European Project Space on Computational Intelligence, Knowledge Discovery and Systems Engineering for Health and Sports

Rome, Italy
October, 2014

Antonello Rizzi and Antonio Pedotti (Eds.)

Sponsored and Organized by INSTICC
Published by SCITEPRESS

Edited by Antonello Rizzi and Antonio Pedotti

Printed in Portugal
ISBN: 978-989-758-154-0
Depósito Legal: 398623/15

Foreword

This book contains the revised and extended versions of papers describing a number of European projects that were presented at the two European Project Space (EPS) events organized in Rome, October 2014, associated with the set of conferences IJCCI (International Joint Conference on Computational Intelligence), IC3K (International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management), icSPORTS (International Congress on Sport Sciences Research and Technology Support), NEUROTECHNIX (International Congress on Neurotechnology, Electronics and Informatics) and CARDIOTECHNIX (International Congress on Cardiovascular Technologies). All these events were sponsored by the Institute for Systems and Technologies of Information, Control and Communication (INSTICC) in collaboration with several international associations and other scientific partners.

The objective of the EPS is to provide a platform and meeting point for researchers and practitioners to discuss and exchange experiences on European Research and Development projects which are funded and developed in the European research area or in collaboration with European partners, thus providing insights into cutting edge research work originating from Academia or Industry, either in Europe or elsewhere.

We aim to establish the EPS as a regular event to create opportunities for knowledge and technology sharing, and establish the basis for future collaboration networks involving current project partners and interested conference delegates.

These two events included a panel discussion with representatives and experts from the European Community. On the October 23rd the panel entitled "The quest for excellence towards 2020" counted with the presence of Antonello Rizzi, Sapienza University of Rome, Italy as chair of the panel, Maria Uccellatore, Director Unit III – European Research

Programmes MIUR – DGIR, Italy; Maurizio Paschero, Polo per la Mobilità Sostenibile, Sapienza University of Rome, Italy; Paloma de las Cuevas, University of Granada, Spain and Gianluca Fabbri, Polo per la Mobilità Sostenibile, Sapienza University of Rome, Italy.

On the October 25th the panellists for the panel entitled "European Project Policy: Past Experiences and Future Opportunities" were Antonio Pedotti, Politecnico di Milano, Italy (chair of the panel), Sergio Cerutti, Polytechnic University of Milan, Italy, Emanuela Teresa Locati, Niguarda Hospital, Italy and Francesco Beltrame, National Research Council of Italy (CNR), Italy.

The EPS technical program included, in addition to an opening panel, the presentation of eight projects which, after the event, have been invited to publish a short report in this EPS book.

We would like to thank the project representatives that decided to take their time and effort to respond to our invitation, whose reports correspond to the seven chapters of this book.

Antonello Rizzi

Antonio Pedotti

Editors

Organization

Panel Chairs

Antonello Rizzi, "Sapienza" University of Rome, Italy
Antonio Pedotti, Politecnico di Milano, Italy

Panel Participants

Francesco Beltrame, National Research Council of Italy (CNR), Italy
Sergio Cerutti, Polytechnic University of Milan, Italy
Paloma de las Cuevas, University of Granada, Spain
Gianluca Fabbri, Polo per la Mobilità Sostenibile - University of Rome "La Sapienza", Italy
Emanuela Teresa Locati, Niguarda Hospital, Italy
Maurizio Paschero, University of Rome "Sapienza", Information Engineering, Electronics and Telecommunications Department, "Polo per la Mobilità Sostenibile - (POMOS)" Laboratory, Italy
Maria Uccellatore, Director Unit III – European Research Programmes -MIUR – DGIR, Italy

Presented Projects

Acronym: TLM-WEB
Presenter: Simone Porru, DIEE, University of Cagliari, Italy

Name: Integration of Metadata, Units, and Uncertainty with Numeric Values for Automation & Cluster Analysis
Presenter: Joseph E. Johnson, University of South Carolina Columbia, United States

Acronym: CRYSTAL
Presenter: Antonio M. Rinaldi, Università di Napoli Federico II, Italy

Acronym: MUSES
Presenter: Paloma de las Cuevas, University of Granada, Spain

Acronym: ACDS
Presenter: Alberto Buschettu, ExperTeam Srl, Italy

Name: Knowledge Management in Healthcare
Presenter: Erja Mustonen-Ollila, Lappeenranta University of Technology, Finland

Acronym: PCE
Presenter: Georgia Sanna, ExperTeam Srl, Italy

Acronym: REPOPA
Presenter: Arja R. Aro, University of Southern Denmark, Denmark

Acronym: PEGASO
Presenter: Giuseppe Andreoni, Industrial Design Faculty and Department of the University of Milan, Italy

Acronym: TRAMA
Presenter: Manuela Galli, Politecnico di Milano, Italy

Acronym: Cupid
Presenter: Alberto Ferrari, University of Bologna Viale Risorgimento 2, Italy

Name: Intelligent wearable devices for home monitoring of cardiac and psychiatric patients, as well as for the detection of parameters correlated to active ageing and well-being
Presenter: Sergio Cerutti, Politecnico di Milano, Italy

Acronym: CAPSYS
Presenter: Lübomira Spassova, Public Research Centre Henri Tudor, Luxembourg

Acronym: MOTION
Presenter: Lübomira Spassova, Public Research Centre Henri Tudor, Luxembourg

Acronym: C2S
Presenter: Gilles Kermarrec, University of Brest, France

Table of Contents

Papers

Papers

An Innovative Web Application for Advanced Library Services

Maurizio Calderamo[1], Simona Ibba[2], Filippo Eros Pani[2], Francesco Piras[1] and Simone Porru[2]

[1] SoSeBi Srl, Via dell' Artigianato, 9, 09122 Cagliari, Italy
{calderamo, piras}@sosebi.it
[2] Department of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi, 09123 Cagliari, Italy
{simona.ibba, filippo.pani, simone.porru}@diee.unica.it

Abstract. The future scenario in which libraries will work will see an increasing centrality of the Internet and of services provided through it, in order to meet the needs of physical and virtual users, which remotely access the library. SoSeBi srl, an innovative company located in Sardinia (Italy), and the Department of Electrical and Electronic Engineering (Dipartimento di Ingegneria Elettrica ed Elettronica, DIEE) of the University of Cagliari (Sardinia, Italy) intend to create a product at the forefront of the evolution of an ever-expanding sector, Information Retrieval, where general-purpose products like search engines, digital archives and social networks are competing. These products are, and will continue to be, a significant presence in our digital lives. Therefore, the objective of this research project is to create a software that responds to the users' tastes, that encompasses the Web 2.0 paradigms, and that is in close contact with technology innovation.

1 Introduction

In the last decade, libraries have gone from a heritage mainly made of paper, to a hybrid one, combining printed resources to multimedia materials of different kinds: music, audiovisuals, databases, electronic books, audiobooks, websites.

The evolution towards digital libraries, where reference services and loan of digital content can take place, is one of the means needed to meet the current and future needs of the citizens, more and more used to peruse digital contents of various types [1]; [2]. An especially important branch of the digital content area is eBooks, which have become increasingly widespread in Italy since the end of 2010. Despite the still small percentages compared to the paper book market, it is going to become a new reading support in the future. eBook is meant to represent the textual work, while the device used to read it is a different, separate topic. Currently, technology offers several types of eBook readers (called e-readers), from new-generation mobile phones (smartphones) to tablets (for example, the iPad), to readers specifically thought and designed with this purpose, among which Amazon Kindle is the most prominent example. An increasing number of Italian libraries is developing the need to provide reference and loan services for eBooks.

The presented project, financed by the Autonomous Region of Sardinia with European funds (Single Programming Document 2007-2013 - P.O. FESR 2007-2013 – Line of Activity 6.2.2.d – Interventions to support competitiveness and innovation, under the Regional Committee Resolution no. 33/41 of 08/08/2013), has the aim of implementing a Web-based application geared towards bibliographic cataloging and reference services in libraries, with innovative functions of semantic search, management of digital contents and the creation of a social network. The proponents also intend to cover the study of future interaction methods among libraries, eBooks, and eReaders.

The creation of the Web application will stem from the strategic partnership between SoSeBi Srl[1] and the Department of Electrical and Electronic Engineering (DIEE)[2] of the University of Cagliari. The purpose of this choice to use the results of fundamental research as well as of industrial research to elaborate an innovative prototype, unique in the domestic market for its features.

The software prototype is designed to have a high-flexibility modular architecture. In this way, the different functions can be independent, and not hinder the essential core of the management program in their development. The new software product aims to be a reference point in the Italian market of management software for libraries, in terms of innovation and completeness of integrated functions. A new-generation application will be created, with an innovative vision on the very concept of "automation program for libraries". The application will become a precursor to a general development trend for this type of software, anticipating and seizing the opportunities in the market in the coming years.

This paper is structured as follows: the second section describes the context in which the proposed project lies, while the section following it describes activities and objectives in detail. In the fourth section, the schedule of the project is outlined, especially concerning the activity of digital content management. The last section hosts our final observations about the project.

2 Context of Research Proposal

Hundreds of thousands of libraries exist throughout the world, of varied sizes and types. The libraries of Italy are estimated to be between 25,000 and 30,000, and the main ones (around 16,000) have a profile in the Italian Libraries Database (Anagrafe delle Biblioteche Italiane, AIB)[3] of the Central Institute for the Union Catalogue of Italian Libraries (Istituto Centrale per il Catalogo Unico, ICCU)[4]. The collection and verification of the data is continuous, also thanks to contributions by Regional governments, Universities, organizations and cultural institutions that work with ICCU.

Currently, those 16,000 public libraries (of which 52% belong to local organizations) provide basic services, also called "public reading", that have gradually grown and become a well-received reference point for thousands of citizens, together with other established public services. Thanks to innovation and automation, many of

[1] SoSeBi. srl, http://www.sosebi.it/.
[2] DIEE, University of Cagliari, http://dipartimenti.unica.it/ingegneriaelettricaedelettronica/
[3] AIB, http://anagrafe.iccu.sbn.it
[4] ICCU, http://www.iccu.sbn.it

those libraries now offer their services through the Internet, for example as remote support or online, digitized versions of part of their collections. The most widespread and important of those services is the free browsing of the Online Public Access Catalog (OPAC)[5].

The Italian IT companies in this sector that market products and services, and that also produce and supply certified applications to be used in the library network of the National Library Service (Servizio Bibliotecario Nazionale, SBN)[6], managed by ICCU with agreements with the Regional governments, are only 15 in number.

As regards the investments in the library sector, according to a nationwide analysis, it is noteworthy to mention that investments in general, thus including those pertaining to automation, have dwindled lately due to the economic-financial crisis.

In Italy, the presence of SBN strongly affects figures and features of the market. The adoption of the SBNMARC[7] protocol has, in fact, forced the companies that wanted to keep being competitive to pursue that direction in their development. The tendency of many public institutions (especially at the regional level, subsequently involving local organizations) was to adhere to SBN, since the communication function is becoming a vital aspect for any software aimed to succeed in Italy

The technological evolution of the current software systems suffered from a slower "innovation pace", due to the low margins offered by the market. Currently, the market is in a phase of "almost-maturity" in technology.

The most important aspect addressed recently is the integration of the information from catalogs in libraries with the information directly available on the Internet. The development now is focused on OPAC, and, more generally speaking, on the search tools available now, in comparison with all-purpose search engines used on the Internet (starting, obviously, from Google).

The solutions are many, but they all tend to take into account several needs. On the one hand, the need to expand and widen the information that can be searched (enrichment, federated search, access to databases); on the other hand, the presentation, that borrows typical systems of the Internet (tag cloud, conceptual associations, users' comments, and so on). They follow two tendencies, sometimes coexisting: the creation of portals and collections (usually customizable) of resources that crowd one page, and the proposal of a very lean and simple, Google-like, interface. As it had begun to happen in the early years of the World Wide Web, some international producers became specialized in building tools that complement or overlap with the actual OPAC, and promote those tools as solutions that can be adopted even by users of other systems, as third-party products. Furthermore, the high specialization and specificity of library standards are a significant technology barrier to the entry of new competitors.

3 Description of Research Project

The proposed project is a software devised for the net, entirely web-based, with the

---

[5] OPAC, http://opac.sbn.it

[6] SBN, http://www.sbn.it

[7] SBNMARC, http://www.iccu.sbn.it/opencms/opencms/it/main/sbn/evoluz_indice_sbn/pagina_147.html

management of all data outsourced in the SoSeBi server infrastructure. It will be supplied to users as a service (Software as a Service, SaaS), according to the increasingly widespread method among IT services that lets the customer be in the position of using the software without having to deal with what is “behind” it. The customer will access the platform in the same way they access a website, using a simple browsing program, and all the data they will enter and the related elaboration received will be handled by SoSeBi through an advanced company technological infrastructure.

The technology used for the software part of the server will be Microsoft's .NET framework, the ASP.NET language on Windows Server platform. Hardware-wise, a cutting-edge server will be used as the base technology, with an Intel multiprocessor and a large quantity of dedicated RAM, combined with a data redundancy system (discs in RAID 1 configuration and backup on external unit). Using this technology, it will be possible to switch to a more powerful infrastructure at a specialized national supplier. Internet connectivity will be of a 2 Mbit/second guaranteed minimum band, which could scale up to 100 Mbit/second on demand.

Regarding product innovation, the objective is to merge four fundamental aspects in one software. These aspects are expected to have a large development in the future.

1. Management of documents according to library science;
2. Management of digital contents and related access rights;
3. Social network;
4. Semantic search into electronic documents.

Some of those aspects, like the digital content management and semantics, have just become a part of citizens' habits. This can pave the way for new scenarios in terms of behavior and use of those technological tools.

The first distinctive feature will be the homogeneous integration of the four aspects stated above, called macro-functions. The dialogue method among the four macro-functions will be innovative, and allow, for example, to integrate the traditional library accuracy with Web 2.0 trends, namely reviews and tagging of books, and social relations between users and libraries.

In particular, each macro-function will have the following functions.

1. Management of documents according to library science. Careful management of bibliographic information according to library science standards and national industry rules. Certification of the application at Level 4 of the SBN dialogue protocol released by ICCU.
2. Management of digital content and related access rights. Integration with multimedia content management platform, with eBooks, audiobooks, digital audiovisuals, images, digitized content. Their management includes a description of their content, search, and viewing, also in streaming for audiovisuals. Access to contents is regulated at the source by a copyright and Digital Rights Management (DRM) control system, so that the legal and commercial boundaries established by the publisher will be respected.
3. Social Network. Creation of a network of potential social relations among the various actors allowed the use of the software: libraries and readers. The main reason for the social network is the sharing of an interest in books, expressed in different ways by the participants, where libraries showcase their catalogs, users

comment and vote books. A functional ecosystem centered on books, where social relations can be built and expressed, giving added value to the traditional catalog search of a library.

4. Semantic search in electronic documents. Integration of a language analysis engine (morphology, grammar, logic), and semantic disambiguation of texts. Use of a semantic network made primarily of a thesaurus of concepts and relations that would provide a conceptual representation of the language. Automatic categorization of documents, interpreting their content, in certain categories. Extraction of data from indexed texts, with normalization and transformation in metadata.

The activities covered in the project were anticipated by a feasibility study in order to analyze the information needs related to the project, which was generally defined during its planning stage. Among the objectives of the study were the following.

1. Find one or more architecture solutions related to the applications, the technologies and organizational solutions;
2. Propose technical-organizational solutions;
3. Provide the management with enough evaluation tools to decide on the operative completion of the project.

3.1 Project Subdivision

The project covers a number of operation stages. Every stage of the working plan is organized in Work Packages (WP), parallel phases in which operation objects are reached with work group activity, through the production of expected Results and Products and the application of a specific Methodology. The WP included in the projects are six:

1. System Architecture
2. Content Management
3. Social Network
4. Semantic Engine
5. Fundamental Research
6. Experimentation

Below is a brief description of each phase in the development of the project.

3.1.1 System Architecture

The future scenario in which libraries will have to work will be analyzed. That scenario will see an increasing centrality of the Internet and its digital services to meet the needs of both physical users and virtual users, who access the library remotely.

Although several software development models exist, we chose to perform a SWOT analysis limited to the Spiral and Agile [3] models, as major representatives – in software engineering – of iterative methodologies and, especially in the latter case, modern ones. The methodology will therefore be the Agile development model, since it is more in tune with the new demands of the market/customers.

The architecture will have to possess a high flexibility, and will have to be able to

incorporate advanced functions, resulted from research, in a modular and incremental way (for example, an evaluation of the adequacy of federated or plug-in based solutions could be useful).

3.1.2 Content Management

An interesting topic to study in this field is certainly the management of the metadata associated with multimedia objects. In this field, a significant role is given to ontologies, which define a formal, shared, and explicit representation of a conceptualization of any knowledge area. They are powerful tools to describe the entities of any domain, and the relations among them. Consistently with this view, it is necessary to guarantee an appropriate management of the semantics of metadata, so that the same entities, properties, and relations of the domain with which the system interacts are denoted.

The guiding purpose of this study lies in the definition of the semantics of metadata associated to multimedia contents by using an ontology. In particular, it is necessary to consider the possibility of defining specific ontologies for the domain of interest in an automatic or semi-automatic way. Alternatively, given an existing ontology, for example manually written by domain experts, the possibility of expanding that ontology with new metadata automatically extracted from a known set of multimedia documents can be considered.

3.1.3 Social Network

The functionality of the Social Network module must be defined taking into account the dialogue between readers and libraries, introducing the well-established tools of library tradition into the Web 2.0: OPAC, loans, reference. It is necessary to improve the dialogue between libraries and their users, providing an innovative service, which will bring them into closer contact with the libraries, thanks to the Web 2.0. The idea is to target Web users that not necessarily are also users of one of the libraries in the network. They would be able to sign up to use the social function and communicate on the Internet, sharing their taste and preferences; they could eventually become users of a library, after having come into contact with its services through the social network.

This macro-function is meant to bring forth a social evolution of the traditional online catalogs, shifting the focus on the customers' views and their expectation in terms of Web 2.0.

3.1.4 Semantic Engine

Semantic search is a field where many studies are focusing, considering the enormous and ever-increasing volume of documents currently available (especially digital documents). Academic research in this field will be fundamental, because semantics is still a young science, with a strong connection to Information Technology, and which takes part to a heated debate with semantic Web. It will thus be necessary to develop a precise outlook to understand what will become of the dialogue between semantics of libraries and semantics of the Web. These two aspects, in fact, are going to become

more and more interconnected in the future, given the increasingly faint line between owned documents and accessible documents on the Web.

3.1.5 Fundamental Research

This phase will encompass all the activities at the DIEE of the University of Cagliari oriented to fundamental research:

1. the definition and study of a metadata ontology for multimedia objects;
2. the definition and study of graph analysis techniques applicable to Social Network;
3. the definition and study of recommendation techniques on Social Network;
4. the analysis and study of techniques and models of semantic analysis;
5. semantic classification from available information coming from thesauri, subject indexes, summaries, and classifications.

The applied research will be based on the methodological results of this research, guaranteeing innovative functions of the product..

3.1.6 Experimentation

The technical activities related to the creation of the software prototype will be verified and validated appropriately, regarding both code and execution. The verification plan will be connected to the definition of software requirements according to the Agile Methodologies. The draft of the validation plan will aim to allow the verification of product requirements along the entire design phase. Once the whole system is broken down into different modules, three test phases will be performed:

1. module test, to test the components of each module and verify the communication of components belonging to the same module;
2. integration test, to test the integration and interaction mechanisms between each module;
3. system test, to test the whole system and come to an acceptance test of the final prototype.

4 Project Schedule

The project officially began on March 6, 2013, and its conclusion is estimated to be on March 5, 2016. Currently, the prototype covered in the project is in its implementation phase.

Among the activities performed up to now, special attention was paid to the the management of metadata associated to multimedia objects, leading to the definition of a taxonomy of metadata for the representation of multimedia objects managed by a digital library.

The approach we followed included the application of 3 distinct phases. The first phase consisted of the analysis of metadata standards used as a reference to represent information associated to multimedia content. In order to cover the reference domain,

made of multimedia objects of interest in a library sector, metadata standards that allow the best and most complete description are used [4][5][6]: the choice fell on the Dublin Core [7][8], XMP [9] e MAG [10] standards. The first provides a general description that can refer to content of any kind (a feature that helped its large scale usage); XMP focuses on more accurate and peculiar characteristics to multimedia content, such as audio and video content, other than images, for which it uses Exif metadata (specialized in digital images) [11]; MAG, instead, is an application profile that, included in our taxonomy, allows to communicate with the Internet Culturale portal without interoperability problems. The portal, in fact, uses MAG for the representation of its contents. The data sheets of the Italian Central Institute for Cataloguing and Documentation (Istituto Centrale per il Catalogo e la Documentazione, ICCD[8]) offer the opportunity to perform a mapping between the metadata in our taxonomy and the PICO application profile used by the portal Cultura Italia, thus increasing the interoperability of the system [12].

We also chose to consider some tags for the management of User-Generated Content (UGC) [13]. Those standard are often complementary, covering information that they could not represent singularly, and sometimes redundant: in that case, a direct mapping was performed.

As regards the second phase, a study on the resources to represent was carried out, using the following sites as reference: Europeana[9], Internet Culturale[10], Cultura Italia[11], Internet Archive[12], Open Library[13], and Project Gutenberg[14].

The third phase involved comparing metadata taken from the standards analyzed during the first phase with the data collected during the second phase. The purpose of the comparison was to verify whether all the characteristics studied during the second phase were represented by the metadata retrieved during the first phase.

Moreover, we studied how to enrich the available metadata in order to manage some often neglected aspects, through a new study on the objects available at the Mediateca del Mediterraneo (MEM)[15] of Cagliari, and through the study of the copyleft licenses Creative Commons[16] and Copyzero X[17]. Topics such as the completeness of information on eBooks, the identification of documents as belonging to gray literature, and the management of rights on the cataloged resources were addressed.

Thanks to the adoption of the MAG application profile, the taxonomy guarantees the interoperability with the portal Internet Culturale, that grants access to the wealth of public libraries and prestigious Italian cultural institutions. The presence of the DC standard allows also to be Open Archive Initiative (OAI)-compliant, with the opportunity to exploit the interoperability protocol Open Archive Initiative Protocol for

---

[8] ICCD, http://www.iccd.beniculturali.it
[9] Europeana, http://labs.europeana.eu
[10] Internet Culturale, http://www.internetculturale.it
[11] Cultura Italia, http://www.culturaitalia.it
[12] Internet Archive, https://archive.org
[13] Open Library, https://openlibrary.org
[14] Project Gutenberg, http://www.gutenberg.org
[15] MEM, http://www.comune.cagliari.it/portale/it/studisardi.wp
[16] Creative Commons, http://creativecommons.org/
[17] CopyZero X, http://www.costozero.org/wai/copy0.html

Metadata Harvesting (OAI-PMH) [14], and ensuring the communication with another important portal, Cultura Italia.

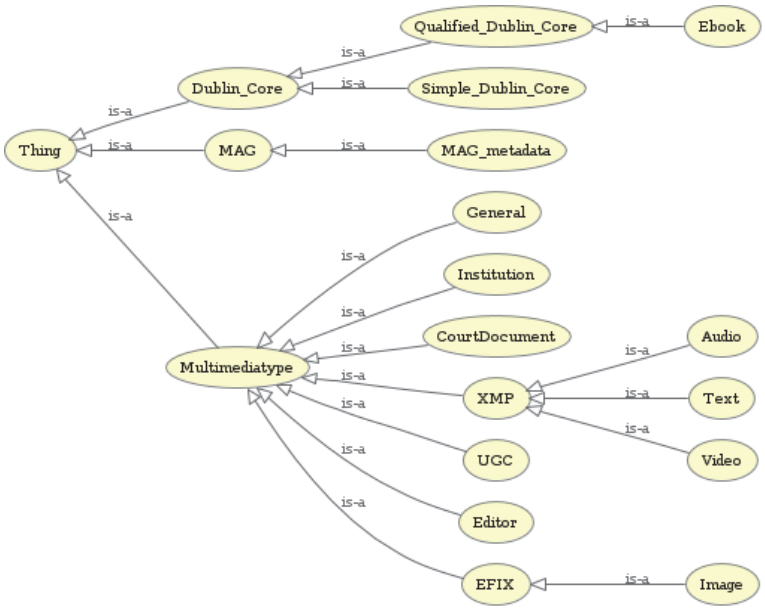


Fig. 1. The resulting metadata taxonomy.

The information network with which the system adopting the taxonomy will be able to interface, and the set of metadata studied to cover the general, main information and the specific information of multimedia content, will make its use suited to a digital library aspiring to modernity. Such a library would become able to manage often neglected pieces of information (gray literature, UGC, management and representation of rights on the resource), and could become part of an information and cooperation network with the most important cultural portals of Italy.

5 Conclusion and Prospects

This project is coherent with the strategic objective of the regional planning in Sardinia, since it aims to implement innovative methods of the ICT sector in the library industry, and it complies with the objectives described in the Regional Strategic Document (Documento Strategico Regionale, DSR) 2007-2013[18] for Sardinia, which states the necessity of promoting the adoption of ICT in order to fill the gap that makes Sardinia less competitive in innovation, although the region stood out in Italy for having started innovative initiatives ahead of its time, marking a sharp passage from a traditional economy to an innovative economy, based on knowledge and inno-

[18] Regione Autonoma della Sardegna, Documento Strategico Regionale (DSR) 2007-2013, http://www.regione.sardegna.it/j/v/17?&s=1&v=9&c=4756&na=1&n=10

vation. The project is meant to become a part, albeit small, of the context of the competitiveness of Sardinian companies, with its purpose to create a software for the automation of cultural services.

The Italian production system is characterized by a high volume of very small-sized companies in every sector, which work outside cooperation agreements or networks. This structural characteristic does not allow these companies to compete on the markets of goods where price, which comes from production costs, is the main factor to acquire and maintain market shares.

The limited availability of funding will strongly affect any new investments made by libraries in goods and services, and will force them to question the worth of maintaining current services, including computer-based services and the automation of bibliographic catalogs. The market prospects regarding this particular aspect, that is the reduction of management costs of technical services, could be favorable for those companies that are committed to creating and supplying more effective software, more competitive in terms of price, and technical support services that would help institutions to save on service cost keeping the same or a higher quality level.

As regards the demand, the product to be created is clearly geared towards a target of customers that have already completed the first steps in service automation, thus it has to meet the needs of a well informed user base with high expectations in terms of performance of management systems.

A new product with the features described in the previous sections has great market prospects, especially in Northern Italy, where the demand for innovative products is stronger and where more libraries, more resources, and more well-established cooperation networks among organizations are located, to provide advanced services to the citizen.

Acknowledgements. Simone Porru gratefully acknowledges Sardinia Regional Government for the financial support of his PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007-2013 Axis IV Human Resources, Objective l.3, Line of Activity l.3.1).

References

1. Candela, L., Castelli, D., Pagano, P.: History, Evolution, and Impact of Digital Libraries. E-Publishing and Digital Libraries: Legal and Organizational Issues, E-Publishing and Digital Libraries: Legal and Organizational Issues, IGI Global (2011) 1-30.
2. Fox, E. A., Gonçalves, M. A., Shen, R.: Theoretical foundations for digital libraries: The 5S (societies, scenarios, spaces, structures, streams) approach. Synthesis Lectures on Information Concepts, Retrieval, and Services, 4(2) (2012) 1-180.
3. Poppendieck, M.P.: Lean Software Development: An Agile Toolkit. Addison-Wesley Professional (2003).
4. Lagoze, C., Van de Sompel, H.: The Making of the Open Archives Initiative Protocol for Metadata Harvesting. Library Hi Tech Volume 21 (2) (2003) 118-128.
5. Scherp, A., Eißing, D., Saathoff, C.: A Method for Integrating Multimedia Metadata Standards and Metadata Formats with the Multimedia Metadata Ontology. International Journal of Semantic Computing, 6(1) (2012) 25-49.
6. Zengenene, D.: Global interoperability and linked data in libraries. New Library World,

114(1/2) (2013) 84-87.
7. Hillmann, D. I.: Using Dublin Core. Dublin Core Metadata Initiative Recommendation (2005). Retrieved from: http://dublincore.org/documents/usageguide/
8. Hutt, A., Riley, J.: Semantics and Syntax of Dublin Core Usage in Open Archives Initiative Data. Joint Conference on Digital Libraries (2005).
9. Adobe Systems Incorporated: Adobe XMP Specifications, Additional Properties (2010). Retrieved November 13, 2014, from http://www.adobe.com/content/dam/Adobe/en/devnet/XMP/pdfs/XMPSpecificationPart2.pdf
10. MAG, Comitato: Metadati Amministrativi e Gestionali: Manuale Utente, Pierazzo, E. (ed.), version 2.0.1. Roma: ICCU (2006).
11. Technical Standardization Committee on AV IT Storage Systems and Equipment:. Exchangeable image file format for digital still cameras: Exif version 2.2. Published by Standard of Japan Electronics and Information Technology Industries Association (2002). Retrieved from: http://www.exif.org/Exif2-2.pdf
12. Buonazia, I., Masci, M. E., Merlitti, D.: The Project of the Italian Culture Portal and its Development. A Case Study: Designing a Dublin Core Application Profile for Interoperability and Open Distribution of Cultural Contents. In Proceedings ELPUB 2007 Conference on Electronic Publishing (2007) 393- 404. Retrieved from: http://elpub.scix.net/data/works/att/114_elpub2007.content.pdf
13. Pani F. E., Concas G., Porru S.: An Approach to Multimedia Content Management. In: Proceedings of The 6th International Conference on Knowledge Engineering and Ontology Development, KEOD (2014) 264-271.
14. Lagoze, C., Van de Sompel, H.: The Making of the Open Archives Initiative Protocol for Metadata Harvesting. In Library Hi Tech, 21(2) (2003) 118-128.

A Numerical Data Standard Joining Units, Numerical Accuracy and Full Metadata with Numerical Values

Joseph E. Johnson

Department of Physics, University of South Carolina, Columbia SC, 29208, U.S.A.

Abstract. Numerical measurements have little meaning without the associated units of measurement, level of accuracy, and defining metadata tags. Yet these three associated items are scattered in rows, columns, references, the title, and other positions in data tables separated from the values themselves and usually immersed in unstructured text. Thus unambiguous electronic readability is not possible but requires human preformatting to link these three data components to the value for computer processing. This paper proposes a tight linkage among numerical values, units, accuracy, and meaning as a string expression that provides a numerical standard which we call a "metanumber" along with a set of requirements for this structure. We require that this metanumber string object be instantly readable by both humans and computers. Our work then develops the requisite algorithms for automated computer processing of the metanumber expressions resulting in a new metanumber where (a) all units and dimensional analysis are automatically processed, (b) numerical accuracy of results are computed, and (c) unlimited metadata tags and structures provide a trace of all historical operations with component meaning providing both the exact evolutionary path of each computed number, and a unique name (as an internet path) for every single measured v4alue. We also propose a flexible table-like structure for archiving all numerical data as metanumbers that allows the automated sharing of data among users. This structure can serve as a foundation for high –speed data exchange removing the costs, errors, and time delays required with human preprocessing and thus serving as a critical component in Big Data processing and as a foundation for advanced artificial intelligence. Other features are explored including an optional proposed expansion of the base units for the SI (metric) system as well as user defined units that are natural for users with commercial, industrial, medical, and scientific problems. These components offer a transformational increase in computational power in every domain of inquiry.

Keywords. Units, Dimensional Analysis, Metadata, Numerical Uncertainty & Accuracy, Metanumber, Error Management, Data Tags, Numerical Information Standards, Si & Metric Units, Big Data.

1 Introduction

Numeric data by itself, except for pure numbers, is meaningless without three associated components: a. units of measurement, b. accuracy level (numerical uncertainty), and c. the associated defining and modifying information. All these are requisite for the unambiguous meaning of numerical values. Yet in every domain, numerical data tables have these three components scattered in the title, row &

column headings, and footnotes. Sometimes such auxiliary data is just assumed to be apparent. There is no universal standard for the attachment of such metadata to numeric values or even a unique computer readable symbol structure for these components. The resulting arbitrary positions, fonts, and formatting styles prevent unambiguous automated electronic reading and requires that all data must be manually converted to a useful form by each user prior to their computer processing each and every time it is used thus incurring substantial costs, time delays, and associated errors. The reason we believe that no standard has emerged is that no single convention seems preferable or natural. Also storage space, in both print and electronic representation, has historically always been a primary consideration, so the compressed formats that are currently used were historically optimal for printing, direct human viewing, and to minimize electronic storage. A substantial component of this problem is that the units, accuracy, and defining tags are not only separated from the values, they are also embedded in text thus making it extremely difficult to extract their content by computer algorithms. Those space considerations are no longer an issue with current technology. Having a numerical metadata standard for automatic electronic data processing would not only remove all of these problems and lead to fully automated sharing and processing of all numerical data, but could lay the foundation for far more advanced and intelligent automated analysis and decision systems, that will not require human preprocessing or intervention. The standardization can be implemented automatically from the initial observation or measurement device when the data is originally recorded. Such a solution is essential for the emergence of a new level of artificial intelligence. When any numerical value is retrieved, using a unique name, it arrives with units, uncertainly level, and complete metadata so that logical and numerical processing is executed not only for dimensional and error analysis but also with advanced metadata tag management and tracing of the historical evolution of numbers. So there are two problems that have to be simultaneously addressed: (a) create a standard for such integrated metanumber formats so that they are easily read by both humans and computers, and (b) create the algorithms for mathematical and logistic processing that evaluates all expressions among such objects.

