensors, sensors on electrical vehicles).They will contribute to a better collect of data and will define specific information services according to their needs.
Partners : TXT Polymedia,Italy (project coordinator), University of Nottingham, University of Readings, BIBA, Hospital San Rafael, CENG, Fing, Vulog.
Site: http://www.elliot-project.eu/

7.3.4. FP7 ICT objective 1.6: FIREBALL Coordination Support Action

Participants: Marc Pallot, Bernard Senach, Brigitte Trousse [resp.].

FIREBALL(*Future Internet Research and Experimentation By Adopting Living Labs - towards Smart Cities*) is a coordination action which establishes a coordination mechanism through which a network of Smart Cities across Europe engages in long term collaboration for adopting User Driven Open Innovation to explore the opportunities of the Future Internet.
The key objectives of Fireball are :

1. To achieve European-wide coordination of methodologies and approaches in the domains of FIRE and Living lab
2. To leverage European-wide available assets for exploring Future Internet opportunities
3. To ensure coordinated development and sharing of best practices of Future Internet innovation in pilot cities and sectors.

In this project AxIS is mainly involved in two workpackages (WP1 and WP2) with the aim of setting up a FI research domain cartography. We first provided a review of main actors and projects at the international level from which we draw a first cartography [64].

The cartography is currently under a validation process through interviews with FI experts.The next step is to set up a cartography of the Smart Cities field and to connect it to a previous analysis of the Living lab field which was presented at the 1st Summer School of Living Labs in august 2010 (see 5.12). The goal is to articulate these three analysis altogether to provide a coherent view of the different domains (Smart Cities, Living Labs and Future Internet). The roadmap for Fireball will be defined through links connecting the different domains of these three research fields.

Partners (17): LTU-CDT (Sweden), AALTO (Finland), AENESCEN (Italy), MCC (United Kingdom), SAIM (Netherlands), ESADE (Spain), ALFAMICRO (Portugal), ISA (Portugal), E-NOVA (Portugal) HK (Finland), INRIA (France), DIMES (Finland), IBBT (Belgium), AUTH (Greece), OY (Finland), IMAGES & RESEAUX (France), BCN (Spain)

Site: http://www.fireball4smartcities.eu/

7.3.5. EU-Asian IDEAS Project (2010-2012)

Participants: Marc Csernel, Yves Lechevallier [resp.].

Ideas (Integrating and Developing European Asian Studies) is an European project which regroup a set of European Institution dealing with Asian Humanities plus INRIA. The overall objective of IDEAS is to make progress in coordinating and bringing together academic research, researchers and policy-makers. IDEAS will make use of the expertise and resources of a recently created network, the European Consortium for Asian Field Study (ECAF) , which comprises 44 research institutions from ten EU countries and nine Asian countries and Russia, which specialize in Asian studies, and a network of 22 field research centers run by ECAF members across Asia. The project started on 1st January 2010 and will run for 30 months.

Project Work Packages:

- WP1: Combining Strengths of the ECAF network
- WP2: Sharing access to a network of 22 field research centers located in Asia
- WP3: Sharing and exchanging access to knowledge resources INRIA is the leader of this workpackage which includes also IsIAO and HAS. The goal is to provide the foundation for the integration of the network library resources through the creation of an IT platform
- WP4: Creating interconnections between EU Asian Studies and policymakers needs

Partners: Ecole française d'Extrême-Orient, University of Turku Institute of Ethnology, Hungarian Academy of Sciences, British Academy, Asien-Afrika-Institut of the University of Hamburg, Istituto italiano per l'Africa e l'Oriente and INRIA.

For a more complete view of IDEAS see http://ec.europa.eu/research/social-sciences/projects/462_en.html or http://www.ideasconsortium.eu.

7.3.6. COST TwinTide Action

Participant: Dominique Scapin.

IC0904-TwinTide is a usability and user experience research community running under the auspices of COST (http://www.cost.esf.org/). TwinTide is the acronym for Towards the Integration of TransectorialIT Design and Evaluation . The main objective of the Action is to harmonise research and practice on design and evaluation methodologies for computing artefacts, across sectors and disciplines. Third-wave human computer interaction (HCI) is characterised by a diversifying user base and use contexts, new emphasis on user experience and new interaction styles. This implies a need for informed method choice sensitive to domains, user groups and system objectives. Effective method use requires complex judgments about applicability across applications and genres, with failure implying significant financial and human costs. The adoption of ICT across ages and abilities further increases the need for sound design and evaluation (D&E) methods, which bring about useful, usable, desirable computing artefacts that improve life quality. Effective cross-sectorial transfer of D&E methods is plausible and demonstrable. Relevant research work, however, is fragmented and scattered. The Action aims to provide harmonization and leadership currently lacking in this field by bringing together researchers and D&E professionals. Their broad experience of D&E methods deployed in different sectors and disciplines enables comparison of method applications, assessing transferability of both established and novel approaches. These collaborative activities in Working Groups and open Workshops will facilitate production of a generic D&E method selection and application framework and scientific publications reaching the wider research community. The Action will also provide young interdisciplinary researchers with systematic training and networking opportunities such as STSMs (short-term scientific missions) and Training Schools.

For the list of very numerous partners, see http://www.irit.fr/recherches/ICS/projects/twintide/memberslist.php

7.3.7. Others initiatives

Participants: Brigitte Trousse, Bernard Senach.

- **KIC ICT Labs (EIT)**: AxIS (B. Trousse and B. Senach) participated to a local contribution with others EPIs from Inria Sophia Antipolis for intergation in a proposal at the EIT call for KICs where Inria was involved at the national level. The proposal called *ICT labs* has been accepted at the european level. B. Trousse has a meeting with J. Chifflet and M. Lacage for innovation prospective about Living labs and the Living Lab ICT Usage Lab.
  Site: http://eit.ictlabs.eu/.
- **Interreg IV Alcotra Mymed** : The project MyMed started on January 25 aims at developing a transborder social network between France and Italia in which users will propose and get different kind of services (translation, booking for boat parking, etc.). AxIS was involved during the submission for a technical assistance for service definition and usage analysis.
  Partners: INRIA (Lognet team, Luigi Liquori), Politecnico di Torino, Universutà di Torino, Università del Piemonte Orientale.
  Site: http://www.mymed.fr
- **France- Germany : the AAI QUAERO contract**
  B. Trousse is involved in the project "Core Technology Cluster" of the AII program "QUAERO" in the multimedia domain (task 3.3 ontology evolution from usage analysis)
  Site: http://www.quaero.org/.

7.3.8. Other Collaborations

- Italy, University of Napoli II: Y. Lechevallier pursued his collaboration with Prof. R. Verde, A. Irpino and A. Balzanella [71]. This year, our visitor F. de A. T. de Carvalho presented there a seminar about "Recent clustering methods for interval data"(november).
- Nederlands: E. Smirnova collaborated on expert finding with Krisztian Balog from the Intelligent Systems Lab Amsterdam (ISLA) at the Informatics Institute of the University of Amsterdam.
- Finland: During the first living lab Summer school at Cité des Sciences, Paris, we validated with one of the three facilitators from VTT the relevance of creating a new ERCIM working Group related to Living Labs Research.
- Nederlands, University of Twente: D. Scapin collaborated with M. Schmettow [36],

7.4. International Initiatives

7.4.1. Introduction

The following international initiatives are briefly described in this section:

- Brazil: FACEPE-INRIA (cf. 7.4.2)
- Morocco: the WRUM Project (cf. 7.4.3)
- Tunisia: STIC program (cf. 7.4.4)
- USA: Explorer Program (cf. 7.4.5)
- Participation to Standards in Ergonomics (cf. 7.4.6)

7.4.2. Brazil: FACEPE-INRIA

Participants: Marc Csernel, Yves Lechevallier.

We continue our collaboration on clustering and web usage mining and we start a collaboration on social network data analysis with F.A.T. De Carvalho from Federal University of Pernambuco (Recife) and his team.

A scientific project *Clustering of Relational Data and Social Network Data Analysis* submitted by F. De Carvalho and Y. Lechevallier has been accepted by FACEPE and INRIA. The project starts on january 2010 and ends on december 2011. Researchers and students are concerned by this project from AxIS and CIn-UFPE side. It aims at developing methods of clustering of relational data and social network data analysis tools.

F. de A. T. de Carvalho was a visiting professor at AxIS project during three months (October to December). He collaborated actively with Y. Lechevallier [41], [40], [24] on partitional clustering algorithms based on multiple dissimilarity matrices in order to submit a paper to the COMPSTAT conference (Paris, August 22-27 2010) which was accepted [32] and an extended version to a international journal.

- This project aims to develop new clustering methods and algorithms for usual or symbolic feature data as well as for relational data. These new methods will apply simultaneously on several feature or relational data tables and they must be able to learn a relevance weight for each data table in each cluster. This kind of clustering method is useful in many situations. Here we will focus on applications in the social network data analysis.
- M. Csernel has visited the UFPE and has modified the N.S.F (Normal Symbolic Form). program to be able to obtain easily all different distance arrays we could need to proceed to the comparison tests concerning the new classification algorithm.

7.4.3. Morocco : the WRUM Project

Participants: Bernard Senach, Brigitte Trousse.

We kept contact with the WRUM project (Web Redesign based on Usage Mining). This project received fundings from the Research and Education Ministery and is managed by the CSPT commission. In this project close to the three topics of AxIS, the chosen application domain is e-Learning. Hicham Behja, Prof at ENSAM Meknès is in charge of the WRUM project with the support of Prof. A. Marzark (Univ. of Casablanca). AxIS is involved in the thesis committees of three PhD students (cf. 7.5.5).

7.4.4. Tunisia: STIC program

Participant: Yves Lechevallier.

During the STIC program, Y. Lechevallier in collaboration with M-A. Aufaure (Ecole centrale) supervised 4 masters and 2 thesis (Riadi Lab, ENSI Tunis) in this project. These masters and thesis subjects are about web mining (usage, content and structure, using different methods) and ontology construction from heterogeneous sources.
Y. Lechevallier is involved in the new STIC program *Exploration des réseaux sociaux pour les systèmes de recommandation* between France-Tunisia [56]. In this STIC project, we welcomed from ENSI (Tunisia) at Rocquencourt: A. Louati (April-August) who participated actively to the themes of our STIC under the co-supervision of H. Baazaoui Zghal, M.-A. Aufaure (Supelec), H. Ben Ghezala and Y. Lechevallier. The report [57] describes the need to design and implement a tool for analysing social networks based on the aggregation graphs. After a state of the art of social networks and their analysis, a tool graph based aggregation *k-SNAP approach* was tested.

7.4.5. USA: Explorer Program

Participant: Chongsheng Zhang.

Chongsheng Zhang, Ph D student, is visiting the WIS team of Prof. Carlo Zaniolo at UCLA from October to November 2010 in the context of the explorer program. This visit is dedicated to the study of bi-streaming data. In such a data stream, each transaction is made by a user with a unique identifier. And it is frequent for one user to make a few transactions within a short time interval. As such, if we observe the transactions by their users, the feature set for each user is unfinished and only partly available at one time point, but new features may be added afterwards. The characteristic that distinguishes it from other data streams lies in the co-evolving of the data points (the feature sets for the users) with the datastreams. That is to say, in a bi-streaming data streams, it is not uncommon that the data points are updating themselves. The dimension of the data can be very large, but one data point only has several features, whereas the features for different data points are not identical. There are many such data streams in real applications. Chongsheng has studied techniques for bistreaming data mining and more particularly itemset mining from such streams.
This work is done in collaboration with Mirjana Mazuran (Polimi, Italy), Hamid Mousavi(UCLA), Carlo Zaniolo (host professor at UCLA).

7.4.6. Participation to Standards in Ergonomics

Participant: Dominique Scapin.

Standardization in ergonomics is increasingly important due to the application of the European directives about the introduction of measures to encourage improvements in the safety and health of workers (e.g., 2006/42CE on security of machinery); as well as taking into consideration national and international legislation, including accessibility. Standardization in ergonomics covers many issues. The contributions from INRIA (D. L. Scapin) concern mainly software ergonomics, in the context of AFNOR X35A, X35E, as well as ISO and CEN mirror groups:

- National: AFNOR X35A (Ergonomie des Logiciels Interactifs) (expert); AFNOR X35E (Ergonomie des Logiciels Interactifs) (chair).
- European: CEN/TC 122/WG 5 (Software ergonomics and human-computer dialogs) (expert)
- International: ISO/TC 159/SC4/WG5 (Software ergonomics and human-computer dialogues) (expert); ISO/TC 159/SC4/WG6 (Human-centred design processes for interactive systems) (expert and co-editor of ISO 9241-230); JWG ISO/TC 159/SC 4 and ISO/IEC JTC1 SC 7 (System and software product Quality Requirements and Evaluation (SQuaRE) - Common industry Format) (expert); ISO/TC 159/SC1/WG1 (Ergonomic principles) (expert).