2 Proposed Solution

The author here proposes standards for attaching the units, error, and all other descriptive metadata, to numerical values along with the associated three software algorithms that mathematically processes such extended information [1]; [2], [3]; [4]. We call such objects “metanumbers” to denote their informational completeness. Our integrated algorithm satisfies our criteria that these metanumbers must (A) be easily readable both by humans and by computers, (B) require the minimum of storage, (C) process at the fastest possible speed, (D) be easily extensible to allow for user defined units, and (E) support computation at all levels from simple calculations as web “apps” on internet devices to unlimited “Big Data” applications in large scale computing. It must also (F) automatically manage all computations among metanumbers with dimensional analysis to trap invalid unit combinations as well as (G) compute the associated accuracy (numerical uncertainty) of the result using error

analysis with user selectable means of combining error. It must (H) provide a simple rapid means for the one-time initial formatting of data into the requisite metanumber format, and (I) allow one to easily retrieve and use all past results. Finally, (J) the algorithms must allow unlimited metadata to be linked to numerical values and still manage that attachment with the absolute minimum number of characters and without any reduction in the processing speed. The author has developed a methodology for a simple one-time reformatting of numerical data sets into such a standardized metanumber form that, with the algorithms we have developed, will satisfy each and all of the listed requirements and still provide unlimited traceability of all metadata linkages and origins.

3 Historical Development & Current Status

Early versions of the author's unit and dimensional management software were built and deployed on a web site that has been in use for over a decade and is still used by students in several university courses but it was limited. Recent developments by the author led to (a) a new optimized design methodology that vastly extends numerical operations with units and dimensionality analysis including (b) the reuse of any previous expression, (c) improved means for output of an expression in any desired units by using a resolution of the identity, (d) an improved means for users to define personal units for their specialized domain of work, (e) a new method of error analysis with an awarded U.S. patent to the author, and (f) a new innovative design for optimally managing unlimited metadata attachments to numerical values. This new algorithm can also support a ten-fold increase in processing speed using a parallelization of the code over multiple processors for Big Data problems. Then a future stage will support a dedicated accelerator chip for ultra-fast processing. The author has completed the full design, programming, and testing of both the units and metadata management components along with several important metanumber tables in a fully operational system. This software and methodology is now fully tested and operational on our central server as a cloud application available from any internet device. The central server approach is necessary in order to support multiple users, rapidly deploy code upgrades, and, most important, to share a rapidly growing interdisciplinary standardized library. Such a cloud system allows use by any internet linked device, (tablets/iPads, smart phones, PCs) as well as large central systems. The software has been developed in the Python language which is open source with extensive additional components, and is a widely accepted very modern framework for rapid development. The addition of a Python algorithm for the processing of numerical accuracy has also been fully integrated using the "Python Uncertainty" system developed by Leibigot and associates. The final phase of building user interface tools is currently active. We will discuss each of these three active components separately.

4 The Units (Dimensional Analysis) Algorithm

4.1 Units Introduction

There are a large number of 'unit conversion programs' and some that can process expressions containing multiple units. Our algorithm however is optimal with both maximum speed and minimum space whereby one simply attaches unit names as variables within any valid algebraic expression such as (4.3*ft/s+7*meter/hour). The 'value' of each unit and constant provides the instantaneous requisite conversion with effortless readability for humans. With this algorithm, each unit name or constant is a unique variable defined in terms of the foundational metric (SI) system units. The units algorithm contains most normally encountered units (over 800 units and essential fundamental constants). It also supports the introduction of other units as defined by users in terms of existing units and constants (such as the lightyear). These results have led us to change a number of traditional notations.

4.2 Units Standards

In order to be compatible with modern computer languages and optimal electronic processing without ambiguity, we define all units in lower case with standard alphanumeric characters, without Greek characters, without superscripts or subscripts, and without any other special characters, symbols, or fonts. The past denotations for unit names and constants lead to ambiguities in different computer languages and thus we seek unit and constant names that are expressible in simple lower case alphanumeric characters. Specifically we remove the capitalization of all proper names assigned to a unit and thus we write "newton" or "nt" rather than "Newton" or "Nt" because the value of a variable in modern computer languages is case sensitive for alphabetical symbols. Likewise the potentially ambiguous encoding of units and constants with multiple fonts, superscripting and subscripting, and of foreign alphabets is removed. Thus we use "ohm" rather than Ω; and hb (h-bar) for Planks constant divided by 2π. Pi is likewise written as pi rather than as π. Our motives are format consistency, adherence to modern software convention (variables are lower case), the use of only standard alphanumeric characters for variables, and to remove the ambiguity that can result from font and format imbedded information. Although a few of the most common units are allowed in the plural form, we generally restrict the spelling to only the singular form thus removing some ambiguities in spelling and simultaneously making the underlying code shorter. Our core units algorithm is essentially the same in any language (C/C++, Java, Ruby, Python…) and can be easily enhanced by parallel processing on a dedicated accelerator chip [5]. Our algorithm catches all exceptions for dimensional errors, and returns expressions in standard format for use in subsequent expressions. These in turn, can be referenced with a unique name for insertion in any future expressions and are functionally valid although they are expressed in different units. All computational history is maintained for each user.

4.3 Units Features

Mathematical expressions of values are returned in metric unless other units are specified using the format "(expression) ! (units desired)" . The speed of light in miles per minute would be written as c!(mile/min). Formally the algorithm multiplies the expression by (desired_units / desired_units) which is technically unity and does not change the value. The software divides the value of the expression (in default units) by the desired_units and then writes the final expression as "final_expression * desired_units". The effect is to both divide by the desired units then multiply by the desired units. So although the value is numerically different when expressed in different units, it is computationally equivalent to the original value since it is now multiplied by different output units and either expression will evaluate equivalently. For example 35*acre*!(ft2) would give the value of 35 acres in square feet and then multiply that value by square feet (ft2).

4.4 Units Base Si System

The units algorithm is defined and based upon the SI (metric) system in terms of meter, kilogram, second, ampere, kelvin, and candela. All the other primary units and core constants are then defined in terms of these allowing also for altered spelling and abbreviations such as m for meter, sec and s for second, etc. Dimensionless prefixes (kilo, mega, centi, million, dozen, etc. can be mixed with units in any mathematically valid manner thus eliminating the need for joined definitions. The 800 or so primary units and physical constants are then each defined in terms of previous units, sequentially back to the fundamental metric units. Values assume the accuracy level of the computational environment while the accuracy is indicated as the median value with an associated accuracy (uncertainty).

4.5 Powers of Base Units

In the original definitions we also define m2= m*m; and m3=m*m2 up to the fourth power in order to assign basic unit variables that have these multiple powers in an abbreviated manner. We likewise define the inverse powers of the base SI units as m_2 = 1/m2, m_3 = 1/m3 also up to the inverse fourth power. If any mathematical expression is then well formed using unit names then the result will be returned in SI units by default. Thus (34.6*acre*(1.2*ft+7.2*inch-3.4*cm)) will give the computed volume (numerically) in cubic meters (m3). In order to obtain that volume in gallons, one would instead enter: (34.6*acre*(1.2*ft+7.2*inch-3.4*cm)) ! gallon.

4.6 Standardized Metanumber Data Tables

The core fundamental constants and primary units are defined within the metanumber program. All other numeric data, user defined units, and special constants are defined in tables of various dimensionalities using the format [server-path_ directory-path_table-name_row-name_column-name…)]. For example the elements table (which has the unique abbreviated name of "e"), contains all elements in rows, with

columns providing over 50 associated properties. This path name (currently using our default server) of [file_row_column] can itself be used as a variable to retrieve that metanumber value from the stored data tables. Thus [E_Gold_Density] / [E_Silver_Density] computes the ratio of the density of gold to that of silver. All stored data is to be standardized and retrievable by its unique internet path name as here described.

4.7 Unique Path Names

These internet path names are unique, and thus exactly define the metanumber of interest as well as provide an unambiguous name for that single unique metanumber. Furthermore the […] structure can be used as a variable in any mathematical expression, function, or algorithm to retrieve that metanumber. For the file name, row name, and column name, the comparison is made with all white space removed and all case changed to lower case. We envision that all numerical data be standardized as metanumbers in tables appropriate dimensionalities: zero dimension (a single value), or one dimension (a list of values such as the masses of objects), or two dimensions such as the elements table, or even other data tables of any higher dimensionality. The metanumbers in each of these tables are to be stored in comma separated values (CSV) formats thus allowing easy viewing and editing in spreadsheets such as Excel and avoiding any extraneous symbols in any type of markup language which could be problematic with the retrieving algorithm.

5 The Metanumber Algorithm

5.1 Metadata Overview

By “metadata” we refer to the collection of all information (other than units and accuracy) that describe, define, and explain the meaning of the associated metanumber. By “metanumber” we refer to a single numerical value with attached units, accuracy level, and its metadata that describes it. We propose that all numerical data be standardized as metanumbers and stored in archived tables as described here. In the last section we concluded that the units are to be attached as variable names in valid mathematical forms to the numerical value as for example with the velocity of 3.759e4*m*s_1. Here we discuss the metanumber archive tables as mentioned above with greater detail on how metadata is stored and associated with a value. When one normally stores structured numerical data, it can be as a single value, a list of values, a rectangular table of values, or an array of higher dimension (such as when economic data has an associated date or location). Our methodology applies to data arranged in any dimensionality. But data which is laid out as a two dimensional table is the most common with a format like “Entity vs Property”. The “ Entity” might be the 110 chemical elements corresponding to the rows, and where the “Property” might be any of 40 to 50 different properties given in columns, such as density, atomic number, atomic mass, thermal or electrical conductivity, heat capacity, melting point, boiling point , etc. Likewise, the rows might specify a person’s ID and the columns might be properties such as medical data for each person such as DOB, weight, pulse, blood

pressure ... We will use the two dimensional table as an example. The format is to be in rows of comma separated values. This is the most common and open format, and it is easily imported and exported from Excel and other editors as well as read by Python. It is also structured so that at any point it can be automatically loaded easily into any relational database such as Oracle, or SQL if desired by a simple program and fully automated.

5.2 Metadata for a Table as a Whole

Some metadata values are associated with (a) an entire table, while other metadata is associated with just the (b) entries in a given column, (c) or entries in a given row, or (d) perhaps only associated with a specific value (such as longitude and latitude or date-time). To capture metadata associated with the table as a whole, we standardize the first two rows of each table with row one containing the names of fields describing the table, while their associated values are contained in the second row in the corresponding column. This 'table metadata' will contain the table name, a unique table abbreviation, the source(s) of the data such as a web address, the table creator's name, and email address, date created & last updated, security level and codes, references, footnotes, and other associated data related to the collective data in the table as a whole. All of this data is metadata that pertains to every metanumber in the table. Thus this data is applicable and linked to all values in that table.

5.3 Metadata Associated with Specific Rows and Columns

The first column (beginning in row three) is to contain the unique row (entity) name/index identifier (such as the element name of gold) while row three will contain a unique column (property) name/index identifier (such as Thermal Conductivity). Other columns (after the first) and other rows (after the third) can contain other metadata associated with that row or column. For example if the third row has an entry of "Thermal Conductivity" for a given column, then several of the rows beneath the heading of Thermal Conductivity could contain references to the source(s) of those data values, web links to discussions of thermal conductivity, associated equations and web links to explanatory videos or notes. Units would preferably be attached to the end of each numerical value or one could have row or column headings labeled as "*units" in which case the units in that column or row would be distributed by multiplication over the values in the associated row or column. The reference [E_Gold_Thermal Conductivity] will find the unique table named "E" referring to the elements, then find the row labeled "Gold" in column 1, and the corresponding column labeled "Thermal Conductivity" in row three. The metanumber string for gold's thermal conductivity will then be retrieved from that table, that row, and that column and the metadata that pertains to that table, associated with that row and that column are linked to the value without having to be transferred into expressions. Although the heading is Thermal Conductivity, the search will compare the name with all case lowered and all white space removed. And thus it would match 'thermalconductivity'. A comparison of lowered case with removed white space is used for all components of the [table_row_column] lookup. Tables of metanumbers cannot contain any set of the symbols ", _ {}()[ ]" as each of these has a reserved use.

When metadata occupies a row or column, this is to be indicated by having the name of that row or column prefixed by a "%". If the metadata values are unique and could be used as an alternative index then the prefix is to be "%%". An example with the elements table is that both the atomic number and the symbol are unique metadata and thus those columns are labeled as %% Atomic Number and %% Symbol.

5.4 Metadata Associated with Individual Values

Finally, there are times when specific metanumbers have an associated metadata just for that one value such as a longitude-latitude or date-time of measurement. This value specific metadata is encoded in the format {var1=value1|var2=value 2|…} which multiplies the associated metanumber and which always evaluates to unity in any mathematical expression and thus disappears. Thus this expression serves as a wrapper to carry any information to be attached to a specific numeric value. By using the reference [E_Gold_Thermal Conductivity] as a variable, one has therewith a unique name for every archived metanumber along with potentially vast associated metadata that is linked but not transferred until requested.

5.5 Data Archived on Other Servers

When data is stored on other servers (as opposed to the central cloud system in the default directory) then the table name is to be preceded with "(the web address of the server) _ (director path to the file)" followed by a double underscore '__' and then "(the table-name_row_name_column-name)". This gives unlimited capacity to the coding for data retrieval as is done with web addresses with a unique name for every single metanumber value.

5.6 A Distinct Pathname for every Metanumber

A powerful feature of the reference string […] is that it gives a unique name to each metanumber via the unique internet path to each separate value of metadata in the table. Thus the metanumber name links to the table metadata, row metadata, and column metadata without having to include that metadata with the value in the expression. This is a critical concept. Furthermore, as all computational history is archived, this structure supports an essentially unlimited unique line of tracing the meaning, method of measurement, and unlimited other metadata via the indirect reference to the numerical value. This can be of great value with the properties of pharmaceuticals, where the non-numeric metadata can give critical associated data such as batch and expiration data, or in accounting for the tracing of the origins of funds and the means by which they were processed to give the current value. A separate application easily reformats existing data tables. Once this conversion to standard metanumber form is done, one can operate with the data at a speed and accuracy never before possible with full automation and no human intervention. This is especially critical with extremely large data sets and even more so when there are a very large number of data tables from multiple sources. Finally, if data is formatted this way as it is created, by that creator, then it can be subsequently used and shared at

maximum speed and confidence by everyone.

5.7 User Defined Units and Archived Results

A users submitted expressions are always archived into a single table under the user's name. This "user's archive table" is created for each user and contains the values: Seq#, DateTime, Submitted Expression, Resultant MetaNumber, and UnitID. The Seq# is the unique submission sequence number for that user: 1, 2, …. The DateTime is in the format YYYYMMDD:HHMMSS. The Submitted Expression is the string that is submitted for evaluation. The Resultant MetaNumber is the metanumber that results from the evaluation of the submitted expression. The UnitID is an integer consisting of 10 single digits that contains a code for the powers of the fundamental metric units in a given order and thus which is unique for each combination of units. This integer is thus representative of the resulting concept and can be used for different types of analysis, classification of results, tracking of work done, and the creation of networks among users, metadata tabs, and concepts studied. One recalls that any string that is bounded by {} will evaluate to unity (1) and thus can multiply any metanumber at any position although the formal structure places it at the end of a metanumber to denote the specific metadata such as Lon Lat for that value. If one prefaces a submitted expression with {= expression name} such as {= Physics 211 Lab problem 4.38}, then later analysis of that users work, when downloaded to a PC into Excel, can be used to filter and sort different computational efforts. Within Metanumber, one can also reference the result as [my_expression name] in order to use it in other expressions. One can also reference past work as [my_ seq#]. More advanced features and operations are also possible. This means that all metanumbers that are used, whether they are tables using [name_row_column] or past input results as [my_name or #] provide unique names for every possible metanumber. In a sense, the ability to create a (one dimensional) table with [my_name or #] means that a users work creates a standardized table of all evaluated expressions by each user. As the form {=name} is maintained as part of the input expression, it follows that it as well as all component parts of the input expression with other metadata can always be searched for content. If one is careful to use unique neumonic names, then such results can constitute a user's own set of 'units' as units are simply values of metanumbers. It is easy to add additional user defined personal units and constants to the system by prefixing any metanumber expression with {=unit or constant name}*(metanumber expression) such as {=bluetruckvolume}*38*m3. A full discussion of these advanced features would take us beyond the scope of this introduction. The expression "[my_string]" is executed by the substitution of the users unique PIN address for "my" and then all the previous discussions on archived table data apply. Thus the term "my" is automatically replaced by the users PIN when used and thus is kept private for that person. Data and special units for a company or government agency can also be kept secure and private and thus shared only among members of a closed user group. The system allows the use of any previous result with the unique sequence number "i", referenced as [my_i] , in a new expression as a variable. Thus all the linkages back to the origin for every numerical value are maintained among contributing users and their work. The software can peel back layer after layer of how a given value was created along with all associated metadata

tags, assumptions, equations, accuracy levels, literally everything, in manufacturing a product, pharmaceutical, account, and scientific measurement, thus supporting levels of artificial intelligence never before possible and thus providing the complete evolutionary history of every number.

5.8 Current Archived Data Tables

Our interdisciplinary team has been creating sample metanumber data tables in several disciplines as examples for training. It is the identification of those tables and that data that is most generally utilized and shared among groups that will enable highly innovative computations crossing multiple disciplines supporting both research programs and student course work. Optimizing our methodology for these tables has been the main research thrust in parallel with refining our algorithms, notations, and user tools.

5.9 Conclusions

To conclude the metadata descriptor section, we have shown that by standardizing all metanumbers in tables with unique file, row and column names, that the internet path to the metanumber [file_row_column …] provides a unique name for entering the metanumber into a mathematical expression and also makes all action sequences fully traceable. The same functionality applies also to each users past work with [my_name or seq#] although this table is private to the user unless more advanced actions are executed for the sharing of data among users.. But an equally important feature is that all table, row, and column metadata is linked by that name enabling intelligent systems to form networks among tags, concepts, dates and locations, units, special constants, and even computational actions. Metadata for pharmaceuticals can describe side effects, dosages, batch numbers, expiration dates, and other metadata which can be automatically linked to recipient patients without the specific transfer of such extensive metadata in the computational process as one only needs the internet path. We will discuss these potential networks in the following.

6 The Numerical Accuracy Standard and Algorithm

6.1 Numerical Accuracy Overview

The most complex domain is the automated processing of numerical uncertainty which is a core research area of the author and is too technical for a full presentation here. The author has jointly published this research with a university colleague and also presented it in the AMUEM international conference on numerical uncertainty in Sardagna, Trento Italy [6] and has been awarded a U.S. patent for this algorithm. Numerical “accuracy” or equivalently “uncertainty” can be computed by several methods: A numerical accuracy algorithm could (a) treat the metanumber value as being exact. But with the exception of the integer counting of objects, values that are represented by real numbers cannot be either measured or represented exactly.

However there are a few “real numbers” that are by definition established as exact such as the speed of light or the number of cm in an inch. We could (b) treat the value with an uncertainty following the last listed valid digit when expressed in scientific notation and follow the standard statistical rules for determining the accuracy of the result for each operation given the accuracy of the operands. This method has the advantage that if one intercepts the raw data prior to processing, then one can count the number of digits present. We could (c) assume a normal (Gaussian) distribution to the value and combine the standard deviations of operands thus basically treating each value as a distribution instead of a real number which is done in the most exacting of cases and which is the most accurate. But the problem with this is that one almost never knows the actual standard deviation. Most measurements are made once or twice and time, effort and cost are not available to reasonably determine the standard deviation of the value except when different scientific teams expend a great effort over many measurements as with the NIST fundamental constants. It is furthermore an assumption that such normal probability distributions accurately represent the true distributions without skewness or kurtosis. Furthermore normal distributions do not close under operations other than multiplication and division meaning that the mean and variance as a pair do not close mathematically except in approximation. This is in contrast to the complex numbers which do close mathematically. We must choose the method of combining values where the accuracy level is simply indicated by an uncertainty of one in the last digit since that information is known and the standard deviation is not known. This method of representing error is always known as one has only a fixed number of valid digits in the value which is always observable. We will choose this method because it is always known and displayed. When a more accurate representation of uncertainty of the probably distribution such as standard deviation is available, the algorithm is able to utilize that framework. The very complex patented technique of the author is so exacting that the supporting information is almost never available and it will not be offered in this release version.

6.2 Use of Accuracy as Measured by Significant Digits

Even this method is in general very complex because one is now representing each numerical value with a pair of values: the known value as represented by a sequence of digits along with an uncertainty value. However it is essential that this error of the last digit be captured in the algorithm at the point when the number is introduced either (a) keyed in by a user or (b) listed in a data table. This is because all mathematical operations between the value of interest and other values in an expression will erase the knowledge of the accuracy of each and provide results of operations that only indicate the limit of accuracy being kept by that computer software. Computer generated results, without an accuracy algorithm, always generate a great overrepresentation of the true accuracy with excessive digits for the value as one can see when dividing say 17 by 3 when the result might only be accurate to one digit. Even this method of accurately computing the number of significant digits of the result is very complex algorithmically as different operations and different functions follow different rules for accuracy combinations and thus the expression must be parsed for the correct operational sequence and the values themselves cannot now be just real numbers but must be a new mathematical class of objects. Here we

are extremely fortunate that several such algorithms have been created for inclusion with the Python language.

6.3 Python Uncertainty Package

We will use the Python Uncertainty package which can be found at http://pythonhosted.org/uncertainties by Dr. Eric O. Lebigot [2]. Our procedure is this: Each data value when loaded from MetaNumber tables, or a keyed entry by the user must be immediately captured (using python code and 'regular expressions' methodology) and converted to a new type of object which is defined as a "ufloat" of a string which contains the known digits. Once this is done then this uncertainty algorithm evaluates the component values and maintains both the full accuracy of the result as though there is no error and simultaneously gives the resulting uncertainty of the numerical result. This algorithm also correctly mediates the operations between the ufloat values and the other values which are known exactly. To utilize the Python Uncertainty algorithm values, they must be identified by the software in order to be converted to the ufloat class of objects or must be identified as exact and thus treated as an exact standard value with no error. Our technique is to identify all uncertain values by numbers which contain a decimal point such as 5.793e5 where the value of 5.793 will be identified via the decimal point and converted to a ufloat and then multiplied by 1e5. An exact value must therefore be entered without a decimal by adjusting the exponent to remove the decimal as 5.793e5, if exact would entered as 5793e2 and thus stored as an exact real number. This methodology requires no new symbol or other indicator and is fully automatic. Thus the coding of numerical certainty or uncertainty does not need an explicit coding but rather uses this implicit coding methodology of our capture algorithm. This algorithm must act on all numeric values in both the data tables as values are brought into an expression and on keyed values that are entered by the user.

6.4 Conventions for Exact & Uncertain Values

We will take all unit conversions to be exact. The current algorithm automatically converts all table values to ufloat objects along with automatic conversion of all user keyed numerical components to ufloat based metanumbers when there is a decimal present as well as decimal present values that are keyed in by the user. For exact values, the user will only need to remove the decimal point by altering the associated exponent in the scientific notation in order to enter a value as exact, either in the tables or in the keyed values. Thus 3.2e4 becomes 32E3 with no decimal in the value to indicate that the value is exact. Keyed or table values that have a decimal present will be assumed to be ufloat objects. Any expression that contains a missing value will automatically return a missing metanumber with the metanumber code. Thus the result of this convention on encoding uncertainty is that no explicit notation is required other than the presence or absence of the decimal point in the value to get the value to be treated as uncertain or exact respectively.

7 Network and Cluster Analysis

7.1 Networks and Cluster Identification

Networks represent one of the most powerful means for representing information interrelationships (topologies) among abstract objects called nodes which are identified by sequential integer 1, 2,... Cluster analysis on these networks can uncover the nature of those structures. We have been able to link objects in this numerical standardization using our past research on the mathematical classification of networks and associated cluster analysis. A network is defined as a square (sparse) matrix (with the diagonal missing) that consists of non-negative real numbers, and which normally is very large, that represent the connection degree between node i and node j as C_{ij}. The mathematical classification and analysis of the topologies represented by such objects is one of the most challenging and unsolved of all mathematical domains.

We have built some far-reaching extensions of the metanumber system that can be constructed on the foundation laid by the standardization described above. Let us first consider the standardization of numerical data as metanumbers in tables with units, accuracy, and metadata descriptors linked as described above. Next consider how each number in such a design has a unique reference name (such as [server path_dir_table_row_column_...]). Likewise there is a unique name for every past computation of each user (such as [my_ref#]), as well as a link for each shared computational path by a project team under a given subject name (or [subject_user_ref#]. We suggest that the resulting system supports vast and powerful automated networks which can be constructed as described in 7.3 and 7.4 below.

7.2 Mathematics Underlying Our Network and Cluster Discoveries

This section will provide a very brief overview of our mathematical results that provide a new foundation for network and cluster analysis. In previous work the author discovered a new method of decomposing the continuous general linear (Lie) group of (n x n) transformations into a Markov type Lie group (with n^2-n parameters) and an Abelian scaling group (with n parameters). Each is generated as is standard, by exponentiation of the associated (Markov or scaling) Lie algebra. The Markov type generating Lie algebra consists of linear combinations of the basis matrices that have a "1" in each of the (n^2-n) off-diagonal positions with a "-1" in the corresponding diagonal for that column. When exponentiated, the resulting matrix M(a) = exp(a C) conserve the sum of elements of a vector upon which it acts, but can take a vector with positive components into one with some negative components (which is not allowed for a true Markov matrix). However, if one restricts the linear combinations to only non-negative values then we proved that one obtains all discrete and continuous Markov transformation of that size. This links all of Markov matrix theory to the theory of continuous Lie groups and provides a foundation for continuous Markov transformations.

Our next discovery was that every possible network (C_{ij}) corresponds to exactly one element of the Markov generating Lie algebra (those with non-negative combinations) and conversely, every such Markov Lie algebra generator corresponds to exactly one network! Thus they are isomorphic and one can now study all networks

by studying the associated Markov transformations and Lie algebra and associated groups. Our subsequent recent discoveries are (a) all nodes in any network can be ordered by the second order Renyi entropy of the associated column in that Markov matrix thus representing the network by a Renyi entropy spectral curve and removing the combinatorial problems so that now one can both compare two networks (by comparing their entropy spectral curve) as well as study the change of a networks topology over time. One can even define the "distance" between two networks (as the exponentiated negative of the sum of squares of difference of the Renyi entropies). We recently showed that the entire network topology of any network can be represented by the sequence of the necessary number of Renyi entropy orders. This is similar to the Fourier expansion of a function such as for a sound wave where each order represents successively less important information. Our next and equally important result was that an agnostic (assumption free) identification of the n network clusters can be found from the eigenvectors for this Markov matrix which not only show the clustering of the nodes but actually rank the clusters using the associated eigenvalues thus giving a name (the eigenvalue) for each cluster.

Our final recent development that is important for the current issues is that one can generate two different networks from a table of values T_{ij} for entities (such as the chemical elements in rows) with properties (such as density, boiling point, …) for each element in columns. To do this we first normalize the table by finding the mean and standard deviation of each column and then rewrite each column value as the number of standard deviations (represented by the table value), divided by the mean for that column, away from the norm. This process also removes any units that are present as the results are dimensionless. We then define a network C_{ij} among the n entities (here the elements listed in rows) as the exp of the negative of the sum of the squares of the differences between T_{ik} and T_{jk} thus

$C_{ij} = \exp\text{-}\Sigma_k(T_{ik} - T_{jk})^2$. This gives a maximum connectivity of '1' if the values are the same and a connection of '0' if they are far apart as would be required for the definition of a network. We form a similar network among the properties for that table. One then, as before, adjusts the diagonal to be the negative of the sum of all terms in that column to give a new C, forms the Markov matrix M(a) = exp (aC) and finds the eigenvectors and eigenvalues for M to reveal the associated clustering. The rationale for how this works can be understood when the Markov matrix is viewed as representing the dynamical flow of an invisible conserved substance among the nodes. This methodology includes and generalizes the known methodology with the Lagrange matrix for a network.

7.3 User-data Type Networks

Users (using the PIN#), can each be linked to (a) unit id (UID) hash values of the results of their calculations, as well as to (b) the table, row, and column names of each value. These linkages can be supplemented with linkages of each user to those universal constants that occur in the expressions which they evaluate. The resulting network links users to (a) concepts such as thermal conductivity, (b) substances such as silver alloys, and (c) core constants such as the triple point of methane, the Boltzmann constant, Planks constant, or the neutron mass. The expansion of this network in different powers, giving the different degrees of separation, can then link

users via their computational profiles, (the user i x user j component of C^2) as well as linkages among substances, metadata tags, and constants. The clustering revealed in different levels of such expansions then reveals groups of users with linkages that are connected by common computational concepts. Users working on particular domains of pharmaceuticals and methodologies are thus identified as clusters as well as groups of astrophysicists that are utilizing certain data and models. At that same time, the clustering can identify links among specific substances, models, and methodologies. Our current research is exploring such networks and clusters as the underlying metanumber usage expands.

7.4 Table Generated Networks among Entities and also among Properties

The methodology for converting a table to a network among row items or among column items was briefly discussed above. We are currently exploring the clusters and network analysis that can be generated from the tables of standardized metanumber values. We have done this for the elements table which was very revealing as it displayed many of the known similarities among elements. The cluster of iron, cobolt and nickel was very clear as well as other standard clusters of elements. We are currently analyzing a table on the properties of pesticides and we have just obtained a table of 56 nutrients for 8600 natural and processed foods. The study of clustering in foods based upon their nutrient and chemical properties as well as the clustering among the properties themselves will be reported as available on the www.metanumber.com web site. As usage of the metanumber system expands, we will also report the results of clustering among similar scientific investigations. These results are expected by the end of August 2015.

7.5 Multiuser Collaboration Application

Often with more complex computations and multi-tier problems, it is useful for a number of engineers, scientists, students, or business workers to collaborate on a problem. This requires that comments be shared as explanatory of the work that is in progress as well as questions and responses to others in the collaboration. This is even true for a user who wishes to document his or her own work. The metanumber application supports documenting text which is archived as an entry just like an expression that is to be evaluated. All that is necessary is to enter a "#" as the first character (similar to the method of entering a remark or comment line in python or other languages). When the "#" is seen, then all processing is bypassed and all text up to the end of that entry (cr lf) is archived as a string. Then when a user's past history is exported or read or shared, then these text entries are seen correctly positioned among the computations giving explanations or recording ones technique. The collaboration part is achieved by entering a command {subject = 'some name'|PW} where 'some name is the name of the subject to be shred. A file is created with that name (in lower case with white space removed and with the attached password. The creator of the subject must then share the name and password (if any) with users in a closed group. The subject stays active until the end of the session or until one enters

{subject = }. While the subject is active, each user with such an active subject has each line that is entered to be written to that new (subject name) file in the form: UserPIN, Seq#. This is all that is needed in the shared file because each user can then open that file to read only all of the entries from other in the group. More information will be available from www.metanumber.com help screens.

8 Conclusions

The use of these algorithms with the approximately 800 units and fundamental constants, can support a standardization of all numerical data. This metanumber environment also presents all data in human and machine readable formats and satisfies the listed set of 10 essential requirements. The resulting system provides a vast saving of time, costs, and a removal of associated errors by supporting the instantaneous use of data without preprocessing. This system can support new levels of artificial intelligence and improved human interaction. It can also support new methods for the creation of novel networks [9] in numerical data that in turn can support new methods of cluster analysis [10].