7.5. Animation of the Scientific Community

7.5.1. Introduction

AxIS members are involved in the animation of the scientific community:

- Reviewing Activites (cf. 7.5.2)
- Organization of Conferences Workshops (cf. 7.5.3
- University teaching (cf. 7.5.4)
- Ph.D. Thesis (cf. 7.5.5)
- Activities of General Interest (cf. 7.5.7).

7.5.2. Reviewing Activities

AxIS is involved in the management and the edition of 4 journals:

- the RSTI scientific committee related to the "ISI, L'OBJET, RIA, TSI" journals (Hermes publisher): B. Trousse is a member of this committee.
- The MODULAD electronic journal: Y. Lechevallier is one of the four editors.
- BIT (Behaviour & Information Technology). D. L. Scapin (Associate Editor:).
- Romanian Journal of Human-Computer Interaction:: D. L. Scapin (member of the International Advisory Board)

AxIS members belong to 8 editorial boards of international journals and two special issues of a national journal (RNTI, RIA):

- the Co-Design Journal (Editor: Janet McDonnel University of the Arts London - Publisher: Swets & Zeitlinger): B. Trousse
- IJDST - International Journal of Design Sciences & Technology (Editors R. besheshti & K. Zreik - Publisher: Europia productions): B. Trousse
- JSDA (the Journal of Symbolic Data Analysis) (Editor:E. Diday, electronic journal: Y. Lechevallier
- UAIS (International Journal of Universal Access in the Information Society). D. L. Scapin.
- IJHCS (International Journal of Human-Computer Studies): D. L. Scapin.
- IWC (Interacting with Computers): D. L. Scapin.
- IJPOP (International Journal of People-Oriented Programming): D. L. Sc apin
- JIPS (Journal d'Interaction Personne-Système). D. L. Scapin.
- RNTI special issue on mining complex data: F. Masseglia
- RIA Special Issue *Intelligence Artificielle et Web Intelligence* (Publisher, Hermes-Lavoisier) Vol. 23 N°1, January-Février, editors-in-chief: Y. Demazeau and L.Vercouter: B. Trousse

AxIS members were reviewers for 3 journals:

- DKE International Journal on Data Knowledge and Engineering (Elsevier): F. Masseglia
- INS International Journal on Information Sciences (Elsevier): F. Masseglia
- Third special issue of RNTI on complex data mining (numéro spécial RNTI sur la fouille de données complexes): F. Masseglia, B. Trousse.

Several AxIS members were involved in 9 international conferences/worhshops as members of Program Committee:

- ECAI'10, 19th European Conference on Artificial Intelligence: B. Trousse
- CSCWD'10, 14th International Conference on Computer Supported Cooperative Work in Design (Santiago, Chile, April 22-24, 2009): B. Trousse
- ICCBR'10, 8th International Conference on Case-Based Reasoning, held in Seattle, Washington, USA from 20 July to 23 July 2009: B. Trousse
- COMPSTAT 2010 [53], the 19th Conference of IASC-ERS COmputational Statistics, CNAM, Paris, August 23-27: Y. Lechevallier , B. Trousse
- HCSE 2010, 3rd Conference on Human-Centred Software Engineering; October 14-15, 2010, Reykjavik, Iceland: D. L. Scapin.
- AHFE 2010, 3rd International Conference on Applied Human Factors and Ergonomics, July 19-22, 2010, Miami, FL, USA: D. L. Scapin.
- IEEE ICDM: the 10^{th} IEEE International Conference on Data Mining, Sydney, Australia, December 14-17: F. Masseglia.
- ICML 20010, the 27th International Conference on Machine Learning, Haifa, Israel, June 21-24: F. Masseglia.
- ACM SAC 2010, the 25th Symposium On Applied Computing, Lausanne, Switzerland, March 22-26: F. Masseglia.

AxIS members were also involved in national conferences/worhshops as members of Program Committee.

- IC'10: 21emes journées francophones d'Ingénierie des Connaissances modèles, représentations, raisonnements, gestion et usages des connaissances, Nimes, http://ic2009.inria.fr/: B. Trousse
- EGC'10: 10èmes Journées Francophones *Extraction et Gestion des Connaissances* (Hammamet, Tunisie, janvier): Y. Lechevallier, F. Masseglia, B. Trousse
- Atelier EGC'10: (Hammamet, Tunisie, 26 January 2010)- La recherche d'information personnalisée sur le web: M-A. Aufaure, Y. Lechevallier, B. Trousse
- EGC-M 10: 1ère édition de la co,férence Maghrébine sur l'*Extraction et la Gestion des Connaissances* (Alger, Algérie, december 13-14): Y. Lechevallier, F. Masseglia, B. Trousse
- BDA 2010 demonstrations: F. Masseglia
- IHM 2010: 22ème. Conference Francophone sur l Interaction Homme-Machine; Septembre 20-23, Luxembourg: D. L. Scapin.
- ErgoIA 13-15 October 2010: D. L. Scapin.
- Ateliers EGC'10
  - Fouille de Données Complexes: complexité liée aux données multiples: Y. Lechevallier, F. Masseglia, B. Trousse
  - La recherche d'information personnalisée sur le web: Y. Lechevallier (co-organisation)

Other Reviewing Activities: B. Trousse is expert for the cluster ISLE (Rhônes Alpes region) related to the project "Web intelligence. F. Masseglia was expert for the BLANC call for project and for the RTRA DIGITEO.

7.5.3. Organization of Conferences / Workshops

AxIS was involved in two major events:

1. **COMPSTAT 2010 (more thant 600 participants)**: Y. Lechevallier was vice-chair of the organizing committee of the 19th COMPSTAT symposium, COMPSTAT 2010 (August 22nd-27th 2010), which is the conference on the Statistical Computing and sponsored by the European Regional Section of the IASC (International Association for Statistical Computing). IASC is a section of ISI (International Statistical Institute). COMPSTAT2010 [53] is locally organized by the CNAM and the INRIA (AxIS). T. Despeyroux was member of this organizing committee.
2. **LLSS2010, First Living Labs Summer School** (around 80 participants from 23 countries)[2]): B. Trousse with K. Pawart and R. Santoro was at the initiative of the idea of organising the first edition of a Living Labs Summer School In France. B. Trousse was co-chair of the scientific committee and chair of the organisation committee of the summer school. The first Living lab Summer School took place at the Cité des Sciences (Paris, august) with the great support of universcience and LUTIN User Lab: http://www-sop.inria.fr/llss2010/

7.5.4. University Teaching

AxIS is an "associated team" for the STIC Doctoral school at the University of Nice-Sophia Antipolis (UNS) and AxIS team members are teaching in various universities (UNS, Univ Paris IX-Dauphine, CNAM, ENSAE):

- Master 2 Recherche Systèmes Intelligents (resp: S. Pinson) of the University Paris IX-Dauphine: Tutorial (12h) on *"Du data mining au knowledge mining"*: Y. Lechevallier.
- Master 2 Pro Mathématiques appliqués et sciences économiques (resp: P. Cazes) of the University Paris IX-Dauphine: Tutorial (15h) on *"Méthodes de classification"*: Y. Lechevallier.
- Master 2 Pro Ingénierie de la Statistique (resp: G. Saporta) of CNAM (12h) on *Méthodes neuronales*: Y. Lechevallier.
- ENSAE (*Ecole Nationale de la Statistique et de l'Administration Economique*): Tutorial (18h) on *Data Mining*: Y. Lechevallier.
- ENSG (*Ecole Nationale des Sciences Géographiques*): Tutorial (12h) on *"Analyse des données"*: Y. Lechevallier.
- "Master CSSR" (resp. Bruno Martin) at Polytech'Sophia: Lecture (3h) on *Data Mining: Sequences, Streams and Security*: F. Masseglia.

B. Trousse made AxIS presentation at Inria during student visits from ENS Lyon, (around 15 students, November 10) and from ENS Cachan (around 23 students), December 10.

7.5.5. Ph.D. Thesis

One Ph.D. defense this year:

- **M. Charrad** (start: end of 2005) [17], "Une approche générique pour l'analyse croisant contenu et usages des sites Web par des méthodes de bipartitionnement", CNAM and University La Manouba (Tunisia). Y. Lechevallier is co-superviser with G. Saporta (CNAM) and in *co-tutelle* with Prof M. Benhamed (ENSI, Tunisia).

Two PhD in progress at Inria :

1. **E. Smirnova** (start: September 2008), "Mining social networks", University of Nice-Sophia Antipolis & INRIA, (co-directors: M. Rueher, B. Trousse). Let us note that E. Smirnova made an internship at Google (Paris) may-august.
2. **C. Zhang** (start: october 2008), "Mining data streams: clustering and pattern extraction", University of Nice-Sophia Antipolis & INRIA, (director: F. Masseglia).

In the context of the WRUM project, B. Trousse and B. Senach participate in the thesis committees of three Ph-D students (Morocco):

- **M. Naamany** from the University of Casablanca supervised by A. Marzark.
- **N. Sael** from the University of Casablanca supervised by A. Marzark.
- **E Zemmouri** from the University of Fes and ENSAM (Meknès, Morocco) supervised by H. Behja and A. Marzark.

AxIS researchers were members of Ph.D. committees in 2010:

- **T. Qureshi**, Ph.D, Contributions to Decision Tree based Learning, June, University Lumière Lyon II: Y. Lechevallier
- **G. Cabanes**, Ph.D, Classification non supervisée à deux niveaux guidée par le voisinage et la densité, December, University Paris 13: Y. Lechevallier (reviewer)
- **P. Salle**, Ph.D, "Les motifs séquentiels pour les données issues des puces ADN" University of Montpellier 2, July: F. Masseglia (reviewer).
- **G. Bonnin**, Ph.D, "Vers des systèmes de recommandation robustes pour la navigation Web: inspiration de la modélisation statistique du langage" University of Nancy 2, November: F. Masseglia (reviewer).

7.5.6. Internships

We welcomed two students this year:

- **A. Louati** [57], (supervised by Y. Lechevallier), AxIS Rocquencourt, Master of ENSI and Manouba university (Tunis) (February-June) on community detection in social networks.
- **M. Yared** [68]. (supervisors F. Masseglia and B. Trousse), AxIS Sophia Antipolis, Master 2, University of Nice-Sophia Antipolis, (from July to September) on mining data streams.

E. Smirnova attended the Machine Learning Summer school (Canberra, Australia, September 27 - October 6) http://mlss10.rsise.anu.edu.au/

7.5.7. Activities of General Interest

- Y. Lechevallier made an invited talk [24] at AAFD'2010 (Apprentissage Artificiel et fouille de données, Université Paris 13, june 29-30th).
- B. Trousse was invited this year to be a member of the "comité de pilotage" of the EGC association.
- AGOS: T. Despeyroux is involved (30 %) as President of AGOS (Inria Works Council).
- F. Masseglia is member of the editorial board of interstices (http://interstices.info). F. Masseglia is author of two documents for interstices:(on Google and on principles and challenges of data mining and usage analysis).
- French national "Science Celebration" ("Fête de la science"):
  - B. Senach and B. dTtrousse prepared a first version of the web site and a flyer which was distributed at CSTB, on Sophia Village (Inria stand) and at UNS Valrose Nice (Mymed stand).
  - F. Masseglia gave two sessions of one hour to students of Audiberti High School of Antibes on research work in Compuetr Science and in Data Mining.
- B. Trousse as Inria representative belonged to the expert committee of the Pacalabs and to the strategic committee of the Pacalabs orientation of the Regional Council.
- B. Trousse and B. Senach are members of the coordination committee of the ICT usage Lab (Inira, CSTB, orange Labs and Unsa).

- B. Trousse as representative of an effectiv emember, ICT Usage Lab was elected as a member of the administration committee of the EnOLL association..
- IEEE: B. Trousse is member of the technical committee on Computer Supported Work in Design in the context of the IEEE SMCS (Systems, Man & Cybernetics Society)

8. Bibliography

Major publications by the team in recent years

[1] E. GUICHARD (editor). *Mesures de l'internet*, Les Canadiens en Europe, 2004, ouvrage collectif suite au Colloque Mesures de l'internet, Nice, France, 12-14 Mai, 2003.

[2] R. BOULET, B. JOUVE, F. ROSSI, N. VILLA. *Batch kernel SOM and related Laplacian methods for social network analysis*, in "Neurocomputing", March 2008, vol. 71, n^{o} 7–9, p. 1257–1273.

[3] M. CHAVENT, F. DE CARVALHO, Y. LECHEVALLIER, R. VERDE. *New clustering methods for interval data*, in "Computational Statistics", 2006, vol. 21, n^{o} 23, p. 211-230.

[4] S. CHELCEA, P. BERTRAND, B. TROUSSE. *Un Nouvel Algorithme de Classification Ascendante 2-3 Hiérarchique*, in "Actes de 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle (RFIA 2004)", Centre de Congrès Pierre BAUDIS, Toulouse, France, 28-30 Janvier 2004, vol. 3, p. 1471-1480, http://www.laas.fr/rfia2004/actes/ARTICLES/388.pdf.