References

1. Johnson, Joseph E., 2009, Dimensional Analysis, Primary Constants, Numerical Uncertainty and MetaData Software", American Physical Society AAPT Meeting, USC Columbia SC
2. Leibigot, Eric O., 2014, A Python Package for Calculations with Uncertainties, http://pythonhosted.org/uncertainties/.
3. Johnson, Joseph E., 2013, The Problem with Numbers – MetaNumbers – A Proposed Standard for Integrating Units, Uncertainty, and MetaData with Numerical Values, USC Physics Colloquium
4. Johnson, Joseph E., 2014, New Software Developed for Interdisciplinary Data Sharing and Computation: Units, Uncertainty, & Metadata, USC School of the Earth, Ocean, and Environment Colloquium
5. Johnson, Joseph E. 1985 US Registered Copyrights TXu 149-239, TXu 160-303, & TXu 180-520
6. Johnson, Joseph E, Ponci, F. 2008 Bittor Approach to the Representation and Propagation of Uncertainty in Measurement, AMUEM 2008 International Workshop on Advanced Methods for Uncertainty Estimation in Measurement, Sardagna, Trento Italy
7. Johnson, Joseph E., 2006 Apparatus and Method for Handling Logical and Numerical Uncertainty Utilizing Novel Underlying Precepts US Patent 6,996,552 B2.
8. Johnson, Joseph E. 1985, Markov-Type Lie Groups in GL(n,R) Journal of Mathematical Physics. 26 (2) 252-257
9. Johnson, Joseph E. 2005 Networks, Markov Lie Monoids, and Generalized Entropy, Computer Networks Security, Third International Workshop on Mathematical Methods, Models, and Architectures for Computer Network Security, St. Petersburg, Russia, Proceedings, 129-135US
10. Johnson, Joseph E. 2012 Methods and Systems for Determining Entropy Metrics for Networks US Patent 8271412
11. Johnson, Joseph E, & Cambell, William 2014, A Mathematical Foundation for Networks with Cluster Identification, KDIR Conference Rome Italy

12. Johnson, Joseph E. 2014, A Numeric Data-Metadata Standard Joining Units, Numerical Uncertainty, and Full Metadata to Numerical Values, EOS KDIR Conference Rome Italy

Improving Automatic Test Case Generation Process with Knowledge Engineering in the Crystal Project

Stefano Marrone[1], Roberto Nardone[2], Antonio Maria Rinaldi[2] and Valeria Vittorini[2]

[1]Dipartimento di Matematica e Fisica, Seconda Università di Napoli, Caserta, Italy
stefano.marrone@unina2.it
[2]Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione,
Università di Napoli Federico II, DIETI, Napoli, Italy
{roberto.nardone, antoniomaria.rinaldi,
valeria.vittorini}@unina.it

Abstract. Recent research advances have brought to a growing interest from both academic and industrial communities in the improvement of existing engineering processes by means of model-driven techniques. This method is starting to demonstrate its effectiveness by raising the level of abstraction and by improving the level of automation of traditional processes. One of these applications is related to V&V processes and in particular to the generation of system level test cases for critical systems. This chapter investigates the possibility to further improve such process by exploiting synergies between model-driven techniques and knowledge engineering ones. This work is developed in the context of Crystal, an EU Artemis funded research project, and focuses on a specific part of its framework. The proposed approaches are demonstrated by means of a case study in the field of railway signalling system.

1 Introduction

Ontologies constitute formal models of some aspect of the world that may be used for drawing interesting logical conclusions even for large models. Software models capture relevant characteristics of a software artefact to be developed. Most often these software models have no formal semantics, or the underlying (often graphical) software language varies from case to case in a way that makes it hard if not impossible to fix its semantics. In this context, ontology-based metamodels constitute a core means for exploiting expressive ontology reasoning in the software modelling domain while remaining flexible enough to accommodate varying needs of software modellers [1]. With this aim, the four-layer modelling architecture provides the basis for formally defining software modelling languages and some open challenges can be recognised: semantics of modelling languages often is not defined explicitly but hidden in modelling tools; to fix a specific formal semantics for metamodels, it should be defined precisely in the metamodel specification; the syntactic correctness of models is often analysed implicitly using procedural checks of the modelling tools; to make well-formedness constraints more explicit, they should be defined precisely in the metamodel specification. Ontologies and the related languages to represent them can be used to improve the expressive power of software metamodels.

In the last two decades, ontological aspects of information have acquired a strategic value. These aspects are intrinsically independent from information codification, so the information itself may be isolated, recovered, organised and integrated with respect to its content [2]. A formal definition of ontology is proposed in [3], according to whom *"an ontology is an explicit and formal specification of a shared conceptualisation*; *conceptualisation* is referred to as an abstract model of specified reality in which the component concepts are identified; *explicit* means that the type of concepts used and the constraints on them are well defined; *formal* refers to the ontology property of being "machine-readable"; *shared* is about the property of an ontology of capturing the consensual knowledge, accepted to a group of person, not only to a single one.

A basic step in the knowledge engineering process is the use of "tools" to represent knowledge, both for inferring and organising it. From this point of view, one of the most important advances in the knowledge representation (KR) applications is derived from proposing [4], studying [5–7] and developing [8–10] languages based on the specification of objects (concepts) and the relationships among them. The main features of all KR languages are the following:

(i) *object-orientedness*, for which all the information about a specific concept is stored in the concept itself (in contrast, for example, to rule-based systems;
(ii) *generalization/specialisation* are basic aspects of the human cognition process [4], the KR languages have mechanisms to cluster concepts into hierarchies where higher-level concepts represent more general attributes than the lower-level ones, which inherit the general concept attributes but are more specific, presenting additional features of their own;
(iii) *reasoning* is the capability to infer the existence of information not explicitly declared by the existence of a given statement;
(iv) *classification* in which given an abstract description of a concept, there are mechanisms to determine whether a concept can have this description; this feature is a special form of reasoning.

Object orientation and generalization/specialisation help human users in understanding the represented knowledge; reasoning and classification guide an automatic system in building a knowledge representation, as the system knows what it is going to represent.

Moreover, we argue that when a KR formalism is constrained in such a way that its intended models are made explicit, it can be classified as belonging to the ontological level [11] introduced in the distinctions proposed in [7], where KR languages are classified according to the kinds of primitives offered to the user.

In recent years, several languages have been proposed to represent ontologies. These languages have a different expressive power and, starting from some considerations from previous authors' works [12, 13], it is our opinion that OWL [14] is the best language for the purpose of the proposed approach.

OWL 2, the web ontology language, is a W3C recommendation with a very comprehensive set of constructs for concept definitions and allow for specifying formal models of domains. Generally speaking, ontologies are conceptual models, that can be described by OWL. Based on its underlying formal semantics and different services could be provided. They vary between satisfiability checking at the model layer, checking the

consistency of instances with regard to the model, or classifying instances (finding their possible types) with regard to instance and type descriptions. Since ontology languages are described by metamodels and allow for describing structural and behavioural models, they provide the capability to combine them with software modelling languages. In our framework we want use ontologies to support the definition of software modelling language semantics and provide the definition of syntactic constraints.

A second concern that is at the base of this Chapter is a growing need coming from industrial settings to introduce advanced modelling approaches into existing development processes. This need is raised not only by industries more prone to innovation in ICT but also by the manufacture industries where the development processes are well assessed and where managers are less prone to change them. Model-Driven Engineering (MDE) [15] is starting to be applied in this contexts; notwithstanding the absence of universally accepted standard processes, it is still one of the most promising techniques to improve productivity. By means of model-driven techniques both requirements elicitation and analysis, design space exploration and verification & validation phases of a product/service life-cycle can be improved: this can be accomplished by processes built upon the two pillars of the MDE:

(i) *metamodelling*, which allows the structure of a domain in terms of abstract and concrete domain models with textual/graphical languages as well as by extending existing languages (e.g., UML profiling mechanism);
(ii) *model transformation*, by means of which it is possible to automatically generate artifacts that can code (as in Model Driven Software Development) as well as further models (e.g., Petri Nets, Bayesian Networks, etc.).

From all the industrial sectors and the development phases that may be improved in this way, this work is focused on verification processes in railway signalling systems. In particular we focus on automated testing processes (at system level) as a way to improve the quality/safety of the product and to reduce costs and time. Several research works proposes improvements of this specific topic by means of model-driven approaches (a great part of this work are framed into the Model-Driven Testing - MDT).

Due to the high number of common aspects between KR and MDE approaches, the objective of this work is to explore the synergies between these two worlds on the specific problem of automated testing process. A mixed KR-MDE approach of the entire system testing process is defined with enabling techniques as well as present issues. This ongoing work is framed into the ARTEMIS Joint Undertaking project CRYSTAL (CRitical sYSTem engineering AcceLeration) [16] that will be further described in the next sections.

The Chapter is structured as follows: Section 2 presents the CRYSTAL project while Section 3 focuses on the specific Use Case of this work. Section 4 describes the overall approach of KR-MDE integration in the improvement of automated testing process and Section 5 details an aspect of the entire process. Section 6 draws conclusions and future developments.

2 An overview of the Crystal Project

CRYSTAL takes up the challenge to establish and push forward an Interoperability Specification (IOS) and a Reference Technology Platform (RTP) as a European standard for safety-critical systems. CRYSTAL is strongly industry-oriented and will provide ready-to-use integrated *tool chains* having a mature technology-readiness-level. Figure 1 depicts this overview.

To achieve technical innovations ("technology *bricks*"), CRYSTAL adopts a user-driven approach based on applying engineering methods to industrially relevant *Use Cases* from the automotive, aerospace, rail and health-care sectors [17] and *increases the maturity of existing concepts* developed in previous European and national projects like CESAR [18], iFEST [19], MBAT [20]. Moreover several product life-cycle or project management phases/concerns are used to group together similar research task in the Crystal project. They are: analysis tools, safety tools, AUTOSAR tools, heterogeneous simulations, product life-cycle management, multi viewpoint engineering, variability management, SW development life-cycle, validation models and simulation models.

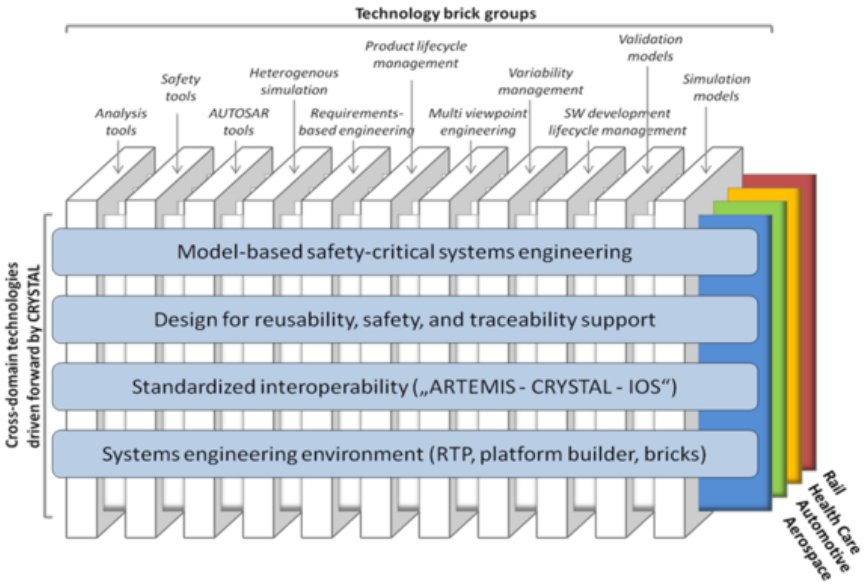


Fig. 1. The Crystal Overview (http://www.crystal-artemis.eu).

Four cross domain technologies cut the entire space of domains and of development life-cycle phases which Crystal embraces: model-based safety critical system engineering, design for reusability and traceability support, standardised interoperability and system engineering environments. As it is clear, model-driven engineering and knowledge engineering are first class citizens in the vision of the Crystal project and hence, finding ways in where these two pillars of the software engineering can express their synergies, is a research task of a great value.

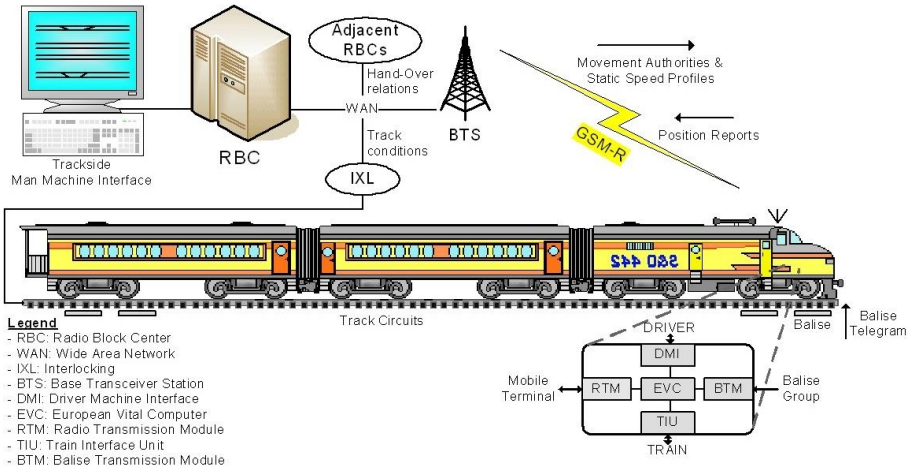


Fig. 2. ERTMS/ETCS Level 2.

The achievement of a good level of interoperability cannot leave aside the definition of specific domain ontologies in such big cross-domain project. In fact, in Crystal, one ontology for each domain are the output of specific work packages. The advantages of these activities can be found in the definition of a common vocabulary in the specific domain, in the simplification of communication thorough the different operators and in the usage of a common glossary in the deliverables and artefacts of the project. Furthermore the application of ontology activities is at the basis of the definition of the IOS which can take advantages from these domain ontologies operating in a cross-domain manner.

3 The RBC Use Case within the Crystal Project

The focus of this work is in the rail domain, and specifically from the needs expressed by Ansaldo STS (ASTS), an international transportation leader in the field of signalling and integrated transport systems for passenger traffic (Railway/Mass Transit) and freight operation. The industrial needs expressed by the ASTS's *Use Case* are oriented to improve the quality and the efficiency of existing Verification & Validation (V&V) processes, with a specific focus on the validation of functional requirement with testing. In fact, testing activities are time-consuming tasks whose efficiency is a primary issue in a global competitive market and whose quality can not be decreased due to the adherence to international standards.

3.1 The RBC Use Case

The ASTS's Use Case is centred on the Radio Block Centre (RBC) system, a computer-based system whose aim is to control the movements of the set of trains on the track area under its supervision, in order to guarantee a safe inter-train distance according to the ERTMS/ETCS specifications. ERTMS/ETCS (European Rail Traffic Management System/European Train Control System) [21] is a standard for the interoperability

of the European railway signalling systems ensuring both technological compatibility among trans-European railway networks and integration of the new signalling system with the existing national train interlocking systems. Each ERTMS/ETCS controlled track is usually divided into several sub-tracks, each of them is supervised by a single RBC in charge of *concurrently and continuously* controlling a number of connections with trains. The main objective of the train control system is to timely transmit to each train its up-to-date Movement Authority (MA) and the related speed profile. The MA contains information about the distance the train may safely cover, depending on the status of the forward track. RBC is also in charge of managing emergency situations if the communication with one or more trains is compromised. Figure 2 gives an overview of the ERTMS/ETCS lev 2 at a glance.

With a particular focus on the validation of the system against functional requirements, a great effort is spent on the generation, execution and analysis of system-level functional test cases. Since these systems are classified as the most dependable in terms of Safety Integrity Level (i.e., they are classified as SIL 4) and according to the applicable international standards and norms (i.e., CENELEC EN50128 [22] and CENELEC EN50126 [23]), these activities must be conducted by a proper "V&V team" which shall be independent from the development team. This team must rely only on high-level behavioural description of the system and on the set of system requirements that the system have to satisfy; its objective, at system level, is the definition of test cases able to functionally validate the overall system against its requirements.

An improvement of the actual V&V approach is hence required for these systems, allowing the automatic execution of some activities. For these reasons our goal in the CRYSTAL project is represented by the definition of a new methodology which must be able to support the execution of these activities: on the basis of a system model is used to drive the process by means of automatic tool. The main activities that have been traditionally done manually are now supported by tools even if the interaction with a *V&V Engineer* is present.

Figure 3 shows, by means of a diagram mixing UML Use Case elements and an architectural schema, the interactions between user and system as well as the tool supporting such functionalities. In the diagram, tick solid lines represent automatic flows, solid thin lines activities that are executed outside this automated process while dotted lines related use cases with automatic tools. A similar approach and supporting architecture has been defined in [24] but, in that paper, the approach is oriented in mixing functional and non-functional properties. With respect to another previous work [25] this description is enriched with more details and it constitutes an improvement.

The flow of activities can be described as follows. The V&V Engineer is in charge to **Model RBC** functions in one **System Model** that is conformant to the Dynamical StaTe Machine (**DSTM**) language. This language, considering both the needs of a strong formal foundation and ease of use of the final user, it is defined according to principles of MDE as a Domain Specific Modelling Language (DSML). Further discussion on DSTM is in Subsection 3.2. Essentially, DSTM is an extension of state machines where the behaviour of the system is represented by states and transitions. Furthermore, the model is annotated with functional requirements (**Model Functional Requirements**): up to date, requirements are mapped onto transitions.

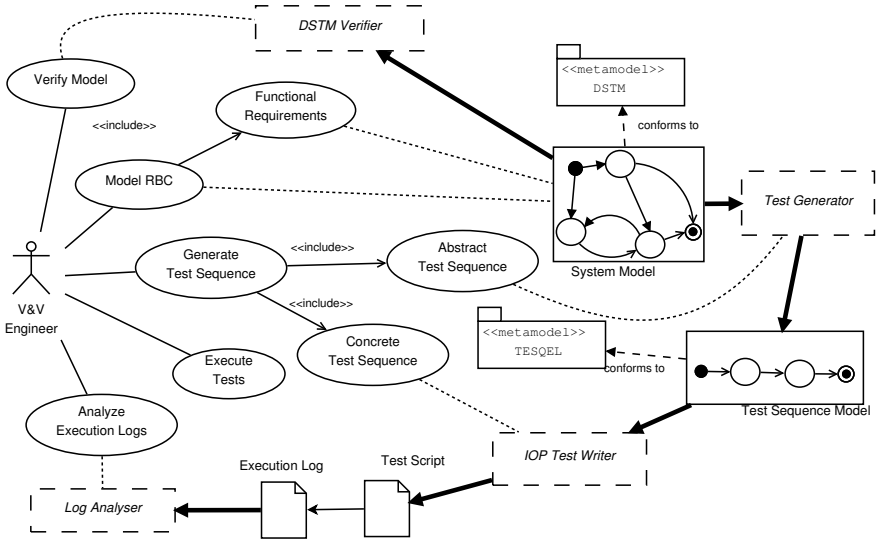


Fig. 3. The RBC Use Case Automated Testing Process.

After the model is created, it should be verified in order to check if it conforms to all the constraint of the language (**Verify Model**): this action is supported by the **DSTM Verifier** tool. Up to date, this tool is essentially a compiler which takes different parts of a DSTM model and verify the consistency of the model itself and its compliance to all the constraints defined in the DSTM language. Different techniques may be used to specify the model: while structural elements of the model itself are better created trough a graphical concrete syntax, for variables and data-types, the best way still is a textual old-style concrete syntax. Both traditional parsing techniques and advanced model-driven manipulation and querying approaches are used.

Then, test-sequences can be automatically generated, with a minimum effort required to the V&V team (**Generate Test Sequences**): this activity can be parted into a phase where "abstract" sequence are generated (**Abstract Test Sequence**) and one where abstract test sequences are realised in a concrete scripting language and able to be executed (**Concrete Test Sequence**).

The generation of abstract test sequences supported by the **Test Generator** tool that works as follows: a test specifications is actually derived from the requirements and it contains the features that a test sequence to generate must own (see for a fully description of this item [26, 27]). At the state, two test specifications are generated for each requirement: a finite set of 'positive' test specifications (i.e., the situations in which the transition must be performed), and one 'negative' test specification (i.e., the situation in which the transition have not to be performed). Starting from these hypotheses, the model and each test specification generate a **Test Sequence Model** which represents one of the many concrete executions on the System Model which fulfils the test specification. This last artefact is conformant to the TEst SeQuEnce Language (TESQEL) which is also built according to model driven principles. At the state, the Test Generator is implemented by exploiting model checking techniques [28]: a DSTM model is hence

translated into a Promela language while the negation of the test specification becomes a CTL property to check. The counterexample is the sequence of execution steps on the model which negates the property (i.e. which satisfies the test specification). This notwithstanding, future developments can consider different approaches for the Test Generator mechanism.

Generated test sequences must be executable and hence TESQEL conformant sequences are translated into an executable language by the **IOP Test Writer** which aim is to translate the "model" of the test sequence into an interoperable language for the execution of ERTMS/ETCS tests (the IOP language itself) (**Test Script**).

Once these scripts are executed, outside of this approach, **Execution Log**s are produced: these logs are analysed (**Analyze Test Logs**) in order to understand if some anomalies are present. This phase is supported by the **Log Analyser**.

It is important to underline that the IOP Test Writer and the Log Analyzer are not in charge of the research units of Seconda Universitá di Napoli and of Universitá di Napoli Federico II. This notwithstanding in this chapter we discuss also on these tools about the possibility to improve them. Such improvements could be done outside the context of the CRYSTAL project.

3.2 The DSTM Language

DSTM extends Hierarchical State Machines [29] specifying an original semantics of fork-and-join. This makes DSTM more powerful than the UML State Machines [30] since it adds, between others, mechanisms for dynamic instantiation and recursive execution of machines. An excerpt of the DSTM metamodel is shown in Fig. 4, where the Ecore diagram is depicted. This Ecore diagram represents the realisation of DSTM, which formalisation have been introduced in [31], in the Eclipse Modeling Framework [32]. The main class is Dynamic State Machine *(DSTM)*, which represents the entire specification model. A DSTM is composed of different *Machine*s, *Channel*s and *Variable*s and allows for the definition of own-defined *Type*s. *Channel*s and *Variable*s allow for communication between machines and with the external environment. A single *Machine* is composed of *Vertex*es, *Transition*s and may have a set of *Parameter*s.

The class *Vertex* is abstract since different kinds of vertexes (with different features and constraints) may be present in a machine. The vertex kinds are similar to those contained in the UML State Machine, but with a different semantics for the *Fork* and *Join* concepts. A fork splits an incoming transition into more outgoing transitions; it allows for instantiating one or more processes either synchronously or asynchronously with the currently executing process. The asynchronous instantiation represents the instantiation of machines without suspending the current executing process, which is enabled to continue its evolution. On the contrary, a join merges outgoing transitions from concurrently executing processes: it synchronises their termination together with the current executing process, if asynchronous instantiation have been performed, and/or allows to force the termination when a process is able to perform a *preemptive* exiting transition. The classes *Fork*, *Join* and *EnteringNode* are inherited from the abstract class *PseudoNode* which encompasses different types of transient vertexes in the machine.

The class *Transition* is specified by many attributes. It can specify its *trigger*, its activation *condition* and a set of *actions*. These attributes are specified by a string that must

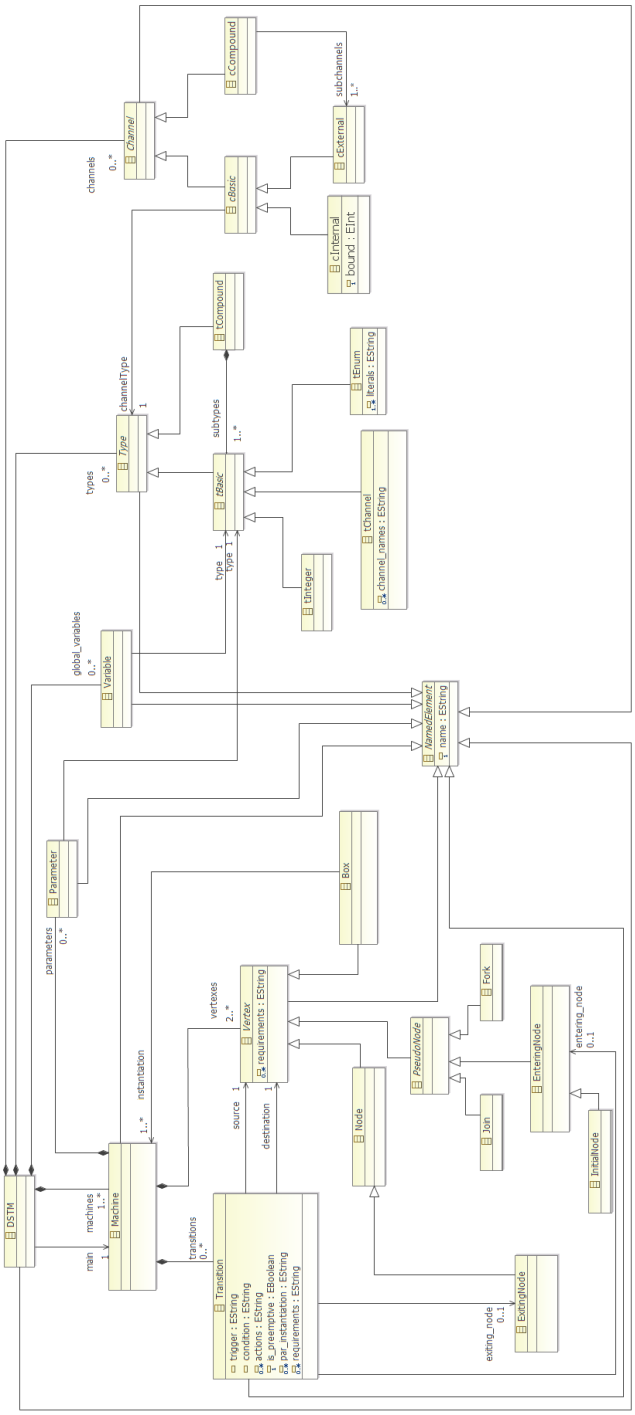


Fig. 4. DSTM4Rail metamodel [31].

comply with a given syntax. Furthermore a transition can be preemptive, i.e. enabled to kill concurrent executing processes, by setting to true the value of the *is_preemptive* attribute. If a transition enters a box, it can specify the set of parameter instantiations by the attribute *par_instantiation*.

Types allowed in a DSTM model are either *tBasic*s and *tCompound*s: the formers represent integer and enumeration types while the latters represent data structures composed by basic subtypes. A specific type, *tChannel*, has been added in order to represent the namespace of channels. Note that *Variable*s and *Parameter*s are associated with *tBasic* since, in this version of the language, only basic types can be specified for both variables and parameters. The set of allowed channels is dived into *cInternal*, *cExternal* and *cCompound*. Each channel has an associated type, either a simple type or a compound. Internal channels allows for internal communication and allows for the specification of a message buffer and are instantaneously updated when a writing action is performed; external channels instead are used for the communication with the external environment and machines are not allowed to remove messages from these channels. Compound channels are also defined in order to group external channels, specifying the set of channels which model the communication with a single external entity, hence the set of channels which can contain at most one message (if one of the grouped channel contains a message, the others must be empty).

The semantics of DSTM is provided by means of a Labeled Transition System containing sequences of a maximal set of transitions. Specifically the messages generated over external channels cannot trigger other transitions in the same step; in addition a node/box cannot be entered and exited simultaneously in the same step. Accordingly to this semantics sequential firings of transitions are not allowed within a step, only transition affecting concurrent processes can be performed within the same step. Furthermore external channels, if empty, can be filled with non-deterministically generated messages (compliant with the specific type allowed on the channels).

The main peculiarities of this language reside in the high expressive power which is also semantically well-defined. In fact, according to the needs expressed in [33], its abstract syntax is given by a metamodel and the semantics is entirely formally defined; in this way multiple developers understand exactly what modelled. Another advantage is that, according to the adopted technology, DSTM can be easily implemented by graphical diagrams, coping with the necessity of usage.

4 Merging Knowledge and Model-Driven Engineering Methods

In this section we introduce our vision on the integration of ontologies and MDE [1]. Generally speaking, we discuss the role of descriptive and structural models, in particular ontologies, in the model-driven process. First, the different role of domain and upper-level ontologies is discussed. In this context an upper-level ontologies can also be used as language descriptions. Second, we integrate parts of the CIM as ontologies into the MDA meta-pyramid (ontology-aware meta-pyramid). In fact, this delivers a first ontology-aware mega-model of MDE [34], and we use its conceptual advantages. On the one hand, the mega-model suggests an extended, ontology-aware software process.

On the other hand, the technologies for tool construction in the MDA and MOF world can be transferred to the ontology world.

The basic idea of the ontology-aware meta-pyramid is that most models in MDE are specifications, but can integrate ontologies on different meta-levels as descriptive analysis models. Since ontologies differ from specifications due to their descriptive nature, the standard M0-M3 meta-pyramid can be refined from using pure specification models to also using ontologies. Depending on the meta-level, an ontology may serve different purposes. In fact, there are different qualities of ontologies in the literature. First of all, the word ontology stems from philosophy, where it characterises Existence. Ontology is a systematic account of Existence [3]. We call such a systematic account of existence a World ontology, a conceptualisation of the world, that is, all existing concepts. Usually, a World ontology is split into an upper-level ontology (concept ontology, frame ontology), providing basic concepts for classification and description, and several lower-level ontologies, domain ontologies describing domains of the world [35, 36].

Usually, concepts of the domain ontology inherit from concepts in the upper-level ontology. For better interoperability and understanding, some researchers try to create a normalised upper-level ontology, from which all possible domain ontologies may inherit [37]. If a standardised upper-level ontology with modelling concepts existed, all domain ontologies could rely on a standardised concept vocabulary.

With this terminological distinction, we can relate the different forms of ontologies to meta-levels in the meta-pyramid. Domain ontologies live on level M1, they correspond to models. An upper-level ontology, also a standardised one, should live on level M2, because it provides a language for ontologies.

We describe two general approaches [33] to bridge software languages and ontology used in the framework of our unit in the Crystal project. In the language bridge approach, the design of an M3 integration bridge consists mainly of identifying concepts in the Ecore metametamodel and the OWL metamodel which are combined. The integration bridge itself is used at the M2 layer by a language designer. He is now able to define language metamodels with integrated OWL annotations to restrict the use of concepts he modelled and to extend the expressiveness of the language. The M3 Transformation Bridge allows language designers and language users to achieve representations of software languages (Metamodel/Model) in OWL. It provides the transformation of software language constructs like classes and properties into corresponding OWL constructs. A model transformation takes the UML metamodel and the annotations as input and generates an OWL ontology where the concepts, enumerations, properties and data types (TBox) correspond to classes, enumerations, attributes/references and data types in the UML metamodel. Another transformation takes the UML model and generates individuals in the same OWL ontology. The whole process is completely transparent for UML users.

Using this mapping, we can transform an Ecore Metamodel/Model into OWL TBOX /ABOX.

In the model bridge approach, software models and ontologies are connected on the modelling layer M1. They are defined in the metamodelling layer M2 between different metamodels. The bridge is defined between a process metamodel on the software modelling side and an OWL metamodel in the OWL modelling hierarchy. The process meta-

Table 1. An example of Ecore and OWL comparable constructs.

Ecore	**OWL**
package	ontology
class	class
instance and literals	individual and literals
reference, attribute	object property, data property
data types	data types
enumeration	enumeration
multiplicity	cardinality

model is an instance of an Ecore (EMOF) metametamodel. A model bridge is defined as follows: (1) Constructs in the software modelling and in the ontology space are identified. These constructs, or language constructs, are used to define the corresponding models in the modelling layer M1. (2) Based on the identification of the constructs, the relationship between the constructs are analyzed and specified. M2 Integration Bridge merges information of the models from the software modelling and from the ontology space. This allows the building of integrated models (on modelling layer M1) using constructs of both modelling languages in a combined way, e.g. to integrate UML class diagrams and OWL. A transformation bridge describes a (physical) transformation between models in layer M1. The models are kept separately in both modelling spaces. The information is moved from one model to the model in the other modelling space according to the transformation bridge. A process model like a UML Activity Diagram is transformed to an OWL ontology. The transformation rules or patterns are defined by the bridge. Thus, having a process model as an ontology we can provide services for reasoning on the semantics of process models. Ontology Reasoning for Behaviour Modelling Languages The model bridge is defined in the metamodelling layer M2 and is used in layer M1 to transform or integrate model entities on layer M1. Process models capture the dynamic behaviour of an application or system. The metamodels of both are instances of Ecore meta-metamodels. The two metamodels provide flexible means for describing process models for various applications. However, due to their flexibility further modelling constraints and semantic descriptions are required for a clearer representation of the intended meaning. There are additional modelling characteristics for process models in the software modelling space which are analysed in detail in literature. (1) A semantic representation of control flow dependencies of activities within a process, i.e. execution ordering of activities in a control flow. Such constraints allow for the description of order dependencies e.g., an activity requires a certain activity as a predecessor or successor. (2) It is quite common in model-driven engineering to specialise or refine a model into a more fine-grained representation that is closer to the concrete implementation. In process modelling, activities could be replaced by sub-activities for a more precise description of a process.

Using these approaches it is possible a representation of behaviour models in OWL and applications of reasoning services in order to provide model management services for example on process models represented by UML Activity Diagrams. In order to use ontology reasoning for process models, a first step is to build a model bridge from process models (software models) in a UML-like representation to an ontology (TBox).

The model bridge is defined in the metamodelling layer M2 and is used in layer M1 to transform or integrate model entities on layer M1. We consider a transformation bridge. We present our process model bridge that defines a transformation from process models given as UML activity diagrams to on OWL ontology (TBox). This requires a thorough consideration of the entities that are represented in process models, their relations like control flow relations and how they are transformed to OWL ontologies. A challenge in this task is to capture the semantics of process models like activity ordering and flow conditions in the ontology.

A task of our unit in the Crystal project is the development of a model bridge between DSTM (M2 language) and an ontology based model to represent in OWL the features of this language. Using a mapping the DSTM notations will Be translate in description logic and by means of a knowledge base some reasoning services will be implemented. In particular, we'll take into account the Automatic annotation of system model and the Log Analyzer bricks developed in the Crystal framework.

5 Improving Test Generation with KE

This Section shows the areas that have been detected in Crystal and more specifically in the RBC Use Case, as possible integration points between MDE and KE. Two of these areas are explored in details and improved version of the MDE-based processes are proposed: achieving these goals would constitute the prime objective of future research efforts.

5.1 Overall of the KE-improved Process

Figure 5 starts from the block level model of the tool-chain proposed within the RBC case study. Furthermore, this schema depicts the point of this tool-chain where Knowledge Engineering techniques can be applied and where synergies with Model Driven Engineering must be searched.

Six main intervention areas are detected:

- *Intelligent Model Verification* (IMV) deals with the problem of adding some advanced features to DSTM Verifier. The proposal here is to improve such level of verification by adding some intelligent features in this phase which can not only verify the model but also suggest to the final user possible improvements. More details on this phase are reported in Subsection 5.2;
- *Requirement Annotation* (RA) means the possibility to automatically propose a mapping between the requirements and a DSTM model. The phase is in charge of analysing the requirements (traditionally expressed in a natural language) and to search the submodels of a DSTM model that best fit to represent these requirement. More details of this phase are reported in Subsection 5.3;
- *Automatic Model Construction* (AMC) is intended to support the modeller into the automatic creation of a DSTM model. This support is constituted by suggest some hints to the modeller: such suggestions may vary from simple expression completion to the suggestion of complex model patterns;

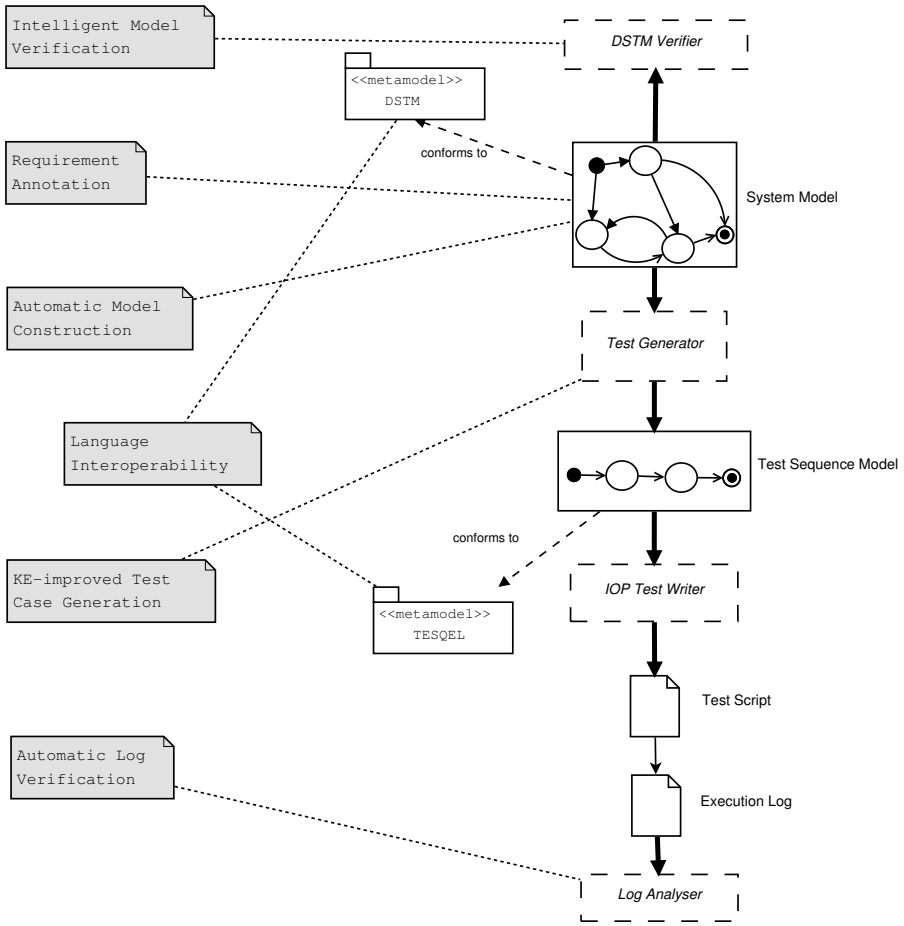


Fig. 5. Points of improvements of the RBC Use Case Automated Testing Process.

- *KE-improved Test Case Generation* (KTCG) aims to improve the test sequence generation phase by defining some assertion which are invariants with respect to the model dynamics. Such assertions may be used in order to restrict the state space where the model checker searches the desired test sequence: these assertions are also called reduction rules and can be inferred from reasoning activities on the DSTM mode by means of automatic reasoning techniques;
- *Automatic Log Verification* (ALV) improves the existing Log Analyzer by adding machine learning techniques in order to understand from the log produced by the execution of the Test Script if the requirement to verify is fulfilled by the trace;
- *Language Interoperability* (LI) can be used to extend the range of influence of the DSTM language to other Crystal's life-cycle phases and/or applicative domain. Since Crystal is a project that strongly promotes the interoperability among different domains, the use of such techniques to apply DSTM and the related tool-chain may be used in order to automatically map concepts in first different among them.

The application of this approach could also be extended to TESQEL as a way to verify if this language may re-used in other contexts.

Some of these application areas can create mutual benefits when synergies are searched: as example, Automated Model Construction may benefit from patterns and anti-patterns defined during the IMV phase while reduction rules of KE-improved Test Case Generation should also be inferred by log analysis.

5.2 Improving Model Verification

Figure 6 shows the schema that will be studied and realised in the next research work: the schema is in charge of defining the main blocks of the IMV functionality. As described, the aim of such functionality is not only to check if the model is correct according to defined syntax and semantics but also to provide a proper support in improving the modelling experience by suggesting best and/or worst practises.

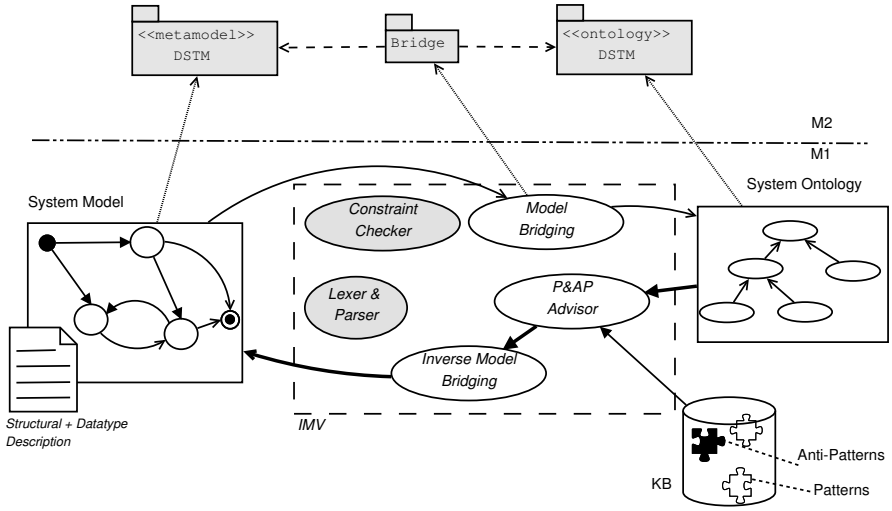


Fig. 6. An integrated reference schema for the IMV approach.

The process works as follows: the DSTM system model (both graphical description of the model structure as well as the textual definition of datatypes and variables) is first processed by traditional techniques. As graphical model structure verification phase can exploit modern model-driven technologies able to generate a model from a graphical user interface that is already conform to a metamodel, the textual definition of the datatypes must by processed by traditional parsers and lexers. After this phase, a validation of semantic constraints are due in order to ensure that the model is well-formed. Some examples of these constraints are: (i) a variable should be defined and assigned to a type; (ii) a variable used in the DSTM model (e.g., in the definition of a trigger) must be declared in the datatype file; (iii) a DSTM transition coming out from a pseudo node must not have a trigger expressed (see for further details [31]).

These activities traditionally retrieve to the user some exceptions in case the model is not con formant to the syntax/semantics of the language. Furthermore, many conformant models (i.e., raising no exception) may be improved by adopting common modelling practises (patterns) or avoiding common pitfalls (anti-patterns). This activity is performed by the Pattern & Anti-Pattern Advisor (**P&AP Advisor**) that is in charge of reasoning on a ontology representing the DSTM model (**System Ontology**, i.e. ABOX) and using available reasoning techniques and technologies. This reasoning activity is in charge of substituting non-efficient sub-parts (the anti-patterns) of this ontology with other more efficient ones (the patterns): both patterns and anti-patterns are contained into a Knowledge Base (**KB**).

The means by which the DSTM **System Model** is translated into the System Ontology are a **Model Bridging** and an **Inverse Model Bridging** transformations. Such bridging functions are built implemented upon the definition of a proper bridge between DSTM metamodel and ontology as described in Section 4.

5.3 Improving Requirement-Model Mapping

A picture of a second integration way is reported in Figure 7 where Requirement Annotation (RA) proposed process is depicted.

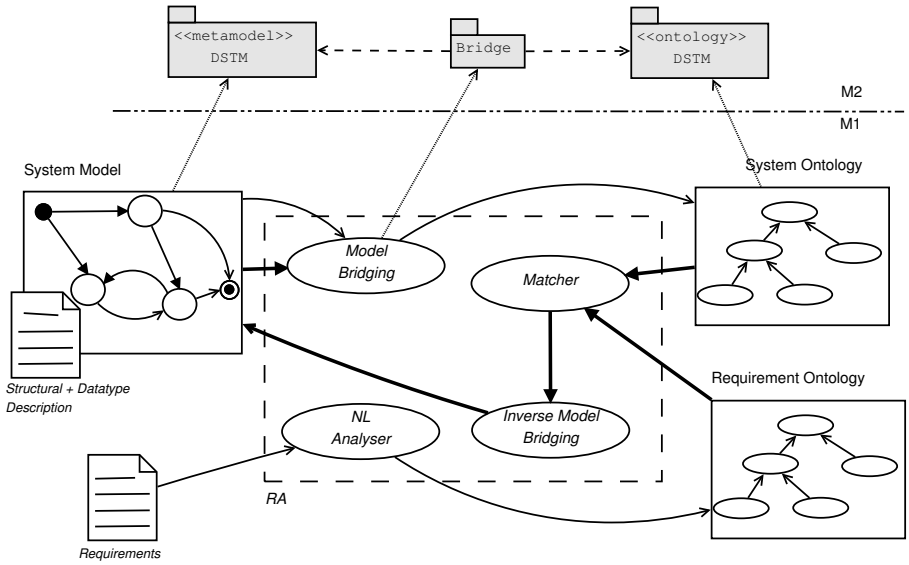


Fig. 7. An inter grated reference schema for the RA approach.

This process has the aim to aid the modeller in annotating the defined DSTM System Model with requirements. Requirements are considered in DSTM for both traceability purpose and to generate automatically Test Sequences on the base of the item of the DSTM that is annotated with the requirement. In other words, when a requirement is mapped onto a model transition, the Test Sequence related to that requirement consider

the passage trough the model transition. Hence, annotating the model in an effective and efficient way is an important task that Knowledge Engineering can improve.

The main idea is to consider a reasoner (the **Matcher**) able to match a **System Ontology** and a **Requirement Ontology**. While the first can be obtained from the System Model by exploiting the bridging technique already defined, the latter can rely on a well-assessed research background on Natural Language Processing (**NL Analyzer**). Once the match is done, an Inverse Bridging reports the annotations on the DSTM System Model.

6 Future Developements

The ontology-aware meta-pyramid offers several other benefits that can all be summarised by the exploitation of the transformational techniques of the MDE and the reasoning techniques of the KE. This chapter has defined a roadmap in the concrete realization of a synergies of these two worlds in the context of an industrial-driven research project: the Crystal project. By selecting one of the many Use Case of the Crystal project, this chapter illustrates the main points where KE and MDE may find their synergy.

Of course, this work describes an on-going research mainly by illustrating the main next activity that involve both the research units of the University of Naples Federico II and the Second University of Naples.

Acknowledgements. This paper is partially supported by research project CRYSTAL (Critical System Engineering Acceleration), funded from the ARTEMIS Joint Undertaking under grant agreement n. 332830 and from ARTEMIS member states Austria, Belgium, Czech Republic, France, Germany, Italy, Netherlands, Spain, Sweden, United Kingdom.

References

1. Aßmann, U., Zschaler, S., Wagner, G.: Ontologies, meta-models, and the model-driven paradigm. In: Ontologies for Software Engineering and Software Technology. Springer (2006) 249–273
2. Rinaldi, A.: A content-based approach for document representation and retrieval. In: Proceedings of the eighth ACM symposium on Document engineering, ACM (2008) 106–109
3. Thomas, R.: A translation approach to portable ontology specifications. Knowl. Acquis. 5 (1993) 199–220
4. Minsky, M.: A framework for representing knowledge. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA (1974)
5. Woods, W.: What's in a link: Foundations for semantic networks. In Bobrow, D.G., Collins, A., eds.: Representation and Understanding. Academic Press, New York (1975) 35–82
6. Brachman, R.: What's in a concept: Structural foundations for semantic nets. International Journal of Man-Machine Studies 9 (1977) 127–152
7. Brachman, R.: On the epistemological status of semantic networks. In Findler, N.V., ed.: Associative Networks: Representation and Use of Knowledge by Computers. Academic Press, Orlando (1979) 3–50

8. Brachman, R., Schmolze, J.: An overview of the Kl-ONE knowledge representation system. Cognitive Science 9 (1985) 171–216
9. Fox, M., Wright, J., Adam, D.: Experiences with srl: an analysis of frame-based knowledge representations. In: Proceedings from the first international workshop on Expert database systems, Redwood City, CA, USA, Benjamin-Cummings Publishing Co., Inc. (1986) 161–172
10. Bobrow, D., Winograd, T.: An overview of krl, a knowledge representation language. Technical report, Stanford University, Stanford, CA, USA (1976)
11. Guarino, N.: The ontological level. In Casati, R., B. Smith, B., White, G., eds.: Philosophy and the Cognitive Sciences. Holder-Pichler-Tempsky, Vienna (1994)
12. Albanese, M., Maresca, P., Picariello, A., Rinaldi, A.: Towards a multimedia ontology system: an approach using TAO_XML. In: In Proceedings of the 11th International Conference on Distributed Multimedia Systems (DMS'05). (2005) 52–57
13. Cataldo, A., Rinaldi, A.: An ontological approach to represent knowledge in territorial planning science. Computers, Environment and Urban Systems 34 (2010) 117–132
14. Dean, M., Schreiber, G.: OWL Web Ontology Language Reference. Technical Report http://www.w3.org/TR/2004/REC-owl-ref-20040210/, W3C (2004)
15. Schmidt, D.C.: Guest editor's introduction: Model-driven engineering. Computer 39 (2006) 25–31
16. (CRYSTAL: CRitical sYSTem engineering AcceLeration) http://www.crystal-artemis.eu/.
17. Pflügl, H., El-Salloum, C., Kundner, I.: CRYSTAL, CRitical sYSTem engineering AcceLeration, a Truly European Dimension. ARTEMIS Magazine 14 (2013) 12–15
18. (CESAR: Cost-Efficient methods and proceses for SAfety Relevant embedded systems) http://www.cesarproject.eu/.
19. (iFEST: industrial Framework for Embedded Systems Tools) http://www.artemis-ifest.eu/.
20. (MBAT: Combined Model-based Analysis and Testing of Embedded Systems) http://www.mbat-artemis.eu/.
21. UIC: ERTMS/ETCS class1 system requirements specification, ref. SUBSET-026, issue 2.2.2 (2002)
22. CENELEC: Cenelec, en 50128: Railway applications - communication, signalling and processing systems - software for railway control and protection systems (2011)
23. CENELEC: Cenelec, en 50126: Railway applications - demonstration of reliability, availability, maintainability and safety (rams) - part 1: Generic rams process (2012)
24. Marrone, S., Flammini, F., Mazzocca, N., Nardone, R., Vittorini, V.: Towards model-driven v&v assessment of railway control systems. International Journal on Software Tools for Technology Transfer 16 (2014) 669–683
25. Barberio, G., Di Martino, B., Mazzocca, N., Velardi, L., Amato, A., De Guglielmo, R., Gentile, U., Marrone, S., Nardone, R., Peron, A., Vittorini, V.: An interoperable testing environment for ertms/etcs control systems. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8696 LNCS (2014) 147–156
26. Amalfitano, D., Amatucci, N., Fasolino, A., Gentile, U., Mele, G., Nardone, R., Vittorini, V., Marrone, S.: Improving code coverage in android apps testing by exploiting patterns and automatic test case generation. (2014) 29–34
27. Gentile, U., Marrone, S., Mele, G., Nardone, R., Peron, A.: Test specification patterns for automatic generation of test sequences. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8718 LNCS (2014) 170–184
28. Gargantini, A., Heitmeyer, C.: Using model checking to generate tests from requirements specifications. SIGSOFT Softw. Eng. Notes 24 (1999) 146–162

29. Alur, R., Kannan, S., Yannakakis, M.: Communicating hierarchical state machines. In: Automata, Languages and Programming. Volume 1644 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (1999) 169–178
30. OMG: (Unified modeling language (uml), superstructure)
31. Gentile, U., Nardone, R., Peron, A., Benerecetti, M., Vittorini, V., Marrone, S., De Guglielmo, R., Mazzocca, N., Velardi, L.: Dynamic state machines for formalizing railway control system specifications. (2015)
32. Steinberg, D., Budinsky, F., Paternostro, M., Merks, E.: EMF: Eclipse Modeling Framework 2.0. 2nd edn. Addison-Wesley Professional (2009)
33. Staab, S., Walter, T., Gröner, G., Parreiras, F.: Model driven engineering with ontology technologies. In: Reasoning Web. Semantic Technologies for Software Engineering. Springer (2010) 62–98
34. Favre, J.M.: Megamodeling and etymology-a story of words: From med to mde via model in five milleniums. In: In Dagstuhl Seminar on Transformation Techniques in Software Engineering, number 05161 in DROPS 04101. IFBI, Citeseer (2005)
35. Guizzardi, G., Herre, H., Wagner, G.: On the general ontological foundations of conceptual modeling. In: Conceptual ModelingER 2002. Springer (2003) 65–78
36. Sowa, J.: Knowledge representation: logical, philosophical, and computational foundations. (1999)
37. Niles, I., Pease, A.: Towards a standard upper ontology. In: Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001, ACM (2001) 2–9

Advanced Cloud Document System

Alberto Buschettu[1], Filippo Eros Pani[2] and Daniele Sanna[1]

[1] Experteam srl, Spin-off of University of Cagliari, Via Zara 11, Cagliari, Italy
{alberto.buschettu, daniele.sanna}@e-xperteam.com
[2] Department of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi, Cagliari, Italy
filippo.pani@diee.unica.it

Abstract. Nowadays, public administration offices are faced with long and complex procedures. Frequently used tools such as email clients and document systems are seldom integrated. This leads to problems like slowness, large quantities of paper documents, redundancies, high operational costs, and a low level of citizen participation. Google Apps try to fill that gap by providing an integrated and usable environment, but it is not enough for typical public administration applications, where the existing IT system has to be taken into account. The aim of this project is to create a Software as a Service (SaaS) platform, based on open-source components and integrated with Google Apps. With such a platform, public institutions and companies could use an environment in which to work on the most frequent tasks. From an innovative point of view, the project uses the latest technologies and the latest patterns in both planning and development.

1 Introduction

The large work volume and the limited time in which public institutions need to provide their services represent a difficult issue to solve. The computerization of services means the reduction of paper documents with time, leading to a substantial step forward. The tasks that an organization has to perform are supported by the use of suitable software, such as email clients, document systems, work-flow managers, online publishing systems, etc. Users access these tools daily, but there are apparent limitations that hinder the improvement of the efficiency of a whole organization (either public or private). This is due mainly to the lack of integration, to the usage of non-open formats in interchanged data and to the impossibility of remote access to the work tools.

The proposed project aims to create, even as a prototype, a typical Cloud software system that can integrate already existing systems with the ecosystem of Google Apps[1]. That would help the resources employed in daily work activity, improving the general efficiency by leveraging on the potential offered by Google Apps, such as their remarkable capability for document circulation among users in environments

[1] Google Apps, https://www.google.com/work/apps/business/

that use email clients and document managers. The advantages are outstanding, including the improvement of the interaction between public offices and other institutions as well as of its communication with citizens, creating an actual Virtual Organization.

The software system at the core of the project is categorized as SaaS, and represents the latest technology in software use and supply. The spirit of the project is to pursue the creation of software artifacts by using typical approaches of Open Source. A project of this kind, making use of innovation in software development through the use of non-traditional models, also needs a new approach to development. Namely, the project includes the use of an Agile Methodology derived from Kanban [1], given the experience in that context gained by the proponents, in particular by the Agile Group[2] of the Department of Electrical and Electronic Engineering (DIEE) at the University of Cagliari.

2 Context of Research Proposal

The project contributes to a substantial advancement to the state of the art in the field. It follows the logic of the process management issues in Service Oriented Architectures (SOA), with an integration to the management of the documents involved in those processes and the management of the data coming from the process together with the documents themselves. At a scientific level, this kind of industry research follows the approach proposed on Service Oriented Computing, Kanban methodologies [2], and evaluation methodologies for software solutions [3]. It focuses, however, on their usage in workflow processes that rely on a Cloud-based Web Service architecture, and pays specific attention to interesting themes such as the integration of a document logic and the management of end-to-end security. The problem of security in applications and data in a SOA will be addressed by taking into account solutions like WS-security. Another aspect that will be under investigation is the study of a sys-tem allowing the optimal management of all applications involved in these processes and elaborated in this SOA. That interaction could be described and formalized also through Business Process Management (BPM) solutions, which would allow for an efficient management and structuring of the information pertaining to analyzed processes and the compliance of the tool to prominent market standards like Business Process Modeling Notation (BPMN) and Business Process Execution Language (BPEL).

It is known that a large part of the data owned by organizations can be found non-structured documents (editable or non-editable text, pictures, videos, etc.), while the current systems, including integrated "data warehousing" systems, are based on structured data belonging to one or more database. A document-centered approach could guarantee the usage of the whole wealth of document, adding in-text search systems, semantic marking, etc., and represents a hot topic in research.

The system adopts an approach that has become a standard in many situations, that is the use of web-based software as a service (SaaS) in Cloud mode. Google can

[2] Agile Group, DIEE, University of Cagliari, http://agile.diee.unica.it/

be said to be the pioneer of SaaS of this kind, creating services like office automation, document management, calendar, web-mail, etc., currently widespread in public and private organizations.

In Cloud systems, the computing paradigm is no more valid, and certain approaches, namely object-oriented design, do not find a place. In fact, in this field it has been replaced by "service-oriented" design.

Another successful approach is the Model Driven Architecture (MDA) [4], based on platform-independent models, which is especially suited in literature for an Agile approach based on Lean-Kanban [5]. The use of this model has remarkable implications in terms of interoperability, a crucial aspect in environments such as True SaaS. It is in the nature of an application that works on the Cloud to possess a high interoperability, which needs to be planned starting with its design when an MDA approach is at its natural place.

The literature has defined some best practices for the creation of a SaaS, whereas research and experience made it possible to find some fundamental points. First of all, software customization tends to be replaced by configuration. In fact, SaaS feature highly configurable, multi-tenant applications. This characteristic is very well finished, so much so that Automatic Tuning (AT) [6] and Dynamic Adaptive Systems are emerging. They allow an automated creation of the optimal configuration for each user (tenant) or group of users. Its modular nature, its templating functionalities and its Role Access Management (which allows to assign specific access policies to users, together with function and resource use policies) are important features for software subject to configuration. The functionalities of SaaS are distributed on different service levels, where each service level to be assigned to a functionality takes into account the use of cloud resources, the business value of the functionality, and of a possible purchase method. Currently, the SaaS that provide an integrated environment where email, document manager, office automation suites and other tools such as certified email, electronic signature, fax, etc, can be used are few, and are usually released as desktop applications. Integration is therefore a key element: according to Forrester Research [7], the main reason why organizations do not adopt SaaS lies in problems with the integration of in-house solutions with the SaaS, especially for real-time application communication.

Another important aspect of the proposed research concerns the comparison with other tools in existence, namely Cloud and non-Cloud. NVAF and its extensions [3] are valid frameworks to make a comparison without discrimination, using metrics taken from scientific literature.

3 Description of Research Project

The purpose of the project is to create an integration solution for typical Cloud tools such as those provided by Google, a solution that would use services and Apps like Google Drive, Google Documents, Google Maps, Gmail, Google Calendar, Google+, etc., to offer the necessary software equipment to circulate documents in an integrated environment from/to other organizations, in a perspective of Virtual Organization. With that Cloud approach, users can provide services without the use of desktop

software, but directly through the Cloud, speeding up their tasks and making data and documents immediately available. This can be applied to typical use cases in the interactions between a citizen and a public office, or between two public offices. It would mean to provide a solution that integrates tools with which the user is familiar, with other tools that add legal standing to the communication (in particular, an integration between tools like certified email, electronic signature, Italian National Service Cards (CNS), fax, IT protocol, but with a more familiar interface for a digital native citizen). It is an area where traditional issues related to the introduction of new technologies merge with new issues related to legal risks in data storage and potential reduction of the costs on competing proprietary software, in a context of continuing variation and advancement of information technology.

The project covers the creation of software artifacts through approaches that are typical of Open Source, and using the tools provided by that paradigm in order to apply the state of the art to the design and implementation of the system.

To create such a project, with its particular problems related to software development on the Cloud, something that differs from traditional models, that is an Agile Methodology derived from Kanban, will be experimented, in light of the experience gained by the proposers in this field.

3.1 Project Subdivision

The project includes a number of operational stages, which involve:

- the analysis and evaluation of the state of the art in the industry relevant to the project (available technologies and services, competitors, etc.);
- the verification of the options in the definition of the software development tools in this type of application-service;
- the evaluation of industry risk and the selection of development options;
- an in-depth study of the selected technology;
- the definition of the basic services to be developed;
- the definition of the integration architecture;
- lastly, the development and validation of a demonstration prototype.

3.1.1 Evaluation of State of the Art

The initial activity entails the study of a broad range of needs related to innovative Cloud applications, in order to perform an appropriate selection and focus on the developments of higher impact. One of the possible options for technology is the use of Google App Engine, the hosting and development platform that allows to create high-traffic web applications without the need to handle their issues. This environment provides a number of development tools: Google App Engine SDK for Java, Google App Engine SDK for Python, and Google Plugin for Eclipse.

A second option involves developing the software in J2EE architecture using the API provided by Google to make its integration possible. Many software houses with a presence on Google Apps Marketplace offer their applications and services on their own portal, in addition to the one provided by Google. Reference technologies will be

picked for an early selection of the architectures for the support system to the development of the services. The following operations will be carried out:

- the evaluation of available technologies and the comparison with the use of a support framework;
- the verification of active services and competitive systems.

3.1.2 Verification of Software Development Tools and Study of Technologies

This activity is based on an innovative Lean approach to software development in a Cloud environment, also in distributed and collaborative environments. Cloud practices will be also used, according to the idea: "developing software for the Cloud using the opportunities offered by the Cloud itself", to determine some relevant modules in a standalone environment. Another subject of investigation is the risk analysis of the use of Cloud solutions of this kind. In fact, one of the factors that acted as a deterrent to the adoption of Cloud solutions was, until recently, the risk stemming from storing company data in third-party remote systems, and the risk of uncontrollable downtime. But the Cloud does not only offer the advantage of online storage: it also allows to create new mechanisms of communication and collaboration.

A study was conducted on Google's environment to analyze its first-level characteristics, such as its data exchange format, its exposed functions, and its authentication method.

3.1.3 Definition of Applications and Services

The issues related to the Cloud, and the effectiveness of the solutions currently available in the market, will be studied, also considering the developments expected in the medium-long term. This phase of the research project will have a scientific character, but also an empirical one, since it requires the knowledge of market orientation and of potential competitors to the current developments as well as of future ones. The activity is thus made of the analysis of the reference market and of the services offered by competitors, the definition of priority requirements of the services, and lastly an evaluation of the evolution expected in the medium term.

3.1.4 Definition of Solution Architecture

This activity aims to study valid solutions and design a technology infrastructure that overcomes the basic issues together with the application issues. The Cloud Computing paradigm needs to be taken into account, from both a technological point of view of cooperation between applications, and from the logic of cooperation between applications and data, which could come from different systems and entities. The studied infrastructure will have to be able to manage document flows supporting an effective engine to manage applications, other than a presentation through Google Apps, in a SOA context.

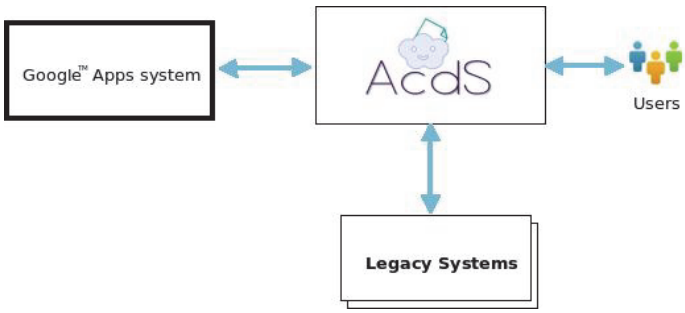


Fig. 1. General structure of the solution architecture.

3.1.5 Prototype Development

In this phase, the software components of the systems will be designed, created and validated, in prototype form. The development will require specific modules that will satisfy the need for modularity, interoperability, flexibility through the use of open standards. Particular care will be given to the building of a prototype that would solve the complex issues raising from the interaction of distinct, yet interdependent, processes in applications.

Another aspect that will be tackled at a prototypical level is the management of the security integrated in SOA supporting the Cloud, in a dynamic form, especially concerning access authorization. An architectural pattern of API Gateway[3] type will be also studied. Such a component provides an intermediate layer that allows the platform components to perform requests to the Google Apps API through calls that have a preset name, independent from Google. The internal routing system, with its rules, allows the transparent use of the correct Google API.

4 Project Schedule

The project officially began on March 3, 2014, and its conclusion is estimated for the end of December 2015.

The current phase is the implementation of the prototype covered by the project. Special attention was given to the study of Document Management and interoperability between systems, as well as of integration techniques. That is a crucial aspect, since the adoption of complex software means having to face existing software systems with which dialogue is needed. Interesting results came to light in the comparison among software systems, which was necessary to evaluate the potential of competitors and of the Open Source software adopted in the project. SaaS is actually a new domain, which requires new approaches and new metrics compared to traditional software.

[3] API gateway, http://microservices.io/patterns/apigateway.html

4.1 Interoperability Issues

The topic of interoperability is fundamental in environments such as True SaaS [8][9]. An application that runs on the Cloud must necessarily have a high interoperability. This feature needs to be planned from the design stage, with and MDA approach and the use of open technologies, as it was found during our study. It is a neutral and open approach for the development of enterprise applications, where modeling leads to the development of the software. MDA encourages the evolution of solutions through subsequent transformations, from high-level models to low-level models, down to the generation point of the code. Due to the considerable impact coming from the sudden changes related to the alternation of new technologies, the creation of Cloud services with a specific technology raises some issues. In fact, even when the same Cloud Service Provider is confirmed, it could update its own technologies. Therefore, MDA makes it possible to develop technology-independent Cloud services where the business logic does not depend on technical details.

4.2 Google Apps Integration

An analysis of competitors was performed, especially of those operating in the ecosystem of Google Apps. The analysis highlighted the great interest many companies have for applications completely based on the integration with Google Apps. The apps that resulted to be most interesting were Google Drive and Gmail. A large part of the systems created by competitors seek the highest integration with those applications, exploiting their potential. One kind of widespread application in terms of Google Apps usage is a Workflow Management (using BPMN) that uses Google Drive as document storage and email as an alert system on the various tasks that make up a workflow. This meant a further confirmation of the validity of the project, and very good prospects for the company that hosts it. The same analysis sheds light on the strong effort from competitors to port the main functions onto a mobile environment.

Real case studies were evaluated in order to set apart the functionalities that would have the biggest impact on a public institution. The research made it possible to review the best practices and the patterns on development in a Cloud environment, giving the opportunity to be based on the state of the art in both design and development. The development of this type of applications requires some features that may not exist in traditional desktop applications [10], such as multitenants, that is the presence of all the users who share the same instance of the application.

Processes represent a fundamental aspect for an organization, and many companies do not have any process management systems. We therefore considered that as a priority functionality to implement.

A SaaS application can improve existing processes, or create new ones. The possibility for users to operate on the move could open new scenarios, which would be impossible were its availability limited to the company offices only.

A process manager is one of the software systems that most require a strong interoperability between systems. Consider, for example, the recurring need processes have for email, document resources, accountancy data, etc.

4.3 Architecture Choices

During the first phases of infrastructure design, our studies led us to define the optimal patterns for both the supply of the SaaS and the functionalities of the application. In particular, an API Gateway-type architecture pattern was used to isolate the application from possible future changes made by Google on Google Apps. Such a component provides an intermediate layer, which allows the platform components to perform requests to the Google Apps API through calls that have a preset name, independent from Google. The internal routing system, with its rules, allows the transparent use of the correct Google API.

The following figure roughly shows the way in which software clients can use an API server with an indirect communication thanks to an API Gateway.

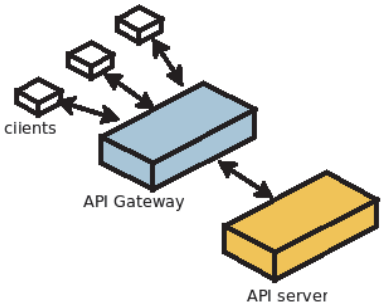


Fig. 2. Communication with API Gateway.