[5] B. CONAN-GUEZ, F. ROSSI, A. EL GOLLI. *Fast Algorithm and Implementation of Dissimilarity Self-Organizing Maps*, in "Neural Networks", August 2006, vol. 19, n^{o} 6-7, p. 855–863, http://dx.doi.org/10.1016/j.neunet.2006.05.002.

[6] A. DA SILVA, Y. LECHEVALLIER, F. DE CARVALHO, B. TROUSSE. *Mining Web Usage Data for Discovering Navigation Clusters*, in "11th IEEE Symposium on Computers and Communications (ISCC'06)", Pula-Cagliari, Italy, IEEE Computer Society, 26-29 June 2006, p. 910-915, http://doi.ieeecomputersociety.org/10.1109/ISCC.2006.102.

[7] T. DESPEYROUX. *Practical Semantic Analysis of Web Sites and Documents*, in "The 13th World Wide Web Conference, WWW2004", New York City, USA, 17-22 May 2004, http://www-sop.inria.fr/axis/papers/04www/despeyroux-www2004.pdf.

[8] T. DESPEYROUX. *Evolution of Ontologies and Types*, in "WWW/Internet 2008 Conference", Freiburg, Germany, October 2008.

[9] G. HÉBRAIL, Y. LECHEVALLIER. *Data mining et analyse des données*, in "Analyse des données", Hermes, June 2003, p. 340-360.

[10] A. MARASCU, F. MASSEGLIA. *Mining Sequential Patterns from Data Streams: a Centroid Approach*, in "Journal of Intelligent Information Systems (JIIS).", November 2006, vol. 27, n^{o} 3, p. 291-307.

[11] A. MARASCU, F. MASSEGLIA. *Parameterless outlier detection in data streams*, in "ACM Symposium on Applied Computing (SAC), Honolulu, Hawaii, USA, March 9-12", 2009, p. 1491-1495.

[12] F. MASSEGLIA, P. PONCELET, M. TEISSEIRE, A. MARASCU. *Web Usage Mining: Extracting Unexpected Periods from Web Logs*, in "Data Mining and Knowledge Discovery (DMKD) Journal.", 2007, DOI 10.1007/s10618-007-0080-z.

[13] F. MASSEGLIA, D. TANASA, B. TROUSSE. *Web Usage Mining: Sequential Pattern Extraction with a Very Low Support*, in "Advanced Web Technologies and Applications: 6th Asia-Pacific Web Conference, APWeb 2004, Hangzhou, China. Proceedings", LNCS, Springer-Verlag, 14-17 April 2004, vol. 3007, p. 513–522.

[14] F. ROSSI, B. CONAN-GUEZ. *Functional Multi-Layer Perceptron: a Nonlinear Tool for Functional Data Analysis*, in "Neural Networks", January 2005, vol. 18, n^{o} 1, p. 45-60, http://hal.inria.fr/inria-00000599.

[15] D. TANASA, B. TROUSSE. *Advanced Data Preprocessing for Intersites Web Usage Mining*, in "IEEE Intelligent Systems", March-April 2004, vol. 19, n^{o} 2, p. 59–65.

[16] F. DE CARVALHO, Y. LECHEVALLIER. *Partitional clustering algorithms for symbolic interval data based on single adaptive distances*, in "Pattern Recognition", 2009, vol. 42, n^{o} 7, p. 1223-1236.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[17] M. CHARRAD. *Une approche générique pour l'analyse croisant contenu et usages des sites Web par des méthodes de bipartitionnement*, CNAM, 2010.

Articles in International Peer-Reviewed Journal

[18] C. BACH, D. SCAPIN. *Comparing Inspections and User Testing for the Evaluation of Virtual Environments*, in "International Journal of Human-Computer Interaction", 2010, vol. 26, n^{o} 8, p. 786-824, http://dx.doi.org/10.1080/10447318.2010.487195.

[19] S. BASHAR, F. MASSEGLIA. *Discovering Frequent Behaviors: Time is an Essential Element of the Context*, in "Knowledge and Information Systems (KAIS)", 2010, DOI 10.1007/s10115-010-0361-5.

[20] S. CAFFIAU, D. SCAPIN, P. GIRARD, M. BARON, F. JAMBON. *Increasing the expressive power of task analysis: systematic comparison and empirical assessment of tool-supported task models*, in "Interact. Comput.", 2010, 10.1016/j.intcom.2010.06.003.

[21] G. HÉBRAIL, B. HUGUENEY, Y. LECHEVALLIER, F. ROSSI. *Exploratory Analysis of Functional Data via Clustering and Optimal Segmentation*, in "Neurocomputing / EEG Neurocomputing", Mar 2010, vol. 73, n^{o} 7-9, p. 1125-1141, 10.1016/j.neucom.2009.11.022, http://hal.inria.fr/hal-00515908/en.

[22] A.-M. MARASCU, F. MASSEGLIA. *Atypicity Detection in Data Streams: a Self-adjusting Approach*, in "Intelligent Data Analysis Journal", 2010, http://hal.inria.fr/inria-00461263/en.

Articles in National Peer-Reviewed Journal

[23] S. CAFFIAU, D. SCAPIN, L. SANOU. *Using K-MADe for Learning Task Modeling: Interests and Difficulties*, in "JIPS (Journal d'Interaction Personne-Système)", 2010, vol. 1, n^{o} 1.

Invited Conferences

[24] Y. LECHEVALLIER, A. DA SILVA, F. ROSSI, F. DE CARVALHO, R. VERDE, A. BALZANELLA. *La fouille des flux d'informations*, in "AAFD, 4 èmes journées thématiques Apprentissage Artificiel et Fouille de Données", Villetaneuse, Paris, june 29-30 2010.

International Peer-Reviewed Conference/Proceedings

[25] S. CAFFIAU, P. GIRARD, D. SCAPIN. *Task modelling techniques and their assessment*, in "W3C Workshop on Future Standards for Model-Based User Interfaces", May 2010, Rome, Italy.

[26] S. CAFFIAU, P. GIRARD, D. SCAPIN, L. GUITTET. *Prise en compte de l'utilisateur dans le processus de conception d'une application d'édition de modèles de tâches*, in "ERGO'IA", October 2010, Biarritz.

[27] M. CHARRAD, Y. LECHEVALLIER, M. BEN AHMED, G. SAPORTA. *On the Number of Clusters in Block Clustering Algorithms*, in "FLAIRS Conference", 2010.

[28] M. CHARRAD, Y. LECHEVALLIER, M. BEN AHMED, G. SAPORTA. *WCUM pour l'analyse d'un site web*, in "Extraction et gestion des connaissances (EGC'2010), Actes, 26 au 29 janvier 2010", Hammamet, Tunisie, 2010, p. 669-672.