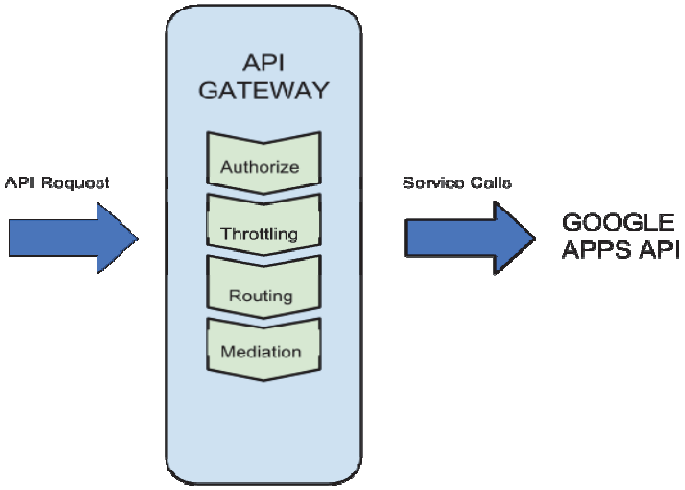


Fig. 3. API Gateway layers.

The API Gateway validates and authenticates requests to Google, transforms data before they are sent (logic validation), behaves as a mediator and performs caching. It makes it possible to manage:

– who is making the request (via OAuth);

- when it requests (Error-rate, payload-size, traffic shaping);
- where requests are managed (in-line transformation, orchestration);
- how to request (flow logic, transformation and validation logic, caching logic).

4.3.1 Authentication/Authorization

Two important actions performed by the API Gateway are authentication and authorization. The gateway has the account data of each user of the platform (tenant). Every request of use of Google Apps passes through an appropriate authentication module of the API Gateway, which performs all the necessary operations on Google's authentication system to authorize a tenant to use the extensions through Google Apps.

4.3.2 Caching

Caching controls the requested resource and the client requesting it. According to that information, in compliance with an appropriate Expiration Model and Validation Model, it is possible to return a local version of data or perform an interrogation towards Google Apps. That local version is called cache entry and can be automatically created by the gateway, or its creation can be inhibited with a no-cache flag related to the call arriving from the client.

We opted to store the cache gateway side, that is, storing it in the gateway. This decision was made from a performance perspective, which would be lacking when using an external module to manage the cache, for example a key managed server.

4.3.3 Mediator

This part of the application executes business operations on the resource request called by a client, executing the manipulations needed to prepare the final request to send to Google Apps. The mediator module includes a listener of the path of the request incoming from the client. If a handler, i.e. an API definition, corresponds to that path, all the operations that prepare data before the next Router/Proxy phase are performed. The operations include a manipulation of the body data of the request coming from the client.

4.3.4 Router/Proxy

This component includes a mapping between functionalities and the URL of Google API, to redirect and launch the final call that will reach the endpoints of Google's API. There is a register inside it, where each entry has the parametric URL and a number of features used to locate the entry according to the request pre-pared by the Mediator. From this module, the request is launched through a Request Rewriting operation, and the response is returned. The response will be appropriately recomposed by the mediator to be eventually returned to the client. A single request from a client could comprise more synchronous requests to Google Apps, the URL of which are included in the router/proxy module.

5 Conclusion

The project represents an advancement of the state of the art in a domain still not completely explored, namely the development of True SaaS integrated with third-party systems like Google Apps. This kind of applications will become increasingly common, and will change the way to work with software tools. The proposed solution aims to become a reference tool for the daily tasks to be performed by users of either public organizations or private companies. This solution also offers basic tools (email managers, office automation, document manager, etc.), using a suite of proven effectiveness such as Google Apps. The innovation that the completion of the project can bring is remarkable, and articulated on more facets. On the one hand, it is the answer to a current need with innovative technologies (SaaS delivered system). On the other hand, innovation is brought into the solution itself. The use of open implementation technologies and of standards of the same kind addresses classic integration and sync issues with legacy systems, and facilitates data migration (document or other) towards the proposed solution. The solution will guarantee a high security, and lower down-time levels compared to other systems in existence. In fact, through synchronous replication, users' data and actions will be copied in real time on a number of data centers, with a switching function between data centers.

Acknowledgements. This paper has been produced as part of the research project entitled ''Servizi Avanzati in Cloud (Advanced Cloud Services)'' developed at Experteam s.r.l. company. The project is financially supported through a research grant funded by the Autonomous Region of Sardinia with European local funds (ROP SARDINIA ESF 2007-2013 – Objective: Regional Competition and Employment, Axis IV Human Resources, Line of Activity l.1.1. e l.3.1.).

References

1. Anderson, D. J.: Kanban: Successful Evolutionary Change for Your Technology Business. Blue Hole Press (2010)
2. Corona, E., Pani, F. E.: A Review of Lean-Kanban Approaches in the Software Development, Transactions on Information Science and Applications, Vol.10, No.1, Print ISSN: 1790-0832, E-ISSN: 2224-3402 (2013)
3. Pani, F. E., Sanna, D., Marchesi, M., Concas, G.: Transferring FAME, a methodology for assessing open source solutions, from university to SMEs. In: D'Atri, M. De Marco, A. M. Braccini, F. Cabiddu, (Eds.), Management of the Interconnected World, ItAIS - The Italian Association for Information Systems, A 1st Edition, XIV, 534 p., Hardcover, LNBIP, ISBN: 978-3-7908-2403-2 (2010)
4. Sharma, R., Sood, M., Sharma, D.: Modeling Cloud SaaS with SOA and MDA, Advances in Computing and Communications - Communications in Computer and Information Science (2011)
5. Poppendieck, M., Poppendieck, T.: Lean software development: An agile toolkit, Addison Wesley, Boston, Massachusetts, USA (2003)
6. Naono, K., Teranishi, K., Cavazos, J., Suda, R.: Software Automatic Tuning, From Concepts to State-of-the-Art Results, Chap. 15, pp. 255-274. Springer, Berlin (2010)
7. Forrester, The Forrester Wave: Hybrid Integration, Q1 (2014)

8. Lewis, G. A.: The Role of Standards in Cloud-Computing Interoperability, Software Engineering Institute, Carnegie Mellon (2012)
9. Almeida, F., Oliveira, J., Cruz, J.: Open standards and open source: enabling interoperability, International Journal of Software Engineering & Applications (IJSEA), Vol.2, No.1, January (2011)
10. Critical Requirements for Cloud Applications: How to Recognize Cloud Providers and Applications that Deliver Real Value - Workday (2015)

Knowledge Management Theory Creation in Healthcare Environment

Erja Mustonen-Ollila[1], Jukka Heikkonen[2], Antti Valpas[3] and Helvi Nyerwanire[1]

[1] Department of Innovation & Software, Lappeenranta University of Technology, Finland
{erja.mustonen-ollila, helvi.nyerwanire}@lut.fi
[2] European Commission, Joint Research Centre, Unit JRC.G1,
Scientific Support to Financial Analysis, Italy
jukka.heikkonen@jrc.ec.europa.eu
[3] Department of Obstetrics and Gynaecology,
South Karelia Social and Health Care District, Finland
antti.valpas@eksote.fi

Abstract. In this research project, the research environment is the central hospital of the South Karelia Social and Healthcare District, Finland. This study is qualitative research project in which the data is collected by semi-conducted interviews in order to create a knowledge management theory in a healthcare environment. The theory consists of conceptual frameworks and their categories and the relationships between the categories. The categories and their relationships are discovered by using the Grounded Theory approach. When discovering a new theory the approach needs both qualitative and quantitative methods. Therefore, this research project will utilize a new methodological approach where both qualitative and quantitative research approaches are applied. The quantitative data will be analyzed with the novel intelligent computing methods which are used to find out relationships of the data and to give deeper understanding of the research domain and its conceptual dependencies.

1 Introduction

The primary goal of this research project is to create a knowledge management theory in healthcare environment based on empirical findings. Health and healthcare are important components for physical and mental well-being and life satisfaction. Only part of the healthcare is medicine where the goal is to retrieve the lost homelike being-in-the-world experience of the patient e.g., by reducing pain and suffering via preventing and curing diseases. Healthcare also consists of nursing and rehabilitation which are lacking from the medicine [1]. In healthcare, the patients are aware of their own health including mental, physical, and social dimensions [2], and transparency, integrated care platform, consumer engagement, patient centeredness, evidence-based medicine, and doctor-patient relationship are needed to be implemented [3], [4], [2], [5].

Healthcare is a knowledge rich sector which experiences rapid growth in order to understand the different diseases and their treatments combined with the ability to apply and access the up-to-date and relevant information. Knowledge itself has many

conceptualizations depending on disciplines [6]. For instance, accountants measure it on the balance sheet; information technologists want to codify it on systems; sociologists want to balance power with it; psychologists want to develop minds because of it; human resource managers calculate a return of investment on it; and training and development officers want to make sure that they can build it. It has been also claimed that knowledge can be divided into creation or construction process, transfer process, storage and retrieval process, and application process [7], [8], [1], [9]. These processes control organizational knowledge in the means of knowledge dissemination, organizational learning, and data management [8], [10].

Among healthcare practitioners, knowledge is captured in a social interaction, e.g. when physicians and nurses meet patients. Physicians transfer their knowledge and expertise in meetings and consultations sessions, and they can express and interpret diagnosis reports, create new expert knowledge by reading or learning in traineeship, and having discussions [7], [8]. Collective knowledge exists in the organizational networks and knowledge is internally ingrained in people [8]. People learn new knowledge through participation with each other, e.g. cardiologists can belong to a community of practice transferring and receiving knowledge on best practices [11]. Knowledge can be made useful to support a specific task or make a decision [12], personalized information [8], awareness, experience, skills, and learning [13], [14], tacit knowledge [15], [16], explicit knowledge [17], and clinical and medical knowledge of the physicians and nurses [18], [19].

Different types of knowledge concepts have been identified and integrated into existing and emerging healthcare information management practices [8]. Knowledge concepts are even expanded to cover clinical medical expert knowledge which is bounded to physicians' medical knowledge and expertise both in practice and theory [18], [19], [20]. Furthermore, collective knowledge is expanded to cover organizational learning from internal and external sources of organizations and sub-networks [20] and embedded knowledge of the members, tools, and tasks of the organization [21]. Mostly healthcare organizations utilize combination of the different types of knowledge and healthcare practitioners are able to apply it properly and efficiently in their work and tasks [20], [21].

Healthcare sector is heavily dependent on knowledge and it needs efficient knowledge management in order to achieve high standards in patient care quality and patent safety and centeredness, but also in cost-effectiveness. Therefore, knowledge management needs information communication technology (ICT) systems support to be applied in knowledge creation, as well as in the capturing, organization, access, and use of knowledge [22], [8]. Based on the ICT systems support, knowledge management includes knowledge acquisition, creation, transfer, storage, and application processes [8], and organizational learning, unlearning, and internal learning processes [23], [24], [25]. It has been stated that evidence-based medicine is a form of organizational learning in the knowledge management context [20]. The British Medical Informatics Society states that the goal of medical informatics, also defined as the health informatics, is to share and promote healthcare through the information technology (IT) with proper skills, knowledge and tools [26]. In healthcare, there exist different types of healthcare information systems, such as electronic health records, electronic medical records, and electronic patient records [26], [8]. These records based information systems contain data about the patient diagnosis, drugs and electronic prescriptions of medical laboratory examinations. The

functionalities and goals of these three types of systems, however, are different from each other and dependent on the hospital or clinic they are used [27], [22], [28].

The physicians' and nurses' tacit, expertise, medical and nursing knowledge, however, cannot be found from the healthcare information systems. This is an unfortunate omission because without such knowledge the other healthcare professionals are not able to carry out the similar diagnoses and nursing decisions in a similar patient care situation. The important role of information and communication technologies (ICT) in healthcare is also ignored due to its problematic nature [26], [29], [30], [31]. Furthermore, studies of knowledge management theory creation in healthcare have been neglected, and past studies have rather focused on knowledge management theory creation in IT context [32], [33].

At our research site practice, external and tacit knowledge is captured and transferred by conducting lectures, and special training sessions to young physicians, and learning by experience. The current healthcare information systems are updated regularly and it causes stress for the physicians and nurses because they need to learn parallel the issues of new healthcare information systems and new medicine and nursing practices. The physicians in the central hospital even claim that the healthcare information systems do not support but on the contrary often disturb and complicate their clinical work. Therefore, there is immense need for information systems modernization, as well as restructuration of knowledge management processes. This is a challenging task because the nationwide medical and nursing practices have to be updated regularly and new guidelines should be easily retrieved from the information systems. In addition, the changes in Finnish and European Union (EU) legislation, national guidelines for management practices, and economic and political situation have to be found from the healthcare information systems [34].

Due to these several important omissions in the past studies, and on the other hand a simultaneous need to modernize the information systems in practice in the research site because of the new guidelines and legislation, this research project is important and carried out in South Karelia Social and Healthcare District, Finland. In this district's central hospital we choose a department as a unit of analysis in order to create a knowledge management theory in a healthcare environment. The theory also includes the healthcare information systems, information systems integrations', and cloud computing adoption's impact to knowledge management. This research project is very significant nationally, because the Finnish healthcare districts are under a great pressure of the growing healthcare costs and organizational changes due to need of specialized medical and nursing care. In the future, the hospitals and social and healthcare districts in Finland have to decide their main quality of service, because the level of service must be balanced with the money and resources available. In Finland, there are nationwide guidelines and operational processes of how to take care of a patient, and the goal of these processes is to provide the same patient care nationally. It should be noted that similar problems have been faced in very many other EU countries making this research also significant internationally.

The rest of this paper is organized as follows. In section two the objectives and research problems are outlined. In section three, we develop the conceptual framework of knowledge management. In section four, we outline research methodology consisting of data collection and categorization, and qualitative data analysis with the grounded theory and quantitative data analysis with the novel

intelligent computing approach. Finally, in section five we communicate the major research findings of the project.

2 Objectives and Research Problems

The goal of this research project is to create a theory in healthcare environment by using both the Grounded Theory (GT) qualitative research approach [35] and novel intelligent computing methods based on general framework called the Cross Industry Standard Process for Data Mining (CRISP-DM) [36]. This research project is planned to take a total of 9 years (2012-2020) in which the output will be five doctoral theses. Furthermore, at least 20 conference and 10 journal articles in high quality conferences and journals will be published.

Table 1. Research areas, research problems in healthcare, and related theories.

Research area	Research problems in healthcare	**Related theory**
Knowledge concepts	What are the knowledge concepts?	[8], [1], [12]
Internal and external knowledge acquisition mechanisms	How do the nurses and physicians acquire knowledge?	[45], [9], [46], [32], [20]
Knowledge use and application	What practical, clinical, medical, and nursing knowledge the physicians and nurses use and apply in patient care situation?	[37], [47], [8], [1], [48], [12], [45], [20], [7]
Knowledge creation	How the physicians and nurses construct knowledge?	[49], [37], [47], [8], [1], [48], [12], [45], [20], [7]
Knowledge transfer mechanisms and transfer problems	How knowledge is transferred? What are the knowledge transfer problems?	[37], [50], [9], [51], [52], [53], [9], [20], [54], [55]
Organizational learning mechanisms, unlearning and internal learning mechanisms	What are the organizational learning, unlearning and internal learning mechanisms?	[45], [37], [9], [32], [56], [23], [24], [25]
Role of information and communication technology	What is the role of information and communication technology?	[43]
Information systems' integration	How to integrate information systems together?	[26]
Information systems' integration approaches and mechanisms	What are the information systems' integration approaches and mechanisms?	[57]
Information systems' standards and technologies	What are the information systems' standards and technologies?	[58]
Cloud computing adoption	How to adopt cloud computing?	[44]
Healthcare services in cloud computing	What healthcare services are available when using cloud computing?	[29]
Data security and privacy issues in cloud computing	What data security and privacy issues must be taken into account in cloud computing?	[29], [59]
Cloud computing risks	What are the risks associated with cloud computing?	[29]

The first primary objective of this study is to discover and conceptualize

knowledge concepts and knowledge management processes in a hospital environment by combining together information systems science, knowledge management science, medical science, nursing science, sociology of knowledge, management science, and computational intelligence [15], [7], [37], [8], [9], [38], [39], [32], [40], [41], [13], [14], [2], [42]. The second primary objective is concerned to discover healthcare information systems' impact to knowledge management in healthcare [43]. The third and fourth primary objectives are to study cloud computing adoption's impact [44], [29], [31], and information systems' integration impact [30], [26] to knowledge management in healthcare. The research questions, research problems in healthcare, and related theories are presented in Table 1.

3 Conceptual Framework of Knowledge Management

The research site is the South Karelia Social and Healthcare District, and the unit of analysis is a department at district's central hospital. As a whole, the study covers currently at least six departments. As shown in Figure 1, knowledge management processes are knowledge acquisition (1), knowledge creation and construction (2), knowledge transfer (3), knowledge storage (4), and knowledge application (5). Their interactions with each other are represented as solid arrows. Internal learning, organizational learning, and organizational unlearning have an impact to knowledge creation and construction process. Their impact is represented as dashed arrows. Cloud computing adoption, information systems' integration, and healthcare information systems' impacts to knowledge management processes, department, central hospital and healthcare district are seen as solid arrows.

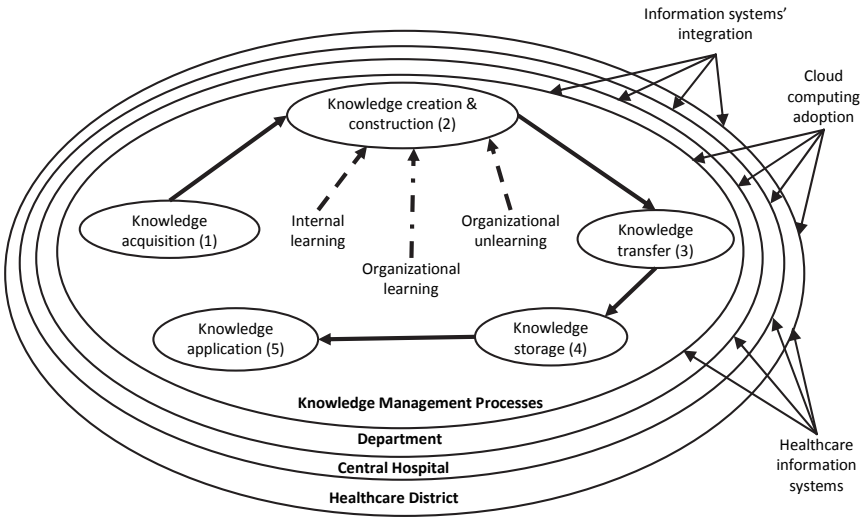


Fig. 1. The conceptual framework of the study.

In healthcare environment the knowledge is ingrained in the practitioners, and the knowledge exists in the medical practice and is stored to healthcare information

systems [1], [60]. In healthcare environment the amount of medical and nursing knowledge accumulates overtime, and the information collected from patients must be stored and updated to the information systems [13] in a way that supports medical practice needs 8], [61].

Knowledge acquisition involves searching for valuable knowledge, and external knowledge may be acquired by importing knowledge components directly or by depending on intermediaries [45], [32]. It has been argued that an organizational gatekeeper is the key individual who connects the organizational members to the external sources of information, and the organizational members are kept up-to-date with the outside information by communicating with the gatekeepers [46].

Learning influences knowledge creation, and knowledge provided by evidence-based medical guidelines and drug information databases help physicians to learn new issues [14]. The electronic patient records enable creation of organizational knowledge, and they are a useful tool to survive in everyday work in primary care [13], [14]. In decision-making and clinical practice, knowledge is transferrable through individual learning for example by observation [9], [62]. It is also possible that knowledge transfer can occur without the individual being aware of it happening [9]. Learning in groups occurs through discussions, meetings and lecture sessions in which people share their experiences [9].

It has been stated that knowledge transfer is the ability to apply knowledge gained in one situation in another similar situation, or to use metacognitive strategies to act in a novel situation [50]. New knowledge is generated by the influx of information into an individual's mind, combined with the existing knowledge of this individual, and then communicated further and made explicit [63]. After knowledge creation, it needs to be transferred throughout the healthcare organization. The knowledge that is relevant and right to be transferred needs to be determined as well [9]. The formal communities of practice include meetings, and the informal communities of practice will include discussion groups, study groups and online communities [53]. When more knowledge is shared between individuals, the more opportunities there are for knowledge creation [9]. Due to the reason that communities are formed with different ways of working and adoption of different vocabularies, they may not understand each other [9]. For example, human actors in IT and the business domain often speak different technical and procedural languages [63]. In addition one domain can articulate requirements, goals and constraints that another domain can think of as being unreasonable and uncooperative [63].

Knowledge storage can be defined as the organization's memory which comprises the knowledge and information that the people working in the organization possess through their skills and experiences. The collective memory of the organization in the organizational culture is expressed through the routines and attitudes inhabiting in groups and networks [48]. Organizational memory can be mental abilities and issues inside the organizational members, but also the information possible to retrieve, such as copies of memos, letters, spreadsheets, and data stored in computers constitute organizational memory [48]. It has been stated that knowledge management systems are a supporting class of systems to the organizational processes of knowledge management and knowledge storage [8]. The organization's computer-based communication and information system applications contain databases, repositories, directories, and networks [8].

Knowledge application is the ability to use the learned material in new and concrete situations by applying rules, methods, concepts, principles, laws, and theories [9]. The use of external knowledge will create new knowledge [45].

Internal learning has two modes of knowledge: tacit and explicit knowledge. Tacit knowledge is gained through clinical and practical experience [9]. It has been argued that the use of healthcare information systems has enhanced individual learning and group learning, and the physicians can achieve and create new knowledge by using information systems [60]. A physician will require also other knowledge than medical and clinical knowledge, such as technical skills, academic knowledge, a hospital's and healthcare organization's cultural knowledge, management know-how, and administrative skills. Healthcare organizations' knowledge can be transferred to an individual, a group or a system [51]. A patient relationship management system is affected by the impact of the knowledge work performed by the physicians in a hospital, and the use of the case system seems to enhance knowledge creation [13]. The knowledge transfer barriers are lowered between the physicians and the patients by enhancing communication through a follow-up system [13].

In organizational learning, knowledge is stored in databases in documents, and the learning entities are both the individual and organization [20]. This kind of organizational learning is an 'old organizational learning'. On the other hand, new organizational learning means discovering new theories, practices and innovation and then distributing or transferring that new knowledge to the organization [32], [20].

In organizational unlearning, the old organizational knowledge is disregarded. The knowledge considered for elimination is the same knowledge that led the organization to its previous success [64], and there is a need to remove or reject previously used practice from the organization [32]. Therefore, change and learning theories are relevant and should be included in a framework in order to draw a comprehensive image of processes at work in the changing organizations [64], [25].

It has been claimed that security and confidentiality issue has slowed down the cloud computing adoption in knowledge management systems in healthcare [31]. When using commercial cloud computing it is not known where the data is physically stored and how it is secured. This issue is problematic especially if the data is confidential [29], [59]. Of course, by creating own cloud computing environment this problematic issue related to the physical location of data is solved, but the security issues still remain. The communication between the healthcare actors and awareness of their relationships with each other is important in information systems' cloud computing adoption [44], [31]. Healthcare personnel must understand how the cloud computing adoption to information systems affects to the processes and how adoption on the other hand is affected by the relationship between personnel and processes [44], [31]. It has been further claimed that cloud computing adoption in its best would minimize costs of healthcare personnel's information retrieval because they can use laptops and mobile phones everywhere when using cloud services remotely, and thus cloud computing helps to simplify the management and access of patients' data [31].

Healthcare information systems' and information systems' integration both need integration requirements gathering in order to evaluate different approaches and industrial integration standards [30]. The goal of information systems' integration is to offer services to disease management [22], to promote and prevent healthcare to identify cause of illnesses, to help in medication and therapy, to offer rehabilitation services, offer long term care, to provide clinical healthcare, and to offer information

communication technologies (ICTs) in order to support management, administrative activities and logistic services [22], [30]. ICTs cover electronic patient records, electronic medical records, picture archiving communication systems, physicians' e-Prescriptions and e-Referrals, and portals with healthcare information and health cards [30]. Healthcare information systems and information systems' integration should also offer infrastructure, research, education, collaboration, healthcare knowledge infrastructure, clinical trials, medical education, local and international platforms and collaborations, efficient IT infrastructure and security, deployment of e-Health with efficient IT infrastructure, establish physical networks which allows connectivity supporting interoperability of various technologies and various systems, data integrity and security, and frame work of security and confidentiality [22], [30]. Thus, cloud computing adoption, information systems' integration, and healthcare information systems all impact to knowledge management in healthcare.

Our conceptual framework in healthcare has many dimensions and due to its complexity and lack of solid and matured knowledge management theories in healthcare it has a lot of challenges both for the academic studies, but also for the practical implementations. As already outlined healthcare is a knowledge driven process and hence knowledge management when properly studied and formalized and implemented provides an opportunity to improve the healthcare performance at its all levels.

4 Research Methodology

4.1 Data Collection

This study is a qualitative inquiry based on a case study approach [65], [66], [67] on the empirical data collected by interviews [16]. This method is best suited for social sciences, as it allows the researcher to interact with the society through interviews and observations for the purpose of acquiring the desired data, such as in our case for creating a knowledge management theory in healthcare. The researcher will be able to combine various data sources such as archival records, interviews, observations, audio recording, and even quantitative data for the analysis without restricting the data formats [68]. The Grounded Theory (GT) approach [35], [69] is used in data collection and analysis.

The individuals considered for the interviews need to have participated in the process or action, and they must be given the time and place to be interviewed [67]. Our study is in line with this, because the central hospital arranges the place and time for the interviews and the research coordinator arranges the interview timetables. It has also been highlighted the importance of type of sampling and the number of interviews needed [67]. Before the interview, permission is asked from the interviewee to use the tape recorder. Audio-recorded unstructured and semi-structured recordings of the interviews will be transcribed.

In our study a department in a central hospital is the unit of analysis and the study covers several departments in order to generalize the results to have a wider impact for the healthcare. A hypothesis was made that via a deep understanding of the selected case departments and the identification of their knowledge management

processes, categories, relationships between the categories, healthcare information systems' impact, cloud computing adoptions' impact, and information systems' integration impact to knowledge management in healthcare would provide a solid background for the generalization of the results. The primary data sources are both open-ended and structured interviews which create the possibility for individuals or groups to express themselves freely in a relaxing atmosphere. Furthermore, archival records are used as secondary data sources, and the background context data includes the annual and financial reports, and press releases of the South Karelia Social and Healthcare District [70], [71].

At first a pilot study was carried out in the Obstetrics and Gynaecology department in January - March 2013 including 10 interviews. The interview questions were predesigned and they were sent to the interviewees in advance [67]. The interviewees were highly motivated, and they were asked to describe how they use their own knowledge in medical and patient care. After this first pilot study, the interview questions were improved based on the received feedback. This reformulation of the research questions was needed in order to match better each of the different department's knowledge management. The chief physician of the Obstetrics and Gynaecology department, who acts also as the research site coordinator, has arranged the research permission and interviews. This is because in the hospitals and healthcare districts in Finland the national laws and regulations are very strict, and the interviews also need a specific time table not to affect the patient care work.

After the pilot study and reformulation of the interview questions, five other departments were included in the research, and new interview rounds have been and will be carried out in the following order (number inside parenthesis gives the number of interviews): in January - April 2014, a second round of interviews was carried out in the Obstetrics and Genecology department (10); in March - May 2014 the first interview round was carried out in the Paediatric department (5); in March - June 2014 the first interview round was carried out in the Paediatric Neurology department (4); in May 2014-June 2014 the first interview round was carried out in the Anaesthesia and Surgery department (10), and finally in May - June 2015 the first interview round will be carried out in the Surgical department (10). Furthermore, more interviews will be carried out in December 2015 - December 2018, including, e.g., hospital IT administration department (10).

The ethical issues related to the data are guaranteed both in the case of the primary data (interviews) and secondary data (the archival material). One of the risks is the difficulty to build up a theory from the empirical case studies. In addition, it may turn out that a single theory is not sufficient to cover all aspects in a required and selected detail level. The validity and reliability of the research can be although affected by the research design, data collection and methods, quality of data, analyses of data, presentation of the results and making the conclusions [72], but simultaneous triangulation from different data sources, such as archival material will improve the validity and reliability of data, and the data is stored and protected according to the Finnish laws. The legal permission to use central hospital as the research site was granted in December 2012 by the Social and Healthcare District's Service Director, because every research which needs attendance and interviews of the staff must be approved by the service director. Each of the interviewees has been and will be asked in the future a permission to use the interviewee material as the data in the study and

their anonymity is guaranteed. The collected data material is not allowed to be taken abroad and hence it will stay in Finland.

4.2 Data Categorization and Analysis with the Grounded Theory

The pre-classified data will be analyzed first with grounded theory (GT) approach [35] which allows the researcher to interact with the society through the interviews and observations for the purpose of acquiring the desired data. This methodology will allow the researcher to combine the various data sources such as interviews, and observations, and even quantitative data for the analysis without restricting the data formats [68].

The knowledge management research problems are the basis for interviews and data collection. The research problems are presented to the interviewees, and they are chosen because their role is to use, create and transfer healthcare-related medical and ICT information, and translate it to knowledge relevant to the healthcare situation at hand. We will use fragmentation and reassembling based on the researchers' own intuition and knowledge in order to categorize our data into thematic categories by trying to capture a broader social system of ideas from the experience of the social actors working in the Social and Health Care District [35], [69]. After the categories have been found, we determine the properties of the categories and propositions (hypotheses) for how the categories were related. The constant comparison between the data and concepts in the past studies in order to accumulate evidence convergence on simple and well-defined categories will led us to a higher level of abstraction of statements about the relationships between the categories. This theorizing is in line with GT approach suggestions in creating a theory [69], [35]. Finally, in the future we will develop several conceptual frameworks of the discovered categories, and their relationships between each other [35], [69]. Furthermore, in our in-depth case studies, we must take carefully take into consideration beforehand who to interview, what to do next, what group to look for, and what additional data we should collect in order to develop a grand theory from the emerging data.

4.3 Quantitative Data Analysis

As this is research is also quantitative in its nature, the collected qualitative data will be converted to quantitative form in order to carry out the needed statistical and computational analyses. The exploratory data analysis approach is needed for generating hypothesis due to weaker assumptions and prior knowledge about the data and the domain. Data mining techniques are tools for exploratory data analysis [36]. The goal of data mining is to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful for the goals of the project. This includes visualization, projection methods, clustering, regression, classification, and association analysis such as association rule mining techniques. Especially data visualization methods, such as the Self-Organizing Maps [73], [74], Bayesian networks [75], and multidimensional scaling and hierarchical clustering [36], are needed for the deeper understanding of domain and variable dependencies. Quantitative analysis covers both linear and non-linear methodologies combined with

variable selection and uncertainty analysis but also classical hypothesis testing to reject or accept hypotheses whenever available.

As a general framework (Figure 2) for quantitative data analysis, the Cross Industry Standard Process for Data Mining (CRISP-DM) [36] approach is followed to guarantee that no important steps in data analysis are missed.

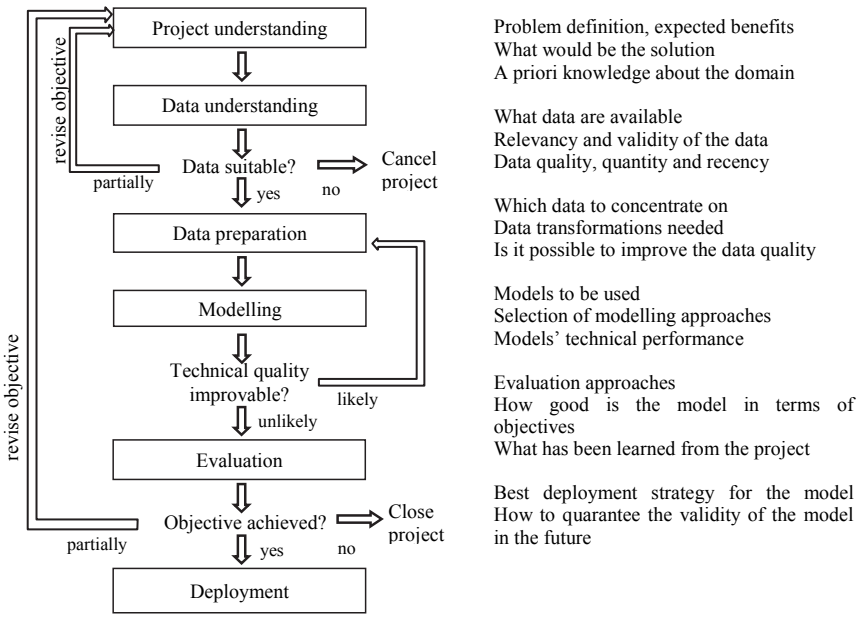


Fig. 2. CRISP-DM framework for data analysis.

5 Conclusions and Discussion

The primary theoretical and scientific outcome and contribution of the project is to create a theory of knowledge management in healthcare in specifically in the hospital environment based on empirical findings. Here the term theory should be understood as 1) a new conceptual framework of the knowledge management processes, and 2) new knowledge management categories, and the relationships between the categories in healthcare environments. Furthermore, the healthcare information systems' impact, cloud computing adoption's impact, and information systems' integration impact to knowledge management in healthcare are also studied. In this research project the research environment is the central hospital of the South Karelia Social and Healthcare District, Finland.

In general, when discovering a new theory, a multiple case study approach should be applied [66], [72]. In our study a department in the central hospital is our unit of analysis. The sample, however, has not been limited to one department, but several departments, because the goal of the study is to achieve a deep understanding of the selected case departments and to identify their knowledge management categories, relationships and processes. Theory creation should also combine multiple data

collection methods due to the triangulation in order to provide stronger substantiation of categories. Collecting different types of data by different methods from different sources produces a wider scope of coverage may result in a fuller picture of the phenomena under study. Thus, both quantitative and qualitative data are used in this study [72]. The flexibility given by Grounded Theory (GT) on the other hand gives respondents an ability to express their views and opinions easily and freely [66], [72]. Therefore, the methodological scientific contribution of this study is to utilize a new methodological approach where both diverse qualitative research methods such as Grounded Theory (GT) [35] and quantitative research analyzing approaches are applied. As the quantitative research approach we use novel intelligent computing and analyzing methods, and as a general framework, the Cross Industry Standard Process for Data Mining (CRISP-DM) [36] approach has been selected. Constant comparison between the data and concepts will be made so that accumulating evidence converges on simple and well defined constructs. The boundary conditions of the theory, however, have to be taken into account, because the phenomenon is so atypical that it holds only in this specific contextual healthcare environment.

The practical and managerial contributions of this study are as follows. First, to help physicians and nurses to understand their own valuable knowledge capital and practice, to understand knowledge management better, and to get familiar with knowledge management practices in the hospital. Second, to develop knowledge transfer from the physicians to nurses, and vice versa. Third, with the discovered knowledge new and user friendly knowledge management processes could be remodeled. Fourth, the hospital based knowledge could be used later to implement more user-friendly healthcare information systems. Of course, it may also turn out that the data will not contain enough information to derive valid and solid knowledge management categories and therefore a very careful analysis has to be carried out to find out if the categories discovered in the data are the correct ones.

The results of this project are expected to gain a lot of interest in other Finnish social and healthcare districts and most probably the results are applicable to many hospitals due to their similarity. Internationally this study can offer guidelines and good practices to follow up in the hospitals abroad, and also to improve their ability to better and safety patient care. The research project will have several collaboration partners which have special knowledge in medical science, nursing science, sociology, intelligent computing and systems, information systems, and software engineering. The collaboration partners include both other Finnish Social and Healthcare Districts and international cooperation.

References

1. Gold, A.H., Malhotra, A., Segars, A. H.: Knowledge Management: An Organizational Capabilities Perspective. Journal of Management Information Systems 18 (2001) 185-214
2. Lahtiranta, J.: Current challenges of personal health information management. Journal of Systems and Information Technology 11 (2009) 230-243
3. Mead, N., Bower, P.: Patient-centeredness: A conceptual framework and review of the empirical literature. Journal of Social Science Medicine 51 (2000) 1087-1110
4. Thornton, T.: Tacit knowledge as the unifying factor in evidence based medicine and clinical judgement. Philosophy, Ethics and Humanities in Medicine 1 (2006) 1-10

5. Leape, L., Berwick, D., Clancy, C., Conway, J., Gluck, P., Guest, J., Lawrence, D., Morath, J., O'Leary, D., O'Neill, P., Pinakiewicz, D., Isaac, T.: Transforming Healthcare: A Safety Imperative. Quality and Safety in Healthcare 18 (2009) 424-428
6. Hernández, J.G., Noruzi, M. R.: How intellectual capital learning organization can foster organizational competiveness? International Journal of Business and Management 5 (2010) 183-193
7. Nonaka, I.: A dynamic theory of organizational knowledge creation. Organization Science 5 (1994) 14–37
8. Alavi, M., Leidner, D.E.: Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. MIS Quarterly 25 (2001) 107-136
9. Dalkir, K.: Knowledge Management in Theory and in Practice. Butterworth-Heinemann Publisher, London (2005)
10. Boateng, W.: Knowledge Management in Evidence-Based Medical Practice: Does the Patient matter? Electronic Journal of Knowledge Management 8 (2010) 281-292
11. Grover, V., Davenport, T.H.: General perspectives on knowledge management: Fostering a research agenda. Journal of Management Information Systems 18 (2001) 5-21
12. Stair, R. M., Reynolds, G.W.: Fundamentals of Information Systems. Cengage Learning, Boston Mass (2006)
13. Oinas-Kukkonen, H., Räisänen, T., Hummastenniemi, N.: Patient Relationship Management: An Overview and Study of a Follow-Up System. Journal of Healthcare Information Management 22 (2008) 24-30
14. Oinas-Kukkonen, H., Räisänen, T., Leviskä, K., Seppänen, M., Kallio, M.: Physicians' user experience of mobile pharmacopoeias and evidence based medical guidelines. International Journal of Healthcare information Systems and Informatics 4 (2009) 57-68
15. Polanyi, M.: The Tacit Dimension. Routledge, London (1966)
16. Heilmann, P.: To have and to hold: Personnel shortage in a Finnish healthcare organization. Scandinavian Journal of Public Health 38 (2010) 518-523
17. Puusa, A., Eerikäinen, M.: Is Tacit knowledge really Tacit? Electronic Journal of Knowledge Management 8 (2010) 307-318
18. Iwai, S., Ishino, F.: Translating tacit medical knowledge into explicit knowledge. Journal of Knowledge Management Practice 10 (2009)
19. Hill, K.S.: Improving quality and patient safety by retaining nursing expertise. The Online Journal of Issues in Nursing 15 (2010)
20. Morr, C.E.L., Subercaze, J.: Knowledge Management in Healthcare. In: Cunha, M.M.C., Tavares, A.J., Simões, R. (eds.): Handbook of Research on Developments in eHealth and Telemedicine: Technological and Social Perspectives. IGI Global, New York (2010) 490-510
21. Wegner, D. M.: Transactive memory: A contemporary analysis of the group mind. In: Mullen, B., Goethals, G.R. (eds.): Theories of group behavior. Springer-Verlag, New York (1986) 185-208.
22. Stroetmann, K.A., Artman, J., Stroetmann, V.N., Protti, D., Dumontier, J., Giest, S., Walossek, U., Whitehouse, D. eHealth Strategies: European countries on their journey towards national eHealth infrastructures- evidence on progress and recommendatiobs for cooperative actions. Final European Progress Report. European Commission DG Information Society and Media, Europe (2011)
23. Hsiao, H-C., Chang, J-C.: The role of organizational learning in transformational leadership and organizational innovation. Asia Pacific Education Review 12 (2011) 621-631
24. Cepeda-Carrión, G., Cegarra-Navarro, J.G., Leal-Millán, A.G.: Finding the hospital in the home units' innovativeness. Management Decision 50 (2012) 1596-1617
25. Venable, J.J., Pries-Heje, J., Bunker, D., Russo, N.L.: Creation, transfer and diffusion of innovation in organizations and society: Information systems design science research for human benefit. In: Pries-Heje, J., Venable, J.J., Bunker, D., Russo, N.L., DeGross, J.I.

(eds.): Human Benefit through the Diffusion of Information Systems Design Science Research Vol. 1-10. Springer-Verlag: Berlin Heidelberg (2010)
26. Sagher, F.: (2008) Healthcare Systems (Part 1) the British Experience. Jamahiriya Medical Journal 8 (2008) 80-83
27. Dobrev, A., Jones, T., Stroetmann, V., Stroetmann, K., Vatter, Y., Peng, K.: Interoperability eHealth is Worth it: Securing Benefits from Electronic Health Records and ePrescribing. European Commission DG Information Society and Media, Europe (2010)
28. Doupi, P., Renko, E., Hämäläinen, M., Mäkelä, M., Giest, S., Dumortier, J.: eHealth Strategies: Country Brief: Finland. European Commission DG Information Society and Media, Europe (2010)
29. Kuo, A.M.: Opportunities and challenges of cloud computing to improve healthcare service. Journal of Medical Internet Research 13 (2011) 2-22
30. Gattu, C.: Information systems in healthcare domain and integration. Lappeenranta University of Technology, Kirjansitomo Hulkkonen (2013)
31. Paiti, T.: What are the opportunities and challenges of cloud computing in the healthcare information systems? Lappeenranta University of Technology, Kirjansitomo Hulkkonen (2013)
32. Mustonen-Ollila, E.: Information system process innovation adoption, adaptation, and unlearning: A longitudinal case study (Diss). Lappeenranta University of Technology, Digipaino (2005)
33. Mustonen-Ollila, E.: Information System Process Innovation Life Cycle Model. Journal of Knowledge Management: In The Knowledge Garden 8 (2007) 1-12
34. NIHCE: National Institute for Health and Clinical Excellence: Behaviour of Change at Population, Community and Individual Levels. NICHE Public Health Guidance, London (2007)
35. Glaser, B., Strauss, A.L.: The Discovery of the Grounded Theory: Strategies for Qualitative Research. Aldine, Chicago (1967)
36. Berthold, M.R., Borgelt, C., Höppner, F., Klawonn, F.: Guide to Intelligent Data Analysis: How to Intelligently Make Sense of Real Data. Texts in Computer Science, Vol. 42. Springer-Verlag, Berlin Heidelberg New York (2010)
37. Argote, L.: Organizational Learning: Creating, Retaining and Transferring Knowledge. Kluwer Academic Publishers, Massachusetts (1999)
38. Mustonen-Ollila, E., Lyytinen, K.: Why organizations adopt IS process innovations: A longitudinal study using Diffusion of Innovation Theory. Information Systems Journal 13 (2003) 275-297
39. Mustonen-Ollila, E., Lyytinen, K.: How Organisations adopt IS process innovations: A Longitudinal Analysis. European Journal of Information Systems 13 (2004) 35-51
40. Mustonen-Ollila, E., Heikkonen, J.: IS process innovation evolution in organizations. Academy of Information and Management Sciences Journal 11 (2008) 65-88
41. Mustonen-Ollila, E., Heikkonen, J.: Historical research in information system field: from data collection to theory creation. In: Cater-Steel, A., Al-Hakim, L. (eds.): Information Systems Research Methods, Epistemology, and Applications. Information Science Reference (an imprint of IGI Global), Hersey New York (2009) 140-160
42. Greig, G., Entwistle, V.A., Beech, N.: Addressing complex healthcare problems in diverse settings: Insights from activity theory. Social Science & Medicine 74 (2012) 305-312
43. Fichman, R.G. Kohli, R., Krishnan, R.: The Role of Information Systems in Healthcare: Current Research and Future Trends. Editorial Review of Information Systems Research 22 (2011) 419-428
44. Mantzana, V., Themistocleous, M., Irani, Z., Morabito, V.: Identifying healthcare actors involved in the adoption of information systems. European Journal of Information Systems 6 (2007) 91-102
45. Huber, G.P.: Organizational learning: The contribution processes and the literatures. Organization Science 2 (1991) 88-115

46. Whelan, E., Donnellan, B., Golden, W.: Knowledge diffusion in contemporary R&D groups: Re-examining the role of the technological gatekeeper. In: King, W.R. (eds.): Knowledge Management and Organizational Learning. Vol. 3-13. Springer-Verlag, USA (2009)
47. Lehner, F., Maier, R.K.: How Can Organizational Memory Theories Contribute to Organizational Memory Systems? Information Systems Frontiers 2 (2003) 277-298
48. Paoli, M., Prencipe, A.: Memory of the organization and memories within the organization. Journal of Management and Governance 7 (2003) 145-162
49. Latour, B.: Materials of Power: Technology is society made durable. In: Law, J. (eds.): A Sociology of Monsters: Essays on Power, Technology, and Domination. Routledge, London (1991) 103-131
50. Lauder, W., Reynolds, W., Angus, N.: Transfer of knowledge and skills: some implications for nursing and nurse education. Nurse Education Today 19 (1999) 480-487
51. Williams, A.M., Baláž, V.: International return mobility, learning and knowledge transfer: case study of slovak doctors. Social Science and Medicine 67 (2008) 1924-1933
52. Lin, C., Tan, B., Chang, S.: An exploratory model of knowledge flow barriers within healthcare organizations. Information and Management 45 (2008) 331-339
53. Wang, W., Lu, Y.: Knowledge transfer in response to organizational crises: An exploratory study. Expert Systems with Applications 37 (2010) 3934-2942
54. Ferlie, E., Crilly, T., Jashapara, A., Peckham, A.: Knowledge mobilization in healthcare: A critical review of health sector and generic management literature. Social Science & Medicine 74 (2012) 1297-1304
55. Eppler, M.J.: Knowledge communication problems between experts and decision makers: an overview and classification. The Electronic Journal of Knowledge Management 5 (2007) 291-300
56. Becker, K.: Facilitating unlearning during implementation of new technology. Journal of Organizational Change and Management 23 (2010) 251-268
57. Khoumbati, K. Themistocleous, M., Irani, Z.: Evaluating the adoption of Enterprise Application Integration in Healthcare Organization. Journal of Management Information Systems 22 (2006) 69-108
58. Spyrou, S., Bamidis, P., Chouvarda, I., Gogou, G., Tryfon, M., Maglaveras, N.: Healthcare Information Standards: Comparison of the Approaches. Health Informatics Journal 8 (2002) 14-19
59. Malin, A. B., Emam, E. K., O'Keefe, C.M.: Biomedical data privacy: problems, perspectives, and recent advances. Journal of the American Medical Informatics Association 20 (2013) 2-6
60. Räisänen, T., Oinas-Kukkonen, H., Leiviskä, K., Seppänen, M., Kallio, M.: Managing Mobile Healthcare Knowledge: Physicians' Perceptions on Knowledge Creation and Reuse. In: Olla, P., Tan, J. (eds.): Mobile Health Solutions for Biomedical Applications. IGI-Global, New York, New Hersey (2009) 111-127
61. Menachemi, N., Brooks, R.G., Schwalenstocker, E., Simpson, L.: Use of Health Information Technology by Children's Hospitals in the United States. Official Journal of the American Academy of Pediatrics 123 (2009) 80-84
62. Hall, A., Walton, G.: Information overload within the healthcare system: a literature review. Health Information and Libraries Journal 21 (2004) 102-108
63. Blumenberg, S., Wagner, H., Beimborn, D.: Knowledge transfer processes in IT outsourcing relationships and their impact on shared knowledge and outsourcing performance. International Journal of Information Management 29 (2009) 342-352
64. Turc, E. and Baurnard, P.: Can organizations really unlearn? In: Mcinerney, C.R., Day, R.E. (eds.): Rethinking Knowledge management. Springer-Verlag, Berlin Heidelberg (2007) 125-146
65. Maxwell, J.A.: Qualitative Research Design: An Interactive Approach. Sage Publications, Thousand Oaks (1996)

66. Yin, R.K.: Case Study Research: Design and Methods. Sage Publications, California (2003)
67. Creswell, J.W.: Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Sage Publications, California (2007)
68. Joan, R., Pastor, A. J.: Applying Grounded Theory to Study the Implementation of an Inter-Organizational Information System. Electronic Journal of Business Research Methods 5 (2007) 71-82
69. Pawluch, D., Neiterman, E.: What is Grounded Theory and Where Does is Come from? In: Bourgeault A., Dingwall, R., De Vries. R. (eds.): The Sage Handbook of Qualitative Methods in Health Research. Sage Publications, London (2010) 174-192
70. Miikkulainen Plus.: South Karelia's Social and Healthcare District. Eksote, Lappeenranta (2010a)
71. Miikkulainen Plus.: South Karelia's Social and Healthcare District. Eksote, Lappeenranta (2010b)
72. Eisenhardt, K.M.: Building Theories from Case Study Research. Academy of Management Review 14 (1989) 532-550
73. Kohonen, T.: Self-Organization and Associative Memory. Springer-Verlag, Berlin (1989)
74. Kohonen, T.: Self-Organized Maps. Springer-Verlag, Berlin (1995)
75. Heckerman, D.: A tutorial on learning with Bayesian networks. Microsoft Research Advanced Technology Division, Redmond (1996)

PEGASO Fit for Future

Renata Guarneri and Giuseppe Andreoni

Politecnico di Milano – Design Dept., Via Durando 38/A, 20158 Milano, Italy
{mariarenata.guarneri, giuseppe.andreoni}@polimi.it

Abstract. Challenging teen-agers in their own fields and areas of interest, PEGASO – Fit 4 Future - aims at promoting a sustainable behavior change towards healthy lifestyles, with a holistic and multidisciplinary approach. PEGASO is based on a mobile, social and networked gaming platform, considered as a powerful tool to actively engage the younger population in activities that will stimulate healthier choices in their daily lives to counterfeit sedentariness, overweight up to obesity. To support this action, in the platform, represented by the smartphone, three main functionalities are implemented: an individual & environmental monitoring through wearable devices, a feedback system for providing a feedback in terms of "health status" changes, the Social connectivity and engagement to support motivation.

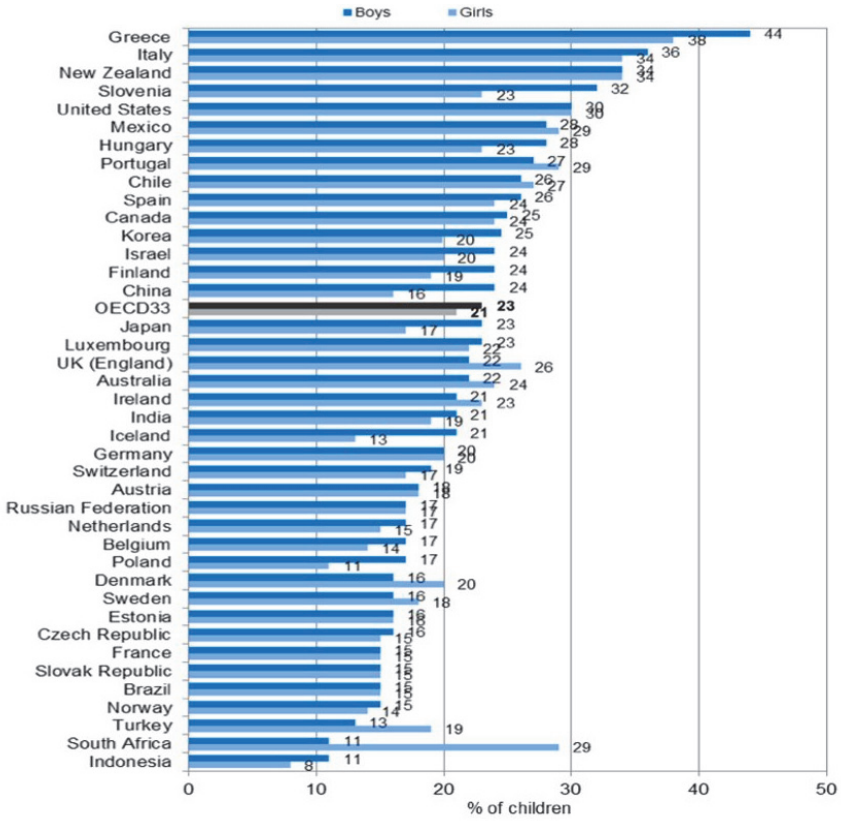
1 Overview and Rationale

The rapidly increasing prevalence of overweight and obesity among children and adolescents reflects a global 'epidemic' worldwide. Recently the US Center for Disease Control and Prevention has evidenced that "Childhood obesity has more than doubled in children and quadrupled in adolescents in the past 30 years."[1, 2].

Fig. 1. The WHO web page on the priority to overweight and obesity prevention in children (last accessed march 18, 2015).

Also WHO recognized prevention of juvenile overweight and obesity as a priority for future health being in 2014 and in the WHO European region 1 on 3 11-years old children in this condition.

The following histogram represents the dramatic situation and dimension of the problem at European and worldwide level.



Source: International Association for the Study of Obesity, 2013; Bös et al. (2004), Universität Karlsruhe and Ministères de l'Education nationale et de la Santé for Luxembourg; and KNHANES 2011 for Korea.

Fig. 2. Measured overweight (including obesity) among children aged 5-17 in 2010 or nearest year.

Due to the associated serious medical conditions, it is estimated that obesity already accounts for up to 7% of healthcare costs in the EU, as well as costs to the wider economy associated with lower productivity, lost output and premature death. Obesity in younger age groups has been recognized as an alarming key predictor for obesity in adulthood, but also entails a number of short term health complications in juvenile age such as hypertension, type 2 diabetes, metabolic syndrome, fatty liver disease, sleep disturbances along with greater risk of social and psychological problems [3, 4].

Sedentariness and over-consumption of high calorie foods and beverages are *a priori* determinants of overweight/obesity and poor health status also in adolescents, according to well-grounded evidences.

> *"Prevention is of obvious importance and there is an urgent need for further research into how physical activity and training, in addition to nutrition, can prevent the steadily increasing average body mass index of Europeans. This proposal includes a vision that integrates a lifestyle of healthy habits with an environment that promotes healthy living by encouraging exercise and making healthy food affordable."* (Visions for Horizon 2020) [3].

Healthy lifestyle habits, including healthy eating and physical activity, can lower the risk of becoming obese and developing related diseases. The dietary and physical activity behaviors of children and adolescents are influenced by many sectors of society, including families, communities, schools, child care settings, medical care providers, faithbased institutions, government agencies, the media, and the food and beverage industries and entertainment industries.

2 Methodological Approach and Concept of the PEGASO Project

Juvenile obesity is a complex disorder with many interrelated consequences. Addressing the obesity issues requires a *comprehensive approach* taking into account the individual's *physical-physiological* characteristics, *personality* as well as the *social* and *psychological* environments influencing decisions and habits in their everyday life. Challenging teen-agers in their own fields and areas of interest, PEGASO – Fit 4 Future - aims at promoting a sustainable behavior change towards healthy lifestyles, with a holistic and multidisciplinary approach. The approach of PEGASO is based on three level of intervention enabling teen-agers to become co-producers of their wellbeing:

1. Generating self-awareness (acknowledgement of risks associated to unhealthy behaviors),
2. Enhancing and sustaining motivation to take care of their health with a short/medium/long term perspective,
3. Changing behavior towards a healthy lifestyle based on healthy diet and adequate physical activity.

In order to achieve the above targets PEGASO applies behavior change techniques to prevention and will develop a mobile-based Behavior Change Platform that can effectively address teen-agers.

The solution proposed by PEGASO comes from the convergence of the need to address through appropriate preventative measures the rapidly increasing prevalence of obesity among children and adolescents on one side and the rapid development of ICT, and in particular mobile technologies, together with their increasing diffusion among the EU population, on the other side.

Indeed the capabilities of ICT technologies (i.e. mobile phones, digital tablets) together with the possibility to integrate them with the new and fashionable additional technologies for data acquisition (e.g. wearable sensors) offer the opportunity to develop an effective behavior change platform. The solution proposed by PEGASO is

therefore based on a **mobile, social and networked gaming platform**, considered as a powerful tool to actively engage the younger population in activities that will stimulate healthier choices in their daily lives.

From the technology point of view, cloud computing, and convergence towards mobile are the key enablers.

The central element of the PEGASO platform is the smartphone. Europe's smartphone penetration is already amongst the highest in the world, with projections indicating a target of 55% in 2015. The smartphone indeed offers unique characteristics that ensure to achieve a cost-effective and scalable solution and may favour its adoption by the target audience.

In addition to the smartphone, wearable sensors - that can be added with a modular approach and a cloud based service offering based on a behavior and situation recognition system - represent the basic elements of the PEGASO platform architecture.

The Figure 3 provides an overview of the PEGASO architecture.

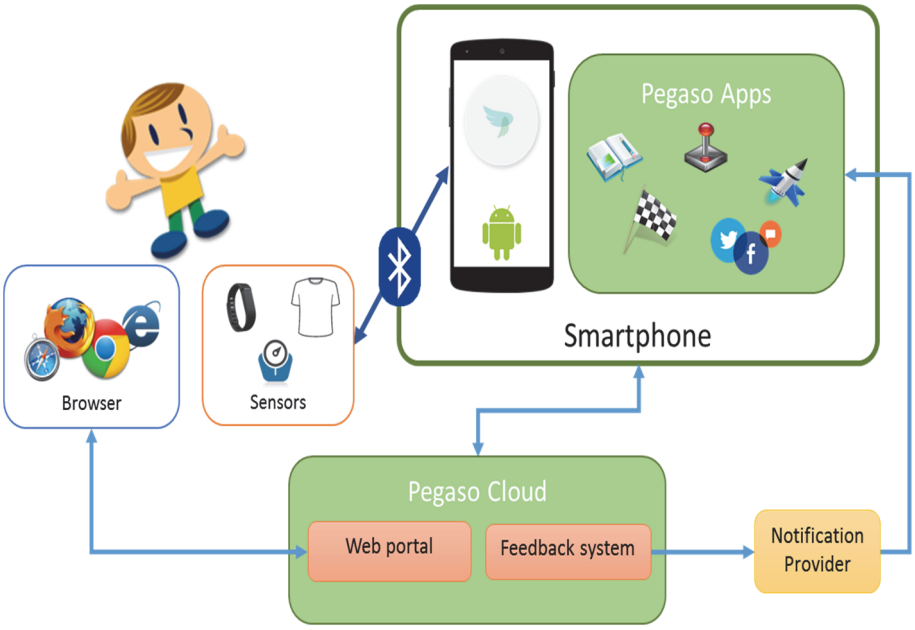


Fig. 3. Key elements of PEGASO architecture and their inter-relations.

2.1 PEGASO Key Functions

The PEGASO system framework is addressing prevention, by offering to teenagers – the primary target of PEGASO - three main functionalities:

1. **Individual & Environmental Monitoring** - This dimension consists of the environmental, behavioral and physiological analysis of young users, through a high level-monitoring platform including wearable sensors and mobile phone as

well as multimedia diaries for the acquisition of physical, behavioral and emotional attitude of adolescent.

2. **Feedback System** - This second functionality is aimed at providing a feedback in terms of "health status" changes, required actions to undertake and so on. This function will also propose personalized healthy modification of the lifestyle (in terms of diet and/or physical activity), thus promoting the active involvement of adolescents in changing their behaviors.
3. **Social Connectivity and Engagement** - The third dimension extends to include a social network where the user can share experiences with a community of peers concerning e.g. physical activity, food consumptions and everyday habits through different gaming strategies.

PEGASO has adopted a User Centred Design approach (UCD) by considering the target population (i.e. teenagers) at the centre of the system in a palingenetic process [5]. The UCD approach integrates three main elements: user involvement in all stages of the problem solving process; multidisciplinary research and development team; and iterative design process to refine the solution set. The main target users in PEGASO project are teenagers; however there are also several actors (who are also secondary users – see Figure 4) and products (as shown by the inner circle in Figure 4) involved.

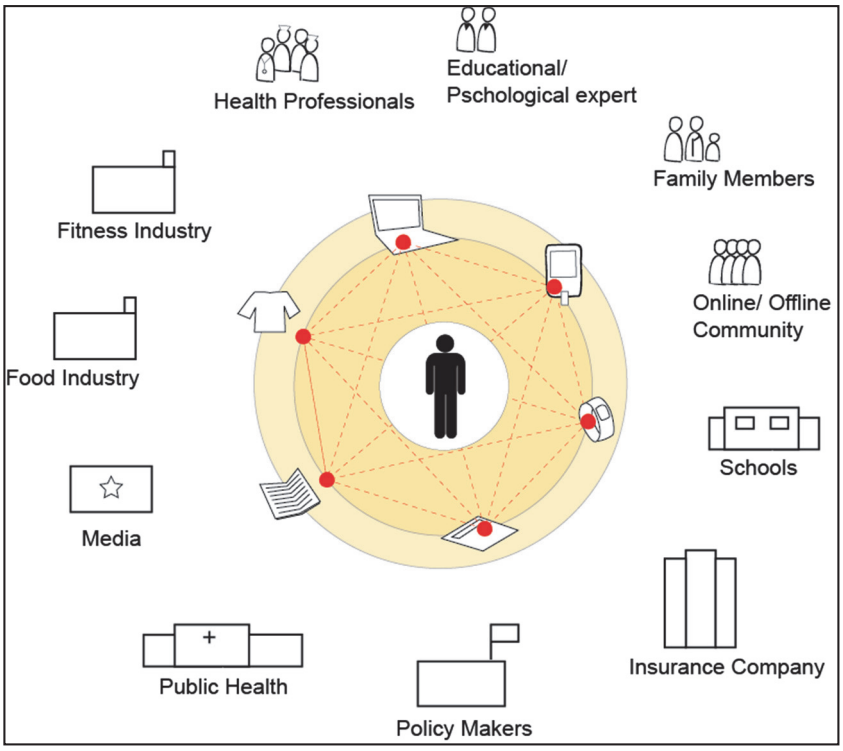


Fig. 4. PEGASO Stakeholders' System.

The ecosystem of stakeholders and enablers is composed of three main parts that are integrated in the user centred PEGASO system: *technological frame* (multimedia

diaries, embedded sensors systems, mobile & web platform), *services frame* (stakeholders services to provide answers to users' needs and desires in real time/not real time, from the health companion to the serious gaming and social experiences) and *experts layer* (which are knowledgeable groups of people from different disciplines - medical/psychological/educational – able to interact with the system, who provide them with filtered accurate and needed information to reach their PEGASO objective).

Technological Frame. Teens are familiar with Internet, social networks, mobile phones and apps, video gaming and, in general, with all the ICT platforms. Smartphones also assures the highest level of technology acceptance. This key issues are assumed as technological starting point to define the PEGASO architecture and to define a successful strategy to empower the teen-agers awareness about healthy lifestyle. The huge amount of personal and social exchanged and/or stored data includes also health records, thus posing severe reliability and security requirements that will be effectively managed through a cloud platform. Finally, PEGASO apps and games from the software layer, as well as wearable sensor and other more traditional systems (balance board for instance) complete the PEGASO technology frame.

Services Frame. Social is the key word for service development: the services created by stakeholders in PEGASO Project promote an individual and social healthier lifestyle through motivating and engaging multiuser serious games. Nevertheless individual support is provided both for data entry through multimedia apps that simplify and engage the users (for instance through multimedia diary compilation or through the health companion interaction). The health companion, developed by PEGASO, constitutes the interface on the smartphone between the guidance system and the teenager. All the stakeholders (including the Food Industry, Public and Private Health Policy actors, Fitness industries, Media, Schools, and Insurance companies) at different levels will offer to users the infrastructure to motivate (and promote) the adoption a healthy behavior.

Experts Layer. in PEGASO motivation and engagement by means of gaming strategies will be integrated with healthier lifestyle. All the information from the users must be "handled" and processed and the corresponding feedback provided. This means building an expert layer that is able to analyse all the data and deliver the resulting answers to the teenagers. A part of this layer will be composed by automatic algorithm (for real-time processing and feedback provision when applicable); a second building block will be the experts' team who will integrate the previous assessment to better stimulate the teenagers' consciousness about obesity and their motivation to adopt a healthy lifestyle. The role of experts in PEGASO project is assumed to be twofold: 1) to personalize information for each individual's physical and psychological models (i.e. personalized care) in order to reach the full acceptance by each teenager and guarantee a correct interpretation; and 2) to follow up of each teenager healthy status.

2.2 The Elements of Behavior Change

PEGASO considers four levels of engagement towards persuasion for user (the teen-agers) empowerment in healthcare [6]: awareness of obesity risks, motivation, affective learning and finally behavior change.

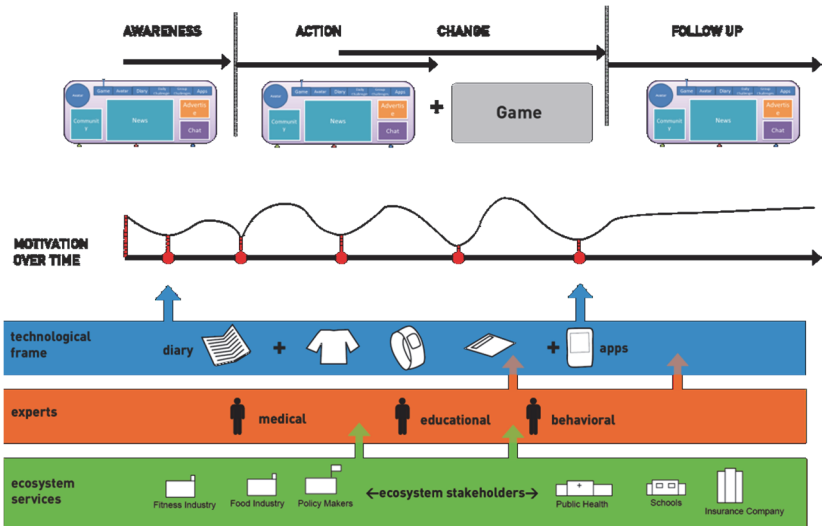


Fig. 5.

Various types of expertise / knowledge and technologies feed these levels of engagement towards healthier lifestyle and empower the teen-agers to take decisions accordingly.

Develop Awareness: teenagers need to be aware of what they are doing; what is right and what is wrong for their healthy living. Some of them are unconsciously and automatically acting, and often under estimate or have no clear notion about information they receive. Monitoring lifestyle of teen's activity, collecting parameters and integrating their own data will enable self-awareness on their current situation. Through developing self-awareness and self-reflection, the user can frame the problem or the opportunity area to act upon or intervene.