[29] A. CIAMPI, A. DYACHENKO, Y. LECHEVALLIER. *Two-way Classification of a Table with non-negative entries: Validation of an Approach based on Correspondence Analysis and Information Criteria*, in "Proceedings of the COMPSTAT 2010 conference", Physica-Verlag, August 2010.

[30] M. CSERNEL, T. CAZENAVE. *Comparing Sanskrit Texts for Critical Editions*, in "Coling 2010", Beijing, China, Coling 2010 Organizing Committee, August 2010, p. 206–213.

[31] G. HÉBRAIL, B. HUGUENEY, Y. LECHEVALLIER, F. ROSSI. *Exploratory Analysis of Functional Data via Clustering and Optimal Segmentation*, in "ESANN 2010", 2010, vol. abs/1004.0456.

[32] Y. LECHEVALLIER, F. DE CARVALHO, T. DESPEYROUX, F. DE MELO. *Clustering of Multiple Dissimilarity Data Tables for Documents Categorization*, in "Proceedings of the COMPSTAT 2010 conference", Physica-Verlag, August 2010, Paris.

[33] A.-M. MARASCU, F. MASSEGLIA, Y. LECHEVALLIER. *A fast approximation strategy for summarizing a set of streaming time series*, in "Proceedings of the 2010 ACM Symposium on Applied Computing, SAC'2010", Sierre, Switzerland, March 22-26 2010, p. 1617-1621.

[34] M. PALLOT, B. TROUSSE, B. SENACH, D. SCAPIN. *Living Lab Research Landscape: From User Centred Design and User Experience towards User Co-creation*, in "1st. European Living Labs Summer School, Collaborative Innovation through Living Labs", August 2010, Cité des Sciences, Paris, France.

[35] F. PETITJEAN, P. GANÇARSKI, F. MASSEGLIA, G. FORESTIER. *Analysing Satellite Image Time Series by Means of Pattern Mining*, in "Intelligent Data Engineering and Automated Learning, 11th International Conference (IDEAL'2010)", Paisley, UK, September 1-3 2010, p. 45-52.

[36] M. SCHMETTOW, D. BACH. *Effizientere Usability Evaluationen mit gemischten Prozessen*, in "Mensch & Computer Konferenz", Duisburg, Germany, Usability Professionals Association, September 12-15 2010.

[37] A. SERNA, G. CALVARY, D. L. SCAPIN. *How assessing plasticity design choices can improve UI quality: a case study*, in "the 2nd ACM SIGCHI Symposium on Engineering Interactive Computing System , EICS 2010", Berlin, Germany, N. SUKAVIRIYA, J. VANDERDONCKT, M. HARRISON (editors), June 19-23 2010, p. 29-34.

[38] E. SMIRNOVA, K. AVRACHENKOV, B. TROUSSE. *Using Web Graph Structure for Person Name Disambiguation*, in "WEPS campaign: Searching Information about Entities in the Web", 2010.

[39] C. ZHANG, F. MASSEGLIA, Y. LECHEVALLIER. *ABS: the Anti-Bouncing Model for Usage Data Streams*, in "the IEEE International Conference on Data Mining (ICDM)", Sydney, Australia, December 14-17 2010.

[40] F. DE CARVALHO, T. DESPEYROUX, F. DE MELO, Y. LECHEVALLIER. *Utilisation de matrices de dissimilarité multiples pour la classification de documents*, in "Proceedings of Première Conférence Maghrébine sur l'Extraction et la Gestion des Connaissances", December 2010, Alger.

[41] F. DE CARVALHO, Y. LECHEVALLIER, F. M. DE MELO. *A Relational Fuzzy C-Means Clustering Algorithm Based on Multiple Dissimilarity Matrices*, in "Tenth International Conference on Intelligent System Design and Applications, ISDA 2010", Cairo, Egypt, December 2010.

National Peer-Reviewed Conference/Proceedings

[42] A. BALZANELLA, R. VERDE, Y. LECHEVALLIER. *Clustering highly evolving multiple data streams*, in "SFC", Saint-Denis, 2010.

[43] COLLECTIF MIDAS. *Résumé généraliste de flux de données*, in "EGC", Jan 2010, p. 255-260, Tunisie, http://hal.inria.fr/hal-00502017/en.

[44] M. CSERNEL, M. LE POULIQUEN. *Comparer les textes Sankrits à des fins d'édition critiques*, in "SFC 2010", St Denis, La Rèunion France, SFC 2010, June 2010, 91.

[45] M. CSERNEL, M. LE POULIQUEN. *Stemma codicum and relation d'intermédiarité, utilisation de la méthode de Don Quentin*, in "JADT 2010", Rome, Italy, JADT 2010, June 2010, p. 309-320, http://www.cavi.univ-paris3.fr/lexicometrica/jadt/jadt2010/allegati/JADT-2010-0309-0320_015-LePoulequin.pdf.

[46] A.-M. MARASCU, F. MASSEGLIA, Y. LECHEVALLIER. *REGLO: une nouvelle stratégie pour résumer un flux de séries temporelles*, in "Extraction et Gestion des Connaissances", Jan 2010, p. 218-228, Tunisie Hammamet, http://hal.inria.fr/inria-00461834/en.

[47] F. PETITJEAN, P. GANÇARSKI, F. MASSEGLIA. *Extraction de motifs d'évolution dans les Séries Temporelles d'Images Satellites*, in "Spatial Analysis and GEOmatics (SAGEO 2010)", Toulouse, 17-19 novembre 2010.

[48] A. SERNA, G. CALVARY, D. L. SCAPIN. *Penser Plasticité peut améliorer la Qualité des Interfaces Homme-Machine : une étude de cas*, in "ErgoIA", October 2010, Biarritz.

[49] C. ZHANG, F. MASSEGLIA, Y. LECHEVALLIER. *MUAR : un modèle anti-rebond pour les flux de données d'usage*, in "Bases de Données Avancées, BDA 2010", Toulouse, France, 19 - 22 octobre 2010, p. 79-99.

[50] C. ZHANG, F. MASSEGLIA. *Discovering Highly Informative Feature Sets from Data Streams*, in "Database and Expert Systems Applications, 21st International Conference (DEXA'2010)", Bilbao, Spain, August 30 - September 3 2010, p. 91-104.

[51] C. ZHANG, F. MASSEGLIA. *Extraction d'itemsets distinctifs dans les flux de données*, in "Extraction et gestion des connaissances (EGC'2010), Actes des 10èmes journées Extraction et Gestion des Connaissances", Hammamet, 26 au 29 janvier 2010, p. 187-198, France, http://hal.inria.fr/hal-00504877/en.