Affective Learning is the "highest" learning goal. The learner should trust in something that happens in several years. That is also a good argument in a new "learning level" to use a constructivist learning model and special media like "social games" are adequate to reach this goal. Giving teenagers information through tools they are affectionate is a strategy to reach their behavioral change.

Create Motivation: it is important to motivate teenagers to change their behavior and keep this activity in a long-term period. The actors in the ecosystem offer healthier benefits and services in the users' environment towards satisfying their needs or desires.

This part is quite challenging, since the motivation depends on many factors as well as emotions, psychological environment and personality of teens. The system needs to provide constant different layouts of motivational activities where experts,

technological frame monitoring and stakeholders services come into the scene.

Enable Behavior Change: once teenagers have the awareness and the motivation, it is important to involve experts and use PEGASO system to support the behaviour change process and reinforce existing virtuous behaviors. The turn from old unhealthy behaviors into healthier new ones has to be monitored through technology on a longer period.

In order to create prevention, it is important to change or stop old unhealthy habits and develop new healthier habits. In this respect, PEGASO takes a holistic approach involving the teenager's environment and specifically the families, by means of an education process empowered by training that will be provided on location (schools) and on line. The expert team will give feedbacks to the users allowing them to change their behavior on a long-term basis. The overall system takes advantage of gaming strategies to persuade users to change their behavior.

2.3 PEGASO Evaluation Strategy

PEGASO will be validated by secondary school students with the support of their schools and families. The reason for involving these students as sample population lies on the assumption that around the age of 14 years old the teen-agers acquire more independency and have increasingly the opportunity to make own and independent decisions. It is therefore important that at this stage they become aware of the consequences of inappropriate, unhealthy lifestyles.

Four validation studies will be carried out in Italy (Lombardy), Spain (Catalonia) and United Kingdom (England/Scotland).

The validation activities will assess the following aspects of the PEGASO Behavior Change Platform:

- System and Technology **acceptance**, usability and long-term use: these will constitute also a secondary assessment of motivation and engagement;
- **Reliability** in assessing the teen-agers lifestyles and their changes (with focus on the eating habits and on physical activities) and related efficacy of the sensing platform (i.e. smartphone and wearable sensors' system);
- **Efficacy** of the system in encouraging lifestyle change;
- **Subjective assessment** for awareness;
- System's **compliance** to Stakeholders' needs.

Further studies are required, and are currently out of PEGASO's scope, to be able to evaluate the longer term outcomes of the intervention, and to perform in particular the evaluation of the user risk awareness regarding the development of obesity and related comorbidities, the evaluation of user environmental factors (family and school), the potential harms and costs, all of them assessable involving the same target groups few years after the project completion.

3 Activities and Results

PEGASO Fit for Future has been running for over a year, with the first year being

dedicated mostly to the consolidation of requirements from the user viewpoint and the general architecture of the platform.

3.1 The Comprehensive Virtual Individual Model

The PEGASO project aims at pushing this concept further introducing the feature of dynamically selecting the opportune tailored interventions based on the user's individual characteristics and interaction context [7]. Tailoring the intervention involves modeling the user's characteristics and for this purpose it has been developed the Virtual Individual Model, which comes from the concept of the Virtual Physiological Human. The latter is a methodological and technological framework for integrated modeling of a living human body that describes the interaction of all the physiological components of individuals from molecular to apparatus level [8]. The Virtual Individual Model aims to include individual's characterization composed of physiological, physical, and psychological determinants. This allows integrating biological aspects of human functioning with lifestyle behaviors and psychosocial externalities that are crucial for the determination of the adoption of a certain life-style. This model is integrated in the system through an ontology-based virtualization. This process allows turning the information contained in the Virtual Individual Model into a structured knowledge that can be dynamically updated and elaborated by the computer to select the best interventions for each individual. Tailored interventions make the information personally relevant and researches demonstrated that computer-tailored health education is more effective in motivating people to make dietary changes [9] and that it could be also a good practice to promote physical activity [10]. 3 Tailored Intervention Forms The Virtual Individual Model characterizes the user's nutritional habits, physical status, and psychological status to provide personalized intervention to foster the adoption of a healthy life-style. Obviously, the interaction between the system and the user plays a crucial role in the tailoring process and to facilitate the effectiveness of the intervention. Since the teenagers are the targets of the PEGASO project, the smartphone has been chosen as the mediator of the interaction. Indeed, the smartphones are already perceived as a companion and it is most likely that this relationship between user and smartphone will strengthen in the future [11]. The smartphone is the perfect companion because it is personal and it is ubiquitous. It will provide the possibility of interacting directly with the user asking to enter some information or in a discreet and implicit manner allowing monitoring the user activity. The sensed data referring to the parameters that concern the selected characteristics modeled for the tailoring will be updated constantly in the Virtual Individual Model. Moreover, with the many connection possibilities, the smartphone can allow accessing the information stored in the cloud and can connect to other devices, such as wearable accessories that can improve the physical activity monitoring. Since it is ubiquitous, it can always provide the appropriate trigger, as tailored messages, to influence the user's behavior. This is very important, since Fogg observed that "without an appropriate trigger, behavior will not occur even if both motivation and ability are high" [6]. Moreover, the many sensors integrated in the smartphone allow capturing the contextual information, which can help to generate the trigger at the opportune moment maximizing its effectiveness. Moreover, the smartphone allows installing many applications as media services and games that will

motivate the teenagers to interact with the system. The mobile game will be designed to promote physical exercise. The integration with social networks will add the social aspect of the users' life to the parameters for the tailoring of the interventions and, most importantly, the social factor represents a very effective motivator. Another mobile application will be a sort of personal food diary, where the user will be able to note his/her alimentary behavior. This diary will help to understand the alimentary behavior of the user in order to provide the right feedback. For example, some data suggest that breakfast consumption is associated with higher intakes of micronutrients, fruit and vegetables and less frequent use of soft drink [12]. This means that the breakfast consumption habit can help to adopt a healthy dietary behavior. The diary allows following this behavior and to intervene through an alarm in order to remind to the teenager to have breakfast. The eating behavior is not only related to homeostatic reasons. In fact, an important factor that influences people's need and choice of food is represented by the emotional state [13]. The diary will allow noting also the mood in order to include the emotional state in the recognition of behavioral patterns. In fact, this information can be used to find some specific behavioral pattern related to emotional eating in order to generate the best intervention.

The introduction of biological models empowering technological actions for the promotion of citizen's health and well-being is considered to lead to a higher user centricity producing a more individualized strategy of health management and a stronger empowerment and engagement of the user [14].

The concepts supporting the PEGASO Virtual Individual Model (VIM) stem from a background of European experiences based on the vision of the individual as a unique multiple organ system, overtaking the traditional approach - in force in the medical practice - of the human body as a set of independent sections. The current modelling of a living human, such as the Physiome model [8], relies on different body functions incorporating knowledge from several biological disciplines and converges into a holistic integrative architecture. However, such an approach does not account for the behavioral and social externalities, which are known to interfere with and determine the biological balance of functions in health and disease.

By contrast, the PEGASO VIM's ambition is to provide a definition of individual's characteristics relevant for the condition of overweight/obesity, including both biological specifications and alimentary and exercise behavior factors, along with their psycho-social drives, specifically analyzed for young people, in a defined age range of 13-16. To this aim, the contribution of experts from different Europe countries has been integrated into a comprehensive view joining competences including medical, exercise physiology and nutritional knowledge, together with psycho-social expertise.

The PEGASO VIM overall structure is depicted in Figure 6 and considers the individual's health and well-being as resulting from the balance between components of physical, functional and psycho-social domains, according to the World Health Organization founding definition [15].

The model is based on elements relevant for overweight/obesity among adolescents. It considers health and well-being as primarily settled on a balance between physical body structure, body functionalities and psycho-social factors which, on turn, influence dietary and exercise behaviors and their possible changes. Alimentary and physical activity habits are key life-styles for preserving good

physical conditions and a well-functioning organism.

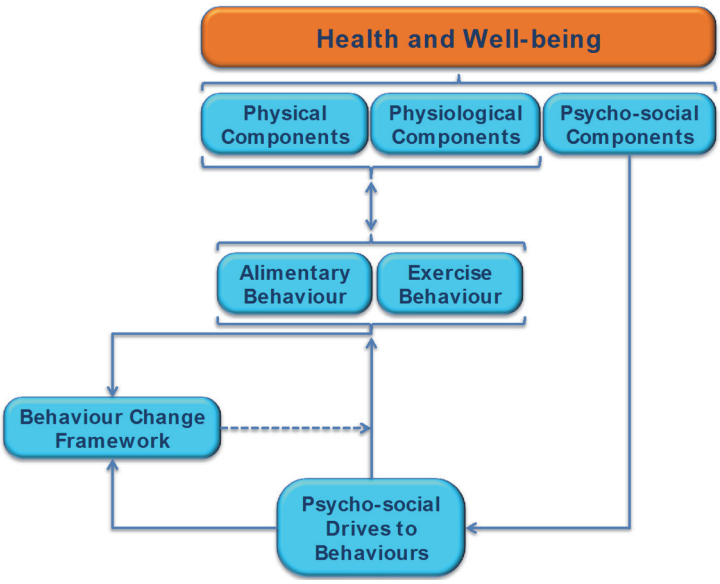


Fig. 6. Overall PEGASO Virtual Individual Model.

Dietary and physical activity behaviors are important life-styles affecting health and well-being through the preservation of optimal physical conditions and functional order. The model thus involves also concepts related to behavioral and psycho-social domains which are only partially covered by current vocabularies and ontologies. Since healthy life-style awareness, motivation and engagement among adolescents are the principal objectives of the PEGASO system, a particular emphasis is given in VIM to the aspects related to behavior change strategies and their relation with psycho-social components driving alimentary and exercise behaviors, especially if liable to be modified.

It is important to consider the psychosocial aspects of overweight and obesity in adolescents because problems associated to these aspects (E.g. peer pressure, low confidence, low self-esteem, depression and attitudes towards food) are likely to have an impact upon their lives and contribute towards differences in physical activity patterns and weight fluctuations. The following sections describe the social and psychological aspects of obesity in adolescents, using a behavioral model based on the analysis of Capability, Opportunity and Motivation as main components of behavior (COM-B). Where relevant in relation to the target behaviors selected for PEGASO adolescents, examples will be given with details.

At the core of the Virtual Individual Model analysis of behaviors is a psychological model of human behavior incorporating the psychological components associated with behavior change. COM-B components are applicable to all human behaviors and are the starting point for developing new behavior change interventions. Each component in the COM-B model directly influences behavior, and interacts with the other components. In combination they can provide the rationale for why the target behavior is not engaged in, and this then identifies the appropriate

components to be addressed to bring about a change in that behavior. In this way all the components of the COM-B model are interdependent, and work in unison to help change a target behavior, or support the maintenance of a target behavior once an individual has adopted it into their regular pattern of behavior. Each component of the COM-B model is divided into sub-components which are used to capture the more refined details of the COM-B components that are specific to the target behavior.

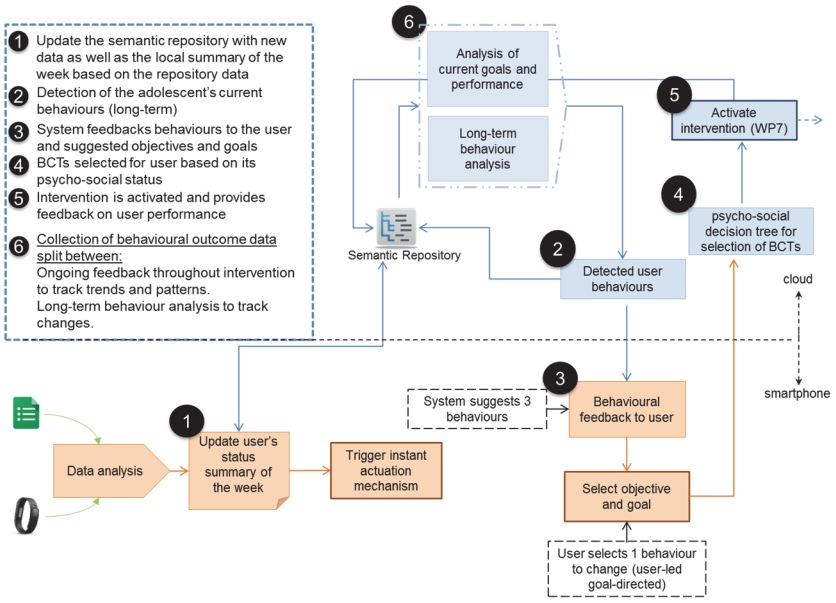


Fig. 7. The functional flow-chart of the PEGASO intervention.

Designing the PEGASO products and service systems the co-design methodology was adopted as key strategy. This is relevant for the specific target population to achieve high acceptance and compliance.

Smartphone, Wearable sensors and Multimedia Diaries have been categorized as sensors of the PEGASO system, and their acceptance and use by the end-user is the first essential requirement for the project success. Together with the sensors two other main fundamental categories constitute the PEGASO system architecture: the Social level of the platform, and the gamification of teens' life.

In this project, Social Network is intended as the creation of a social community that shares the same objective, i.e. healthy lifestyle habits. The End-User requirements aimed at retrieving opinion about the Information Sharing, in particular what teenagers want to share, people they'd prefer to be connected with, how they would like to share those information (feedback system and notification)

Finally, the Social platform represents also the gamification of teens'life and the role of experts is to monitor results from physiological data and to send specific feedback information on the health status of the user through the Health Companion.

Schools play a particularly critical role by establishing a safe and supportive environment with policies and practices that support healthy behaviors. Schools also provide opportunities for students to learn about and practice healthy eating and

physical activity behaviors. For this reason we involved and collaborated with schools in organizing the focus groups and recruiting teenagers in the user requirements definition phase. 16 focus groups in three nations were carried out with about 200 teenagers to gather this information and to iteratively design services and devices, according to the identified directions of PEGASO interventions.

Fig. 8. PEGASO interventions assets.

Fig. 9. The prototypes of garments and wearable devices.

This is a sort of life compass that the PEGASO project aims at helping teenagers to build, through a positive message and tool dedicated to teenagers, for improving their lives. The identified 4 pillars/directions are: a) move, i.e. the adoption of an active lifestyle, b) play, i.e. the serious game sustains engagement and motivation in the choice, c) eat, i.e. food education and pleasure, and d) share, a community and social approach to mutually engage in the actions.

The initial systems have been designed and prototyped and are currently undergoing pre-pilot tests.

A key feature is the PEGASO companion. The smart companion is a Personal Digital "Friend" acting as a daily-life guide for Coaching, Caring for, and Empowering teenagers in their activities toward healthy habits. From this definition the companion has multiple facets fostering behavior change.

- Digital: the companion exists in the smartphone
- Personal: the companion is customized to the single user
- Friend: the companion would establish an affective relationship with the user
- Daily-life guide: the companion accompanies the user (coaching, caring and empowering) during her daily activity
- Toward healthy habits: supporting behavior change to promote healthy lifestyles is the main goal of the companion and PEGASO project in general

Figure 10 highlights the most relevant services that the companion, represented by the circle in the center, should offer to the user in relation to the COM-B model.

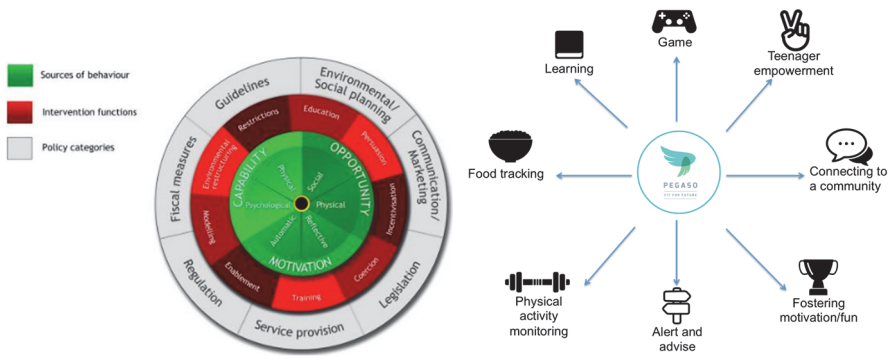


Fig. 10. The reference COM-B model and related services for Companion design.

The main goal of the Companion is to support teenagers' behavior change increasing Capacity, Opportunity and Motivation to achieve a set of target behaviors.

As we can see from the previous definition, the COM-B model of behavior is at the base of the design of the companion. In particular, the target behaviors have been selected in collaboration with PEGASO experts and taking into account the results of the focus groups. The selected target behaviors are:

- Physical activity: 10.0000 steps, 60 minutes of physical activity per day, Community sports teams, School sports teams, Screen viewing duration, Sleep duration & quality, Walk or cycle to school;
- Alimentation: Breakfast eating, Fruit consumption, Vegetable consumption, Sugar sweetened drinks.

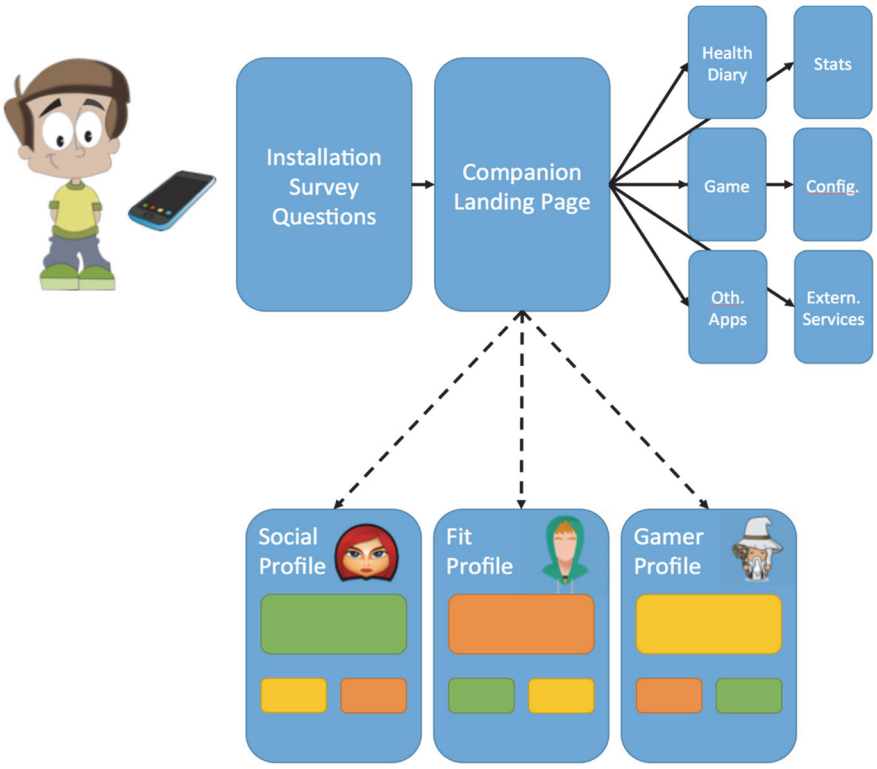


Fig. 11. The first conceptualization of the PEGASO Companion.

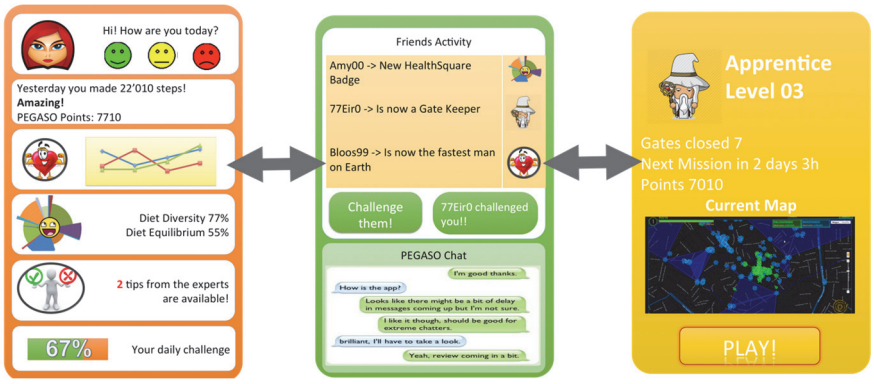


Fig. 12. The first draft PEGASO Companion.

4 Conclusions and Expected Impact

PEGASO believes that we are at a key turning point in the history of the Internet. Convergence of major trends is occurring which is driving changes in people behavior and expectations. These trends include the exponential rise in use of smartphones and

tablets, increased Internet access speeds, new business models driven by online commerce and app stores, the impact of social online communication, and software delivery transitioning from prior PC/internet models to cloud-based services accessed with touch-based devices (smartphones and media tablets).

With more than five billion mobile users worldwide and a massive global network, for the first time in history mobility is attracting significant attention among the healthcare and life sciences community.

Integrating mobility, gamification and life science has the potential to motivate individuals to adopt healthy lifestyles, with personalization techniques and incentives that will be delivered through the PEGASO system.

The following key issues however have to be adequately addressed and are at the center of the PEGASO rationale:

- Knowledge of how to stay healthy is ubiquitous; however, obesity and lifestyle-related illness are still among the top healthcare challenges in Europe.
- Although clinical content and health information have been available through the Internet for years, there has been no improvement in overall health in Europe.
- Motivating individuals to change behavior is not just a clinical issue. Successful programs include incentives along with personalized programs and, increasingly, the inclusion of behavioral science.
- Gamification has emerged as a recognizable trend that can have a significant positive impact on all businesses and is yet to be leveraged by wellness and healthcare.
- Because gamification's goal is to change human behaviors, PEGASO as a wellness player will confront both opportunities and risks, requiring a clear vision of the value of the system.

The strategy proposed by PEGASO, based on behavior changes, is expected to have favorable effects in reducing overweight/obesity and associated diseases and social costs in proportion to the national prevalence of body mass excess in this age class taking advantage of the possibilities offered by innovative ICT and of teens' affection to mobile and social network.

4.1 Enhancing Self-awareness of Younger People for Health Issues, by Means of Inclusive Approach Integrating Individuals into a Community Context, and Promoting Behavioral Changes in Favour of Physical Activity and Healthy Diets

Recommendations specific for children and young people have been released in 2005 also by the Commission of European Communities in the Green Paper on healthy diets and physical activity for the prevention of obesity among EU citizen [16]. Reducing the risks deriving from unhealthy diets and physical inactivity and increasing awareness and understanding of the influences of diet and physical activity on health are the core objectives of the global strategy dictated by the World Health Organization (WHO) against non-communicable disease and changes in behavior has been indicated by WHO among the **outcome indicators** for assessing actions fostering such a global strategy [17, 18]. The relationship between diet, physical

activity and health is based on strong scientific evidence. Studies using motion sensors have shown that children who spend less time in physical activity are at higher risk to become obese during childhood and adolescence [19, 20]. Television and video games contribute to more sedentary leisure activities and are reported to favour concurrent consumption of energy-dense snacks and beverages [21]. In the greatest majority, behavioral aspects related to lifestyle and diet (with a relevant interference of socioeconomic factors) play an important causative role. The findings from the 2009/2010 survey in EU countries from Health Behavior in School-aged Children (HBSC) international report indicate that young people who are overweight/obese are more likely to exhibit unhealthy alimentary patterns, are less physically active and watch television more [22].

4.2 Preventing Juvenile Overweight/Obesity and Reducing Morbidities Associated to Juvenile Overweight/Obesity in the Short Time, and Long Term Health Consequences, including Adult Obesity, and Associated Medical, Social and Personal Costs

According with the 2007 report of the EU Public Health Program Project "Global Report on the Status of Health in the European Union - EUGLOREH", the number of EU children affected by overweight and obesity is estimated to be rising by more than 400,000 a year, adding to the over 14 million of the EU population who are already overweight (including at least 3 million obese children) [23]. Overall, across the entire EU, overweight affects almost 1 out of 4 school age children/adolescents. Childhood obesity has physical, psychosocial and economic consequences. Overweight and obesity in the juvenile age are associated with a number of serious medical conditions such as sleep-related breathing disorders and asthma, fatty liver disease, poor glucose tolerance, increased risk of non-insulin-dependent diabetes, hypertension, and can lead to increased rates of non-communicable disease in adulthood such as cerebro-vascular disease, diabetes, certain types of cancer, osteoarthritis, gall bladder and endocrine disorders, in relation with the high probability for obese children to become obese adults [24]. Recent estimates indicate that more than 27 000 children in the EU have type 2 diabetes and more than 400.000 have impaired glucose tolerance. About 1.1 million suffer from hypertension and 1.2 million suffer from metabolic syndrome (consisting in a clustering of hypertension, abnormalities of lipid metabolism, and raised blood glucose levels)[23]. Moreover, body mass excess in the juvenile age leads to low self-esteem, depression and social exclusion, with immediate consequences in the psychological and social domain, possibly leading also to concurrent or subsequent psychiatric pathology. Specifically, social exclusion in childhood has been associated with reduced psychological functioning in adulthood [19].

4.3 Reducing Medical, Social and Personal Costs Associated to Juvenile and Adult Overweight/Obesity

The economic impact of overweight and obesity on health care and social systems is definitely sizable. Such an economic burden has been recognized in terms of direct

medical costs (including preventive, diagnostic, and treatment services related to obesity), indirect costs (attributable the value of income lost from decreased productivity, restricted activity, absenteeism, and bed days), and intangible costs (which can be referred to as the social and personal costs or losses associated with obesity) [25, 26]. In recent years, health costs deriving from obesity among adults in different Western European countries have been conservatively estimated to amount up to over 10 billion Euros with relative economic burdens ranging from 0.20% to 0.61% of each country's gross domestic product [25]. A fairly higher economic load, although scarcely definable with accuracy, can be considered to arise from the impact of obesity on the wider economy in terms of loss of productivity [26]. It is even more difficult to estimate the overall cost of obesity among children and young people. In the perspective of health care costs, in fact, it has been assessed that overweight and obese children exhibit significantly higher expenditures for outpatient visits, drug prescriptions, and emergency room admissions, and have a higher probability of being high utilizers of health care services [27]. Additionally, in the personal context, it has determined by different studies that individuals who were obese, as adolescents become adults with lower educational attainment, earning less money, experiencing higher rates of poverty and having a lower likelihood of marriage, compared with thinner peers [28]. Furthermore, as childhood obesity is a risk factor for adult diseases, mortality costs in terms of the value of future income lost by premature death in adulthood, should be considered.

4.4 Developing a System Suitable for Interventions based on Equity and Inclusivity

The PEGASO system, integrating state of the art technologies within an holistic approach including social and human aspects, fully complies with WHO guidelines indicating the priority of comprehensive and coordinated multiple-strategy interventions across the whole population promoting behavioral changes in favor of physical activity and healthy diets in order ensue an effective obesity prevention in childhood [4]. An increased prevalence of overweight/obesity is also significantly associated with low family affluence [23]. The principles of the intervention included in the present proposal, based on a "virtual individual" model considering as basic features also social status and social behaviors, are in accordance with the main outcomes from the 2009/2010 HBSC report recommending to address not only health and health behavior issues, but also the social context in which young people live, providing equal opportunities for all [22].

4.5 Providing a Transnational Opportunity for a Coordinated Effort to Tackle a Transnational Issue

The transnational relevance of the growing prevalence of overweight/obesity among younger population in industrialized and developing countries worldwide, prompt also to an EU co-ordinated effort in research and industrial development to face such an epidemic in member countries, which is another main feature entailed in the present proposal. Indeed, the promotion of research for the prevention and control of

non-communicable diseases (which include overweight/obesity) is one of the main objectives recently focused by the WHO 2008/2013 Action Plan, proposing for international partners the action of work jointly on "research on socioeconomic determinants, lifestyle and behavior modification as well as community-based interventions" [29].

Acknowledgements. This paper is based on the discussions and on the material produced during the preparation and the activities of the PEGASO project to which all project partners have contributed. The PEGASO project is co-funded by the European Commission under the 7th Framework Programme. PEGASO is part of the cluster of projects in the ICT for health area, it has started in December 2013 and will run for 42 months.

The authors wish to thank all the project partners for their contribution to this work.

References

1. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. Journal of the American Medical Association 2014; 311(8): 806-814.
2. National Center for Health Statistics. Health, United States, 2011: With Special Features on Socioeconomic Status and Health. Hyattsville, MD; U.S. Department of Health and Human Services; 2012.)
3. Visions for Horizon 2020 - Copenhagen Research Forum, 2012.
4. World Health Organization. Population-based prevention strategies for childhood obesity: report of a WHO forum and technical meeting, Geneva, 15–17. December 2009.
5. Sanders, E. B., (2002). From User-Centered To Participatory Design Approaches. Design, 1-7. Taylor &Francis.
6. Fogg, B.J., (2009). A Behavior Model for Persuasive Design, Proceedings of the 4th International Conference on Persuasive Technology Persuasive 09, ISBN: 9781605583761.
7. Caon M, Carrino S, Guarnieri R, Andreoni G, Lafortuna CL, Abou Khaled O, and Mugellini E. A Persuasive System for Obesity Prevention in Teenagers: a Concept. Second International Workshop on Behavior Change Support Systems (BCSS 2014) pp 17-20, http://ceur-ws.org/Vol-1153/Paper_2.pdf)
8. Fenner JW, Brook B, Clapworthy G, Coveney PV, Feipel V, Gregersen H, Hose DR, Kohl P, Lawford P, McCormack KM, Pinney D, Thomas SR, Van Sint Jan S, Waters S, Viceconti M. The EuroPhysiome, STEP and a roadmap for the virtual physiological human. Philos Trans A Math Phys Eng Sci. 2008; 366: 2979-2999.
9. Brug J., Oenema A., and Campbell M., Past, present, and future of computer-tailored nutrition education., Am. J. Clin. Nutr., vol. 77, no. 4 Suppl, p. 1028S–1034S, Apr. 2003.
10. den Akker HO, Moualed LS, Jones VM, and Hermens HJ, A self-learning personalized feedback agent for motivating physical activity, Proc. 4th Int. Symp. Appl. Sci. Biomed. Commun. Technol. - ISABEL '11, pp. 1–5, 2011.
11. Carrino, S., Caon, M., Abou Khaled, O., Andreoni, G., Mugellini, E. (2014). PEGASO: Towards a Life Companion, in the Proceedings of the 16th International Conference on Human-Computer Interaction, 22 - 27 June 2014, Heraklion, Crete, Greece.
12. Merten MJ, Williams AL, Shriver LH. Breakfast consumption in adolescence and young adulthood: parental presence, community context, and obesity. Journal of the American Dietetic Association, 2009, 109(8):1384-1391
13. C. Science, A. Kapoor, P. Johns, K. Rowan, E. A. Carroll, M. Czerwinski, and A.

Roseway, “Food and Mood : Just-in-Time Support for Emotional Eating,” ACII2013, pp. 252–257, 2013.
14. Honka A, Kaipainen K, Hietala H, Saranummi N. Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health. IEEE Rev Biomed Eng. 2011; 4 :119-139.
15. World Health Organization, 1948. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948. Available at: http://www.who.int/about/definition/en/print.html.
16. Commission of European Community. Green Paper - Promoting healthy diets and physical activity: a European dimension for the prevention of overweight, obesity and chronic diseases. COM(2005) 637, Brussels, December 2005.
17. World Health Organization. Global Strategy on Diet, Physical Activity and Health, Geneva, 2004.
18. World Health Organization. Global Strategy on Diet, Physical Activity and Health. A framework to monitor and evaluate implementation, Geneva, 2006.
19. Consensus statement: Childhood Obesity. J Clin Endocrinol Metab. 2005; 90: 1871-87.
20. Lobstein T et al. Obesity in children and young people: a crisis in public health. Obes Rev. 2004; 5 Suppl 1:4-104.
21. Rey-López JP et al. Food and drink intake during television viewing in adolescents: the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) study. Public Health Nutr. 2011;14(9):1563-9.
22. Currie C et al. eds. Social determinants of health and well-being among young people. Health Behavior in School-aged Children (HBSC) study: international report from the 2009/2010 survey. Copenhagen, WHO Regional Office for Europe, 2012 (Health Policy for Children and Adolescents, No. 6).
23. EUGLOREH2007 project The Status of Health in the European Union: Towards a Healthier Europe, available at: http://euglorehcd.eulogos.it
24. World Health Organization. Obesity: Preventing and Managing the Global Epidemic. Report of a WHO Consultation. WHO Technical Report Series no. 894. WHO: Geneva, 2000.
25. Müller-Riemenschneider F et al. Health-economic burden of obesity in Europe. Eur J Epidemiol. 2008; 23: 499-509.
26. Dent M. The economic burden of obesity. Oxford: National Obesity Observatory, 2010.
27. Trasande L & Chatterjee S. The impact of obesity on health service utilization and costs in childhood. Obesity (Silver Spring). 2009; 17: 1749-54.
28. Gortmaker SL et al. Social and economic consequences of overweight in adolescence and young adulthood. N Engl J Med 1993; 329: 1008–1012.
29. World Health Organization. 2008-2013 Action plan for the global strategy for the prevention and control of noncommunicable diseases. Geneva, 2008.