Scientific Books (or Scientific Book chapters)

[52] G. CALVARY, A. SERNA, J. COUTAZ, D. SCAPIN, F. PONTICO, M. WINCKLER. *12*, in "Envisioning Advanced User Interfaces for e-Government Applications: a Case Study", Springer, 2010.

Books or Proceedings Editing

[53] Y. LECHEVALLIER, G. SAPORTA (editors). *COMPSTAT 2010—Proceedings in Computational Statistics*, Physica-Verlag, Heidelberg, Germany, 2010, 619 pages [*DOI :* 10.1007/978-3-7908-2604-3], http://extras.springer.com/2010/978-3-7908-2603-6/ESM_COMPSTAT_ebook.pdf.

Research Reports

[54] K. AVRACHENKOV, N. LITVAK, D. NEMIROVSKY, E. SMIRNOVA, M. SOKOL. *Monte Carlo Methods for Top-k Personalized PageRank Lists and Name Disambiguation*, INRIA, 09 2010, n^{o} RR-7367, http://hal.inria.fr/inria-00510991/en/.

Other Publications

[55] J. ESCHENBÄCHER, A. HESMER, M. LEITZELMAN, B. TROUSSE, ET AL.. *State-of-the-art report for IOT oriented user co-creation*, november 2010, n^{o} D3.1, FP7 STREP Elliot Délivrable D3.1.

[56] M. HADDAD, H. BAAZAOUI, M.-A. AUFAURE, C. CLARAMUNT, Y. LECHEVALLIER, H. BEN GHEZELA. *Proposition d'une architecture pour la personnalisation de l'information spatiale sur le web*, in "Atelier EGC", Jan 2010, Tunisie.

[57] A. LOUATI. *Analyse des réseaux sociaux: développement d'un outil basé sur l'agrégation des graphes*, 2010.

[58] P. MARIE-DESSOUDE, D. L. SCAPIN. *Etat de l'art « Atelier Fonctionnel Collaboratif »*, november 2010, ANR Process 2.0 Délivrable D1.5.

[59] P. MARIE-DESSOUDE, D. L. SCAPIN. *Expression des besoins ergonomiques*, august 2010, n^{o} D1.2, ANR Process 2.0 Délivrable.

[60] P. MARIE-DESSOUDE, D. L. SCAPIN, M. WINCKLER. *Espace Individuel d'Informations « dites personnelles » Enquête pilote par questionnaire*, march 2010, ANR MyCitizSpace Délivrable.

[61] F. MASSEGLIA. *Rapport final sur le démonstrateur permettant de résumer les navigations sur le portail mobile*, 2010, CRE Orange labs Contract deliverable T1.3, 104 paaes.

[62] M. PALLOT, A.-L. NEGRI, B. TROUSSE, ET AL.. *D1.1 KSB Experience Model Overall Framework*, december 2010, n^{o} D1.1, FP7 STREP Elliot Délivrable D1.1.

[63] M. PALLOT, B. TROUSSE, B. SENACH. *Future Internet Research Landscape: From User Centred Design and User Experience towards User Cocreation*, september 2010, internal AxIS document.

[64] H. SCHAFFER, SÄLLSTRÖM, N. KOMNINOS, P. TSARCHOPOULOS, B. SENACH, B. TROUSSE, M. PALLOT, H. HIELKEMA. *Landscape and Roadmap of Future Internet and Smart Cities*, september 2010, n^{o} D2.1, FP7 CSA Fireball Délivrable D2.1.

[65] B. SENACH, C. GOFFART, B. TROUSSE. *Protocole d'expérimentation*, November 2010, n^{o} D1.3, Pacalabs Ecoffices Déliverable 1.3.

[66] B. SENACH, C. GOFFRAD, B. TROUSSE. *Ecoffices: protocole d'expérimentation*, 2010, n^{o} Lot 1.3, Pacalabs Ecoffices Déliverable 1.3.

[67] B. SENACH. *Exigences à satisfaire pour une conception centrée utilisateur*, juin 2010, n^{o} SP0.1, Web 2.0 Déliverable.

[68] M. YARED. *Classification de données d'usage par flux*, 2010, n^{o} Tâche 3.4, Internship report and ANR Midas T3.4 Déliverable.

References in notes

[69] H.-H. BOCK, E. DIDAY (editors). *Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data*, Springer Verlag, 2000.

[70] K. BALOG, T. BOGERS, L. AZZOPARDI, M. DE RIJKE, A. VAN DEN BOSCH. *Broad expertise retrieval in sparse data environments*, in "SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval", ACM, 2007, p. 551–558.

[71] A. BALZANELLA, Y. LECHEVALLIER, R. VERDE. *A new approach for clustering multiple streams of data*, in "Classification and Data Analysis 2009", S. INGRASSIA, R. ROCCI (editors), 2009, p. 417-420.

[72] P. BLACKBURN, J. BOS, K. STRIEGNITZ. *Learn Prolog Now!*, Texts in Computing, College Publications, 2006, vol. 7.

[73] D. BLEI, A. NG, M. JORDAN. *Latent Dirichlet Allocation*, in "Journal of Machine Learning Research", 2002, vol. 3, 2003.

[74] M. CHAVENT, Y. LECHEVALLIER, O. BRIANT. *DIVCLUS-T: A monothetic divisive hierarchical clustering method*, in "Computational Statistics and Data Analysis", 2007, vol. 52, p. 687-701.

[75] S. CHELCEA. *Agglomerative 2-3 Hierarchical Classification: Theoretical and Applicative Study*, University of Nice Sophia Antipolis, March 2007.

[76] W. F. CLOCKSIN, C. S. MELLISH. *Programming in Prolog*, Springer Verlag, 2003, 5th edition.

[77] B. CONAN-GUEZ, F. ROSSI. *Speeding Up the Dissimilarity Self-Organizing Maps by Branch and Bound*, in "Proceedings of 9th International Work-Conference on Artificial Neural Networks (IWANN 2007)", San

Sebastian (Spain), F. SANDOVAL, A. PRIETO, J. CABESTANY, M. GRANA (editors), Lecture Notes in Computer Science, June 2007, n^o 4507, p. 203–210.

[78] B. CONAN-GUEZ, F. ROSSI, A. EL GOLLI. *Fast Algorithm and Implementation of Dissimilarity Self-Organizing Maps*, in "Neural Networks", August 2006, vol. 19, n^o 6–7, p. 855–863, http://hal.inria.fr/inria-00174196.

[79] M. CSERNEL, F. DE A. T. DE CARVALHO. *The normal symbolic form*, in "Symbolic Data Analysis and the SODAS Software", E. DIDAY, M. NOIRHOMME-FRAITURE (editors), Wiley, 2008, p. 93-108.

[80] A. DA SILVA. *Analysis of evolving data: application to the Web usage data (in French)*, University of Paris IX Dauphine and INRIA-France, 2009.

[81] A. DA SILVA, Y. LECHEVALLIER. *Analyse exploratoire des indices pour la détermination du bon nombre de clusters : application aux données évolutives*, 2009, Groupe de Travail EGC sur la Fouille de Données Complexes (GT EGC-FDC): Complexité liée aux données multiples..

[82] A. DA SILVA, Y. LECHEVALLIER. *Axis Tool for Web Usage Evolving Data Analysis (ATWUEDA)*, INRIA-France, 2009, AxIS internal document.

[83] F. A. T. DE CARVALHO, R.M.C.R. DE SOUZA, M. CHAVENT, Y. LECHEVALLIER. *Adaptive Hausdorff distances and dynamic clustering of symbolic interval data*, in "Pattern Recognition Letters", February 2006, vol. 27, n^o 3, p. 167–179.

[84] T. DESPEYROUX. *Developing efficient parsers in Prolog: the CLF manual (v1.0)*, INRIA, 2006, n^o RT-0328, http://hal.inria.fr/inria-00120518/en/.

[85] M. E. FAYAD, D. C. SCHMIDT. *Object-Oriented Application Frameworks*, in "Communication of the ACM", 1997, vol. 40, n^o 10, p. 32-38.

[86] GET-ENST, LIRMM, CEREGMIA, INRIA, LGI2P, EDF R&D, FT R&D. *Summarizing data streams by means of patterns extraction*, 2009, n^o Tâche 1.3.

[87] M. HADJOUNI, M. R. HADDAD, H. BAAZAOUI ZGHAL, M.-A. AUFAURE, H. BEN GHEZALA. *Personalized Information Retrieval Approach*, in "WISM/CAISE'09 : Sixth International Workshop on Web Information Systems Modeling, held in conjunction with CAISE'09", 2009.

[88] M. JACZYNSKI. *Modèle et plate-forme à objets pour l'indexation des cas par situation comportementale: application à l'assistance à la navigation sur le Web*, Université de Nice Sophia-Antipolis, Sophia-Antipolis, December 1998.

[89] M. JACZYNSKI, B. TROUSSE. *Patrons de conception dans la modélisation d'une plate-forme po ur le raisonnement à partir de cas*, in "Revue l'Objet", 1999, vol. 5, n^o 2, Numéro Spécial sur les patterns orientés objets, D. Rieu et J-P. Giraudon (guest editors).

[90] R. E. JOHNSON, B. FOOTE. *Designing Reusable Classes*, in "Journal of Object-oriented programming", 1988, vol. 1, n^o 2, p. 22–35.

[91] A. MARASCU. *Extraction de motifs séquentiels dans les flux de données*, University of Nice Sophia Antipolis, 2009.

[92] A. MARASCU, F. MASSEGLIA. *Mining sequential patterns from data streams: a centroid approach*, in "J. Intell. Inf. Syst.", 2006, vol. 27, n^{o} 3, p. 291-307.

[93] M. NOIRHOMME-FRAITURE. *User manual for SODAS 2 Software*, FUNDP, Belgique, april 2004, version 1.0.

[94] R. A. O'KEEFE. *The craft of Prolog*, MIT Press, Cambridge, MA, USA, 1990.

[95] F. ROSSI, B. CONAN-GUEZ. *Functional Multi-Layer Perceptron: a Nonlinear Tool for Functional Data Analysis*, in "Neural Networks", January 2005, vol. 18, n^{o} 1, p. 45–60, http://hal.inria.fr/inria-00000599.

[96] F. ROSSI, B. CONAN-GUEZ. *Un modèle neuronal pour la régression et la discrimination sur données fonctionnelles*, in "Revue de Statistique Appliquée", 2005, vol. LIII, n^{o} 4, p. 5–30, http://hal.inria.fr/inria-00001190.

[97] F. ROSSI, B. CONAN-GUEZ. *Theoretical Properties of Projection Based Multilayer Perceptrons with Functional Inputs*, in "Neural Processing Letters", February 2006, vol. 23, n^{o} 1, p. 55–70, http://hal.inria.fr/inria-00001191.

[98] F. ROSSI, N. DELANNAY, B. CONAN-GUEZ, M. VERLEYSEN. *Representation of Functional Data in Neural Networks*, in "Neurocomputing", March 2005, vol. 64, p. 183–210, http://hal.inria.fr/inria-00000666.

[99] F. ROSSI, A. HASENFUSS, B. HAMMER. *Accelerating Relational Clustering Algorithms With Sparse Prototype Representation*, in "Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 07)", Bielefeld (Germany), September 2007, ISBN: 978-3-00-022473-7, http://dx.doi.org/10.2390/biecoll-wsom2007-144.

[100] F. ROSSI. *Model collisions in the dissimilarity SOM*, in "Proceedings of XVth European Symposium on Artificial Neural Networks (ESANN 2007)", Bruges (Belgium), April 2007, p. 25–30, http://apiacoa.org/publications/2007/dsom-collision-esann.pdf.

[101] P. SERDYUKOV, D. HIEMSTRA. *Modeling Documents as Mixtures of Persons for Expert Finding*, in "ECIR", 2008, p. 309-320.

[102] D. TANASA. *Web Usage Mining: Contributions to Intersites Logs Preprocessing and Sequential Pattern Extraction with Low Support*, University of Nice Sophia Antipolis, june 2005.

[103] B. TROUSSE, M.-A. AUFAURE, B. LE GRAND, Y. LECHEVALLIER, F. MASSEGLIA. *Web Usage Mining for Ontology Management*, in "Data Mining with Ontologies: Implementations, Findings and Frameworks.", N. HÈCTOR OSCAR, S. E. G. GONZALEZ CISARO, X. DANIEL HUGO (editors), Information Science Reference, 2007, chap. 3, p. 37-64.

[104] N. VILLA, F. ROSSI. *A comparison between dissimilarity SOM and kernel SOM for clustering the vertices of a graph*, in "Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 07)", Bielefeld (Germany), September 2007, ISBN: 978-3-00-022473-7, http://dx.doi.org/10.2390/biecoll-wsom2007-139.

[105] F. DE A. T. DE CARVALHO, Y. LECHEVALLIER, R. VERDE. *Clustering methods in symbolic data analysis*, in "Symbolic Data Analysis and the SODAS Software", E. DIDAY, M. NOIRHOMME-FRAITURE (editors), Wiley, 2008, p. 181-204.