The TRAMA Project

Manuela Galli[1,2], Veronica Cimolin[1], Chiara Rigoldi[1] and Marcello Crivellini[1]

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, P.zza Leonardo Da Vinci 32, 20133 Milano, Italy
[2] IRCCS "San Raffaele Pisana", Tosinvest Sanitá, via della Pisana, Rome, Italy
{manuela.galli, veronica.cimolin, chiara.rigoldi, marcello.crivellini}@polimi.it

Abstract. The Project TRAMA Network was born with the aim of training researchers in the field of quantitative analysis of movement creating a Network among all the Institutions involved in the Project. The aims of the Project were: a) Training about the use of the equipment usually present in a Motion Analysis Lab, about the use of new experimental set-ups for movement analysis and about the development of new methodologies; b) Training via web and in practical session in the Labs, about the evaluation of the data acquired in motion analysis Labs and their clinical meaning; c) Network realization for a continue teaching assistance within different Institutions involved in the Project. TRAMA Project was one of the project of Programme αLFA (America Latina - Formacion Academica), a programme of co-operation between higher education institution of the European Union and Latin America.

1 Introduction

Motion Analysis (MA), or computerized multifactorial and integrated analysis of human movement, is a rapidly expanding field of considerable interest from a clinical perspective: the study of postural and motor changes in patients with movement disorders can yield, in fact, crucial information in establishing the degree of functional limitation, associated to a specific pathology, in identifying the rehabilitative program specific for a particular patient, and in following its evolution over time. Furthermore, posture and motion assessment can provide elements important for the evaluation of the effectiveness of rehabilitative programs aimed at reducing the functional limitation due to pathology.

In clinical settings, the movement evaluation is generally conducted using video recording; however, this method has some limits, as it is only able to supply a qualitative description of the movement. This approach is sufficient to evidence gross abnormalities in movement; however, as functional limitation and movement complexity increase with organic pathology, objective analysis becomes necessary. The availability of MA with innovative techniques and advanced equipment for the description, quantification and evaluation of motion achieves precisely this objective.

MA is in fact able to supply clinicians with quantitative, non-invasive, three-dimensi -onal information relating both to kinematic and kinetic aspects of motion and to the pattern of muscle activation during movement. Thanks to these features, MA has a great

role in clinical applications: in fact, the quantitative assessment of a patient's movement provides crucial information of functional limitation related to the pathology. It provides elements useful for the identification of rehabilitative and rehabilitative programs and for the evaluation and monitoring of their effects over time, too.

The great importance of MA in clinical centres is demonstrated by the increasing number of Motion Analysis Labs (MALs), placed in clinical settings; in recent years, a large number of clinical centres, especially those involved in rehabilitation, have set up MALs and carried out motion evaluations in different pathological situations, both in Europe (EU) and in Latin America (LA).

Even though MA is a powerful tool in a clinical environment, its use requires the clinicians and the operators working in a MAL to solve from a practical point of view a lot of obstacles, which sometime limit the routine use of this methodology. Standardized experimental sets-up, data representation and common data evaluation are more and more required.

For this reason, a shared operational methodology is necessary to allow the MA to spread throughout clinical centres and to overcome the existing difficulties related to the use of different instruments and work practices. Furthermore, key managerial and organizational skills are required to assure sustainability of MALs service role to clinical centres. Competencies needed to manage and operate them effectively require the establishment of devoted training programs.

In this panorama, our idea has been the realization of an international network aimed to:

1. the training of specialized personnel able to operate in MALs
2. the exchange of data and methodologies to establish standardized working practices.

From this idea the TRAMA (TRAining in Motion Analysis) Project has been thought. TRAMA Project was an international project, approved and financed by the European Community in the field of Programme Alpha (a programme of co-operation between higher education institutions of the European Union and Latin America; `http://ec.europa.eu/europeaid/where/latin-america/regional-cooperation/alfa/index_en.htm`), lasting three years (from 2007 to 2010).

2 Participants

The Coordinator of the project has been the Bioengineering Department (now Department of Electronics, Information and Bioengineering) of Politecnico di Milano (Italy) (Fig. 1) and partners from EU (Italy, Sweden and Belgium) and from LA (Chile, Colombia and Mexico) have been involved, both Higher Educational Centers (Full Partners) and Clinical Centers (Associate Partners) (Fig. 2).

This international network has worked during these three years in order to train researchers about the use of the equipment of MALs, the clinical use of MA data, the use of new experimental technologies for movement analysis and their transferability in clinical field and development of new methodologies. In particular, the Full Partners

Fig. 1. Politecnico di Milano.

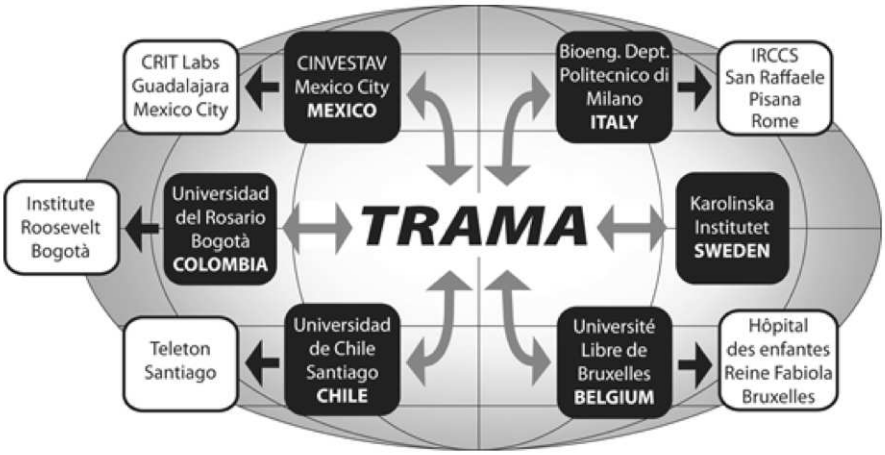


Fig. 2. Partners of the TRAMA project. In grey the Full Partners and in white the Associate Partners are highlighted.

with valuable experiences in MA have shared protocols and practices with the Associate Partners already utilizing MALs. On the other hand, Associate Partners used the protocols and the learned concept in their daily clinical practice, improving their clinical services.

The partner involved in the Project are described following, country by country.

2.1 Italy

POLITECNICO DI MILANO

The Politecnico di Milano was established in 1863 by a group of scholars and entrepreneurs belonging to prominent Milanese families. Its most eminent professors over the years have included the mathematician Francesco Brioschi (its first Director), Luigi Cremona, and Giulio Natta (Nobel Prize in Chemistry in 1963). The Politecnico di Milano is now ranked as one of the most outstanding European universities in Engineering, Architecture and Industrial Design, and in many disciplines is regarded as a leading research institution worldwide. In Italy, the term ”Politecnico” means a state university consisting only of study programmes in Engineering and Architecture. The Politecnico di Milano is nowadays organized in 12 departments and a network of 6 Schools of Engineering, Architecture and Industrial Design spread over 7 campuses over the Lombardy region with a central administration and management. The 6 schools are devoted to education whereas the 12 departments are devoted to research.

The number of students enrolled in all campuses is approximately 40.000, which makes the Politecnico di Milano the largest institution in Italy for Engineering, Architecture and Industrial Design.

Inside the Politecnico di Milano, the Department of Electronics, Information and Bioengineering is present. The mission of Bioengineering section (ex Bioengineering Department) is to progress the knowledge of biomedical engineering through the multidisciplinary research, starting from the molecular and cellular level up to the complex living organism, aiming at the design, realization and optimization of devices, equipment and systems in the studying of different physiological and clinical aspects for diagnosis, therapy and rehabilitation. Theoretical and practical contributions are also intended to be dedicated towards the structures and services involved in the management of health and environment. Further, it constitutes the coordination of the intellectual and material resources of Polytechnic University in Milano for developing and providing didactical and training activity at the level of Bachelor (3 year track), Master Degree (overall 5 year track), Master Courses (generally 1 year track), PhD Courses (3 year post-graduate course) and continuous training activity to students and professionals in biomedical engineering, in other areas of engineering studies, in biology, medicine and living sciences. Finally, the section fulfils the task of making available proper methods, tools and knowledge to hospitals and private and public health organizations, both at a national and international level, in technical supporting systems, advisory, consulting, research and development as well as transfer of innovative products, systems and technologies.

IRCCS ”SAN RAFFAELE PISANA”, TOSINVEST SANITÁ, ROMA

IRCCS ”San Raffaele Pisana” belongs to a select and distinguished circle of Institutes (Tosinvest Sanitá) which are highly specialized and represent a reference point at the national level, capable of providing treatment to patients suffering from any type of disabilities.

Thanks to the intense and distinguished clinical and research work it is carried out for years in the area of the rehabilitation, it has been recognized by the Health Ministry as a Scientific Institute for Research, Hospitalization and health Care (IRCCS).

In a building surrounded by greenery and provided with 298 beds for inpatients as well as outpatients, the following working units operate:

- cardiologic rehabilitation
- neuro-motor rehabilitation
- otolaryngological rehabilitation (for hearing, balance, voice, speech and deglutition disorders)
- paediatric rehabilitation - Centre for Child Development
- respiratory rehabilitation
- internal medicine

The clinical and research activities are supported by the presence of the following laboratories and services:

- Neuro-physiopathology laboratory
- Gait Analysis laboratory
- Cardiac functionality laboratory
- Audiology, vestibology, speech therapy and deglutition laboratory
- Respiratory functionality laboratory
- Clinical pathology lab oratory
- Hydrokinesitherapy service
- Occupational therapy service
- Functional educational service
- Diagnostic imaging service

The IRCCS ”San Raffaele Pisana” carries out intense and distinguished research work in the area of rehabilitation and, more in general, of Neuroscience. For this reason, it has equipped itself with a modern Research centre with clinical and basic research laboratories which avail themselves of the most recent technologies and of the collaboration of numerous Italian as well as foreign researchers.

In particular, the Child Adult Aging Development Centre turns to children, teenagers and adults with cognitive retard, motor coordination problems and behaviour-learning difficulties. These symptoms could be caused by different inborn factors (i.e. genetic syndrome) or acquired factors that act in pre, peri or post-natal age: patients are followed in a longitudinal prospective, from childhood to adulthood during aging.

The Child Adult Aging Development Centre has a multidisciplinary staff dedicated to the evaluation and spotting of a diagnosis and to the elaboration of a multilevel rehabilitation program that involves therapists, psychologists, social assistants and caregivers.

Inside this aim, Gait Analysis lab plays a fundamental role giving an important help in decision-making process in term of quantitative movement analysis that, together with biomechanical data, studies neuro-physiological parameters during movement.

2.2 Sweden

THE KAROLINSKA INSTITUTET, STOCKHOLM

Karolinska Institutet (KI) is a leading medical University, dedicated to improve people’s health through research and higher education. It developed from a school of army surgeons in 1810 to a medical university, celebrating 200 years anniversary 2010.

KIs mission statement is ”to be Europe’s leading medical university and the Nordic region’s foremost innovation centre in the life sciences, and as such, it comprises an important driving force for the development of the country and the Stockholm region”. According to an international ranking in December 2009, KI is the 30iest best university in the world.

Karolinska Institutet has two main campuses, one in Solna and one in Huddinge.

A considerable amount of teaching and research is also carried out on other sites in Stockholm in collaboration with the Stockholm County Council and the health care sector. This includes primary health care facilities and the main hospitals: Karolinska University Hospital in Solna and Huddinge, Danderyd Hospital, Söder Hospital, St Göran’s Hospital and St Erik’s Eye Hospital. Some courses are also run in cooperation with Stockholm University, The Royal Institute of Technology (KTH) and Södertörn University College. Karolinska Institutet has two Science Parks, one at Campus Solna and one at Campus Huddinge in Flemingsberg.

In keeping with Alfred Nobel’s testament, the Nobel Assemble at Karolinska Institutet selects the winner of the Nobel Prize in Physiology or Medicine. Actually, five of eight Swedish Nobel Prize Laureates in Physiology or Medicine are from Karolinska Institutet.

Research and education bridging from molecule to patient is carried out in 22 departments with 9 research fields such as Cancer, Circulation and respiration, Infection, Inflammation and immunology, Neuroscience, Public and international health, Reproduction, growth and development, and Tissue and motion. In 2008 researchers at KI published 3.000 original articles and 1.000 other publications. External research funding accounts for 80 % of Karolinska Institutet’s total income. There are 3.600 employees and 600 research groups with 1.500 researchers/university teachers. About 2.100 PhD students are enrolled at the different departments and 5.300 students are enrolled in higher education at KI stretching from undergraduate programs (Bachelor level), Advanced programs (Master level), Specialist nursing programs, Single-subject courses and contract education. Karolinska Institutet has developed an integrated infrastructure for health care, education, research and development with the Stockholm County with the goal of increasing the caregivers’ competence and reducing the time from experimental discovery to clinical application.

The University library is the largest medical library in the Nordic countries and has premises in Huddinge and Solna. The mission of the university library is to support scientific communication, support the learning process, manage scientific information resources, and provide a forum for study, dialogue and networking. The library is visited by an average of 3.000 people a day and provides access to over 10.000 journals and periodicals, about 100 databases and a large number of e-books.

2.3 Belgium

THE UNIVESITÉ LIBRE DE BRUXELLES (ULB)

The city of Brussels is the capital of a federal state which has three different administrative regions based on language and has been at the heart of Europe since 1957. It naturally follows that the city should have a university in keeping with its standing and

the ULB, with its 21.000 students, 29 % of whom come from abroad, and its very cosmopolitan body of staff, is an intrinsically international institution open to both Europe and to the whole world.

It was at the heart of the creation of a network of major universities from different European capitals - UNICA - and is involved in international programs for research and development and for mobility. ULB is a multicultural institution, which has 8 faculties and a range of schools and institutes and is, at the same time, a comprehensive university providing academic tuition in all disciplines and study cycles. With its 3 Nobel Prize winners, a Fields medal, three Wolf Prize, two Marie Curie Prizes and 29 % of the Francqui prizes awarded, the university is also a major research centre which is recognized by the academic community the world over. Nor does it shirk its social, societal and scientific commitments, which it meets through combining broad access to higher education with excellent quality research and through its role in furthering economic development in the regions where it is located (Brussels and Wallonia). ULB also has a teaching hospital - Hôpital Erasme, a specialist institute for studying cancer - Institut Bordet, and an extensive hospital network.

For about a decade now the university has been actively involved in maximizing research potential in both Brussels and Charleroi, where it has set up a biotechnology park around its renowned Institute for Biology and Molecular Medicine (IBMM) & Institute of Medical Immunology (IMI) In terms of partnerships, it is part of the Alliance for Higher Education and Research, together with the Mons University and, in conjunction with 5 Hautes écoles, the Royal military school, 2 institutes for architectural studies and 2 colleges for fine arts, it also makes up the Brussels partners of the Alliance. As a private university, which is recognized and subsidized by the Belgian authorities, ULB receives government funding today to the tune of 58 % of its overall budget. Founded on the principle of freethinking analysis that advocates independent reasoning and the rejection of dogma in all its forms, ULB has remained true to its original ideals - an institution free from any form of control which is committed to defending democratic humanist values, an approach it also extends to the way that it is run.

HOSPITAL UNIVERSITAIRE DES ENFANTS REINE FABIOLA (HUDERF), BRUXELLES

Inaugurated in 1986, the HUDERF is the only Belgian university hospital entirely reserved for children's medicine: all is conceived for them and for their parents. From birth to adolescence, the children receive there the most complete care in respect of the charter of hospitalized child's rights. As a medico-surgical hospital of 168 beds, the HUDERF accommodates more than 11.000 children per annum in hospitalization. The ambulatory sector (consultations and emergencies) is one of most important in Belgium and receives more than 100.000 patients per annum. HUDERF is also a public hospital (Brussels network IRIS) guaranteeing quality care and modern medicine accessible to all children. HUDERF is a reference centre for children with cerebral palsy. This centre is called "CIRICU". The goal of CIRICU is to optimize the follow-up of the children with cerebral palsy by elaborating an individualized treatment plan. It is necessary to have a good communication between the centre, the patient, his family and all the therapists in charge of the child. The intervention of CIRICU is organized in this way:

- Elaboration of a multi-disciplinary assessment, which is the base for the treatment plan;
- Regular evaluations and upgrade of this treatment plan;
- Occasional advises in the field of communication, adaptations . . .

2.4 Chile

UNIVERSIDAD DE CHILE, SANTIAGO DE CHILE
The University of Chile with more than three centuries of history is a Public University founded in 1842 as a continuation of the Universidad Real de San Felipe (1738). Its first Rector, Don Andrés Bello, Chilean-Venezuelan humanist, knew how to give a seal guarantor of classical culture, humanist and secular. The history of the University of Chile is parallel to the country's history. It has become progressively until our days, in one of the main and largest Universities in the country. In their classrooms have been formed most of Chilean Presidents, their National Awards and two Nobel Prizes we have had in our country.

Originally had five Faculties; currently has fourteen Faculties, four Institutes and a Clinical Hospital, covering all the areas of knowledge.

Its nearly 30.000 students are divided into Pregraduate, Magister and Doctor Degrees.

Ranks in first or second place in virtually all national and international rankings, according to the parameters used. It has the largest number of accredited doctoral programs in the country, in all disciplinary areas, currently training more than 900 doctors in 30 accredited programs.

It is the first University in research in our country, representing 37 % of the ISI index of the country and 40 % of research competitive funds.

Faculty of Medicine: The Faculty of Medicine creation goes back to the opening of the University of Chile, being one of the five Faculties that gave origin.

With about 5,200 pre-graduate students, has in our days eight careers in health area: Medicine, Nursing, Nutrition and Dietetics, Medical Technology, Physical Therapy, Speech therapy, Obstetrics Nurse and Occupational Therapy.

His extensive postgraduate activity is represented with five accredited doctoral programs, several Magister and more than 60 programs of Medical Specialties. Currently are now 1.100 physicians in training, representing 51 % of the country's medical training. It has the only specialist-training program in Physical Medicine and Rehabilitation of the country; since 1964 more than 150 physicians have acquired this specialty. Research is one of the largest institutional missions of the Faculty of Medicine that has the largest trajectory of research in the country. Currently has multiple work lines, numerous laboratories and a scientific productivity of front line.

Hospital Clinico Universidad de Chile: Founded in 1952, is the main University Hospital of the country. This is a highly complex hospital, with over 600 beds, 300 medical journey and 32 postgraduate programs, with 240 residents.

This place gives attention to 400.000 outpatients, 26.000 discharges and more than 23.000 surgeries every year. The University of Chile has no laboratory for motion analysis. For the quantitative evaluation studies the MAL of the Teletón centre is used, which

is described in the next paragraph. Since many years, a scientific collaboration is active between the University of Chile and Teletón.

INSTITUTO DE REHABILITACIÓN INFANTIL TELETÓN, SANTIAGO DE CHILE
The ”Instituto de Rehabilitación Infantil Teletón Chile (IRI Teletón Chile)”, is a non-profit Institution, for the rehabilitation of children and young people up to the age of 20, with motor disabilities. However, the upper age limit is 24 years for spinal cord injury patients and other traumatic acute injuries. It was founded in 1947 in Santiago Chile, and from 1978 it began to grow quickly due to large annual, televised fund raising event known as TELETÓN Campaigns. Now they are ten Institutes around the country and give medical rehabilitation assistance to around 26.000 (Dec 2008) patients throughout the country, with approximately 3.500 new patients each year and the 52 % of patients before 3 years old.

The most frequent diagnoses are Cerebral Palsy (9575 patients), Neuromuscular Diseases, Congenital Spinal Cord Injuries (Mielomeningocele) and Amputees. The socio-economic situation of our patients is 80 % low income families, 63 % are living in conditions of extreme poverty. During the last year (2008) gave 122.565 medical consultations and 961.339 therapeutic attentions.

The Mission is the Comprehensive Rehabilitation of children and young people with invalidating diseases. Our strong emphasis is on their independence and autonomy in order to improve their integration into the family, school, social and work environment: ”To Rehabilitate in order to Insert into the Community”.

The Future Vision is to be the leader in Chile in Comprehensive Rehabilitation. The strategic objectives are:

- Quality service
- Effective model of rehabilitation
- Modern and efficient administration
- To be an agent of change within the community
- Qualified human resources; continuous improvements policy performance management
- To maintain community support through the Teletón.

The main activity is the comprehensive rehabilitation, but work too, in academicals activities at pre graduate and post graduate levels, clinical researches and community activities. The therapeutic model is a Bio Psycho Social Model, with the followings programs:

- Medical Programs: Diagnosis and Treatments
- Psycho-Social Education Programs
- High Motivation Programs Arts and Sports

The aims are to achieve the maximum development of functional, physical, psychological, emotional and social abilities, independence, autonomy, and familial and social integration of our patients and, to establish support networks within the community.

2.5 Colombia

COLEGIO MAYOR DE NUESTRA SEÑORA DEL ROSARIO, BOGOTÁ
The Universidad del Rosario is an autonomous, private, secular, non-profit institution founded in 1653, accredited for its high quality and evaluated in 2006 for European University Association. Since then it has fulfilled the mission of forming individuals with a strong sense of responsibility in the benefit of society and has established five fundamental purposes: the integral, ethical and humanistic education; the academic requirements and academic quality; the investigation, the consolidation of the educative community; and the social responsibility.

The organizational framework of the university consists of the Board of Trustees (rector, vice-rector, counsellors, trustees, and the secretary general), Academic Board (deans of respective schools) and the Executive Board (Chancellery). The Board of Trustees and the Academic Board are responsible, with a participative model, for short, medium and long term planning; the Executive Board is in charge of applying the strategies and programs defined by said planning. The University has a staff close to 1.500 people and is financed mainly (73%) by student tuitions.

The University has seven Schools: Jurisprudence; Medicine and Health Sciences School; Natural Sciences, Economy; Political Science, Government and Foreign Affairs; Business; and Human Sciences. Each School is in charge of undergrad programs (22 in total), graduate programs (97 in total) and house research centers for 24 different workgroups.

In the present Integral Development Plan 2004-2015, the vision expresses three emphases: growth, which refers to the qualitative and quantitative development of the University; identity; strengthening the differentiating characteristics of the institution; and quality, understood on the basis of global, national and institutional referents and elaborated under the principle of university autonomy with social responsibility.

Since the middle of the XX century, the University has concentrated its activities in the social, human and health sciences. The present PID foresees the fact that the university should give priority to, strengthen and consolidate the existing facts, and, at the same time, opens up new action fronts in natural and exact sciences. For the beginning, these will help impelling the current programs, and in the next future, it will constitute options to broaden the offer of academic programs. This qualitative growth will lead to a more comprehensive realization of the ”being of the University”.

Concerning the investigation areas that the University emphasizes, each School works in the definition and consolidation of the priority and strategic areas of their own investigation groups and lines. Seen under a wider perspective, the most salient developments in investigation are found in jurisprudence, medicine and economy.

The Universidad del Rosario’s School of Medicine and Health Sciences, created the Health Sciences Research Centre to develop, to adapt and transfer new knowledge in the field of the health sciences with a commitment towards promoting, encompassing and developing research projects and looking for the resolution of high-priority health problems fulfilling the most demanding national and international research regulations regarding research with human beings.

The GiSCYT research group is attached to the Health Sciences Research Centre, the GiSCYT research group studied the problematic of health in work environments from

a dual approach: health and work. First, it is necessary to adopt an external perspective that explains the dynamics of this relation, integrated at the same time the point view that emerges from the logic of the worker, from the collective and the organization productive.

We can then consider what health and work results of a co-construction of two opposing logics: the first which refers to the pursuit of productive efficiency through the involvement of the workers (the logic of productivity) and the other hand the pursuit of health in the middle of changes the tasks and activities (the logic of work) which raises a questions concerning the consequences of this involvement of the worker. It is therefore necessary to know how the sense of the individual involvement in the pursuit of productive efficiency (requested) from exposure of every worker to the risks in the work (defined by each worker).

The GiSCYT research group, have the motion laboratory ”ergomotion” to apply the motion analyses in work environments. This is a research unit oriented to study of human movement in productive activity from an ergonomics and biomechanics perspective; the laboratory is responsible for producing models and protocols for the study of human movement in work activities. These models can contribute to development of theories about the action strategies for manage occupational hazards and also to design tools for prevention of lesions associated with human movement and demand intensive joint structures.

INSTITUTO DE ORTOPEDIA INFANTILE ROOSEVELT, BOGOTÁ

The Institute of Pediatric Orthopaedics Roosevelt is a hospital open to children since 50 years ago; at the beginning it was built to help children with orthopaedic diseases mainly poliomyelitis, a frequent disease at the middle of the last century. When poliomyelitis began to disappear as result of vaccination campaigns, the next in frequency neurological disease was cerebral palsy. Soon we notice that the knowledge about this disease and the results was not what we expected; eighteen years ago the study of cerebral palsy and of the gait disease produced by this pathology took our team to follow publications by Dr James Gage, who strongly recommended the use of gait laboratories to analyse these patients and to have diagnosis and a plan of treatment. The most interested physician at that time was Dr Camilo Turriago, who started, 20 years ago, to use gait analysis and multilevel orthopaedic surgery following Dr Gage concepts.

2.6 Mexico

CINVESTAV, ELECTRICAL ENGINEERING DEPARTMENT, BIOELECTRONICS, MEXICO

The Centre for Research and Advanced Studies of IPN (CINVESTAV) is a public organism dedicated to promoting, developing and teaching scientific investigation. The Institution counts 37 Academic Departments, separated in 4 areas of research: Exact Sciences, Biology and Medicine, Technology and Engineering and Social Sciences & Humanities.

CINVESTAV is integrated by 9 Centres across the country, offering 53 Academic programs and more than 500 research topics. All the Academic Programs are considered

by the Mexican National Council of Science and Technology with high level and twenty one programs are classified as competent at an International level.

CENTRO DE REHABILITACION INFANTIL TELETÓN, MEXICO CITY
The Children's Rehabilitation Center Teletón of the State of Mexico (CRIT-EM) provides care to disabled children carriers of neural-muscle-skeletal diseases. The main goal of the CRIT-EM is to offer an integral rehabilitation program for the patients of the centre, promoting their development and integration into the society. To achieve this goal the medical model of the CRIT-EM is based in the follow premises: prevention, interdisciplinary care and attention in a process, which includes the patient, family, school and social environments. Within this model, the use of high technology has had a prominent place as a tool for functional assessment and treatment.

The CRIT-EM is part of system rehabilitation centres of the Fundacion Teleton, which consists of 13 rehabilitation centres. The centres of Fundacion Teleton are located in different cities of the Mexican Republic and actually give attention to more of 18.200 children with ages ranging from 0 to 18 years old. Only the CRIT-EM provides services to 3.500 (19%) children.

CENTRO DE REHABILITACIÓN INFANTIL TELETÓN OCCIDENTE, GUADALAJARA
The Centro de Rehabilitacion Infantil Teletón offers integral management for the rehabilitation of children with neuromuscular problems. The attention model is based in the management by clinical groups, where children with pathologies as: light to moderate cerebral palsy, severe cerebral palsy, neuromuscular diseases, osteoarticular diseases, congenital and genetic diseases and spinal cord injury are treated.

It counts with auxiliary studies for Diagnosis in Specialized Radiology, Electroneurophysiology, Urodynamic and Gait Analysis and Human Movement Laboratory.

3 Activities

As concerns the activities of the three years of the project, courses and seminars devoted to training specialized MA personnel and to share the knowledge about the use of MA for clinical applications have been conducted. In addition, the mobility of researchers between European (EU) and Latin American (LA) countries has been promoted.

In particular, the Grant Holders (GH) - i.e. researchers involved in a MAL activity in a Full or Associate Partner, with an experience in the field of rehabilitation or with a technical education, like medical doctors, physiotherapists, engineers... - have been selected, at the beginning of the project, 3 for each LA country, 1 for each EU country. During the three years, as the interest for the project increased more and more, the number of the GHs involved in the project rose, mainly in LA, and many people took part to the actions proposed during the project.

They were mainly involved in the didactical activity, both theoretical and practical. They were trained not only in the basic technical and clinical competencies needed for everyday operation of MALs, but also in advanced technical topics, including the definition and implementation of new protocols. Both EU and LA GHs spent a period

in LA and EU Labs respectively, to take part in everyday MALs activity, to learn and share the practical management of experimental session and data interpretation. All the activities and the meetings are summarized in Table 1.

Table 1. Summary of all the activities of the TRAMA Project year by year.

MEETING	PLACE	PERIOD
FIRSTYEAR		
Start-up meeting	Milan	11th May 2007
Basics in Motion Analysis	Milan	10th-12th September 2007
Practical session	Milan	13th-22nd September 2007
Motion Analysis and clinics	Bruxelles	14th-17th January 2007
Motion Analysis Lab set up and running	Chile	10th-14th March 2008
SECONDYEAR		
The role of Motion Analysis in rehabilitation: decision making and treatment outcomes evaluation	Rome	21st-23rd May 2008
MALs management and organization	Milan	4th-6th June
Practical sessions in EU MALs		May, June, September 2008 (1 month)
Practical sessions in each own MAL	Own MAL	September 2008-February 2009
THIRDYEAR		
Final practical activity in LA MALs	Mexico City, Bogotá, Santiago	June 2009
MAL business plan simulation	Milan, Mexico City	September/October 2009
Final meeting	Bogotá	10th-12th March 2010

In particular, the training program of the project had a three-year schedule, compromising of alternating theoretical courses and practical sessions. The first year courses were focused mainly on training LA GHs in the basic skills of MA in the clinical setting. The aim was to supply the EU and LA GHs with a common background of basic competencies needed to establish a MAL from a technical point of view and to perform MA evaluations. At the beginning of the first year a start-up meeting was organized in May 2007 in Italy with all the partners' coordinators for in-depth analysis of required competencies and expertise for GHs selection and to agree on basic course contents. In the first year of training, there were three theoretical courses. The first two courses

were held in EU countries (in Italy, September 2007; in Belgium, January 2008) with two practical staying in EU after each course; the third course was held in LA (Chile).

Table 2. Summary of the theses prepared by Grant Holders during the TRAMA Project.

Name of the Institution	Title of the thesis
ITALY	
Politecnico di Milano - IRCCS "San Raffaele Pisana",Roma	Postural control in children, teenagers and adults with Down Syndrome.
COLOMBIA – SWEDEN	
Universidad Colegio Major de Nuestra Señora del Rosario, Bogotá, Colombia Karolinska Institutet, Stockholm, Sweden	Movement analyses in load lifting tasks-comparison of two methods for capturing and analyses of trunk kinematics.
Instituto de Ortopedia Infantil Roosevelt, Bogotá, Colombia	Intertrochanteric extension osteotomy to treat hip flexion deformity in walking children with spastic cerebral palsy.
CHILE	
Faculdad de Medicina, Univerisad de Chile, Santiago Sociedad pro ayuda del Niño Lisiado Instituto de rehabilitación infantil Teletón Santiago	Description of kinematic characteristics in children with lumbar and lumbosacral mielomeningocele and calculation of new indexes for a comprehensive evaluation.
MEXICO	
Centro de Investigacin y de Estudios Avanzados del IPN - CINVESTAV, Mexico City	Falling risk in elderly.
Centro de Rehabilitación Infantil TELETÓN CRIT, Mexico City	Kinematic upper limb assessment of children with hemiparetic CP during a reaching functional task.
Centro de rehabilitacion infantil Teleton Occidente, Guadalajara	Analysis of trunk mobility in children with scoliosis.

The second year was characterized by two theoretical courses in Italy (in Rome, May 2009; in Milano, June 2009) and by repeated practical sessions both in EU and LA labs for training in selected MA experimental set-ups and performance. The first course was aimed to give more detailed information about the use of motion analysis for clinical application to the LA GH; the second one was aimed to produce trained staff able to manage each aspect of MAL organization from fund-rising through to experimental sessions organization. The LA GHs spent their practical staying in Italy and in Sweden.

The third year was focused on checking the outcomes of previous years teaching and practical training. In particular, the activities were mainly dedicated to prepare educational material to give a scientific and didactical support to the GHs, as well as to

disseminate the results of the Project, and to prepare the GHs' theses, which were presented during the Final Meeting and included in the final hand book of the project. The GHs were involved in two practical sessions: the first one focused on completing their training in data collection and interpretation and the second one focused on simulating business plan to assure sustainability of a MAL. Both of sessions were aimed to the production of their personal thesis. The year was closed by a Final Meeting in Colombia in March 2010 where all the GHs presented their thesis.

During the period spent in their own MALs, all the GHs worked on new protocols and the results of this activity have been presented in theses (Table 2), collected in a book [1].

During the Final Meeting the general satisfaction for the three years project has been investigated by the TRAMA Project satisfaction, a questionnaire asked to be filled by all the participants (both professors and GHs) at the end of the Final Meeting; the same questionnaire was sent by e-mail to the TRAMA Project's participants not present at the event.

Following the questions were reported (score 1= poor, 2= adequate, 3=good, 4= excellent):

1. Which is your opinion regarding the TRAMA Project organisation (staff, travel booking, residence)?
2. Which is your opinion regarding the contents and the argument of the Project?
3. Were the arguments of Courses/practical sessions/Seminars adequate for your knowledge level?
4. Will be the contents learned during the Project useful for your activity in your Institution?
5. Are you generally satisfied of the Project?
6. If in the future there is a new opportunity of an international project on Movement Analysis, are interesting in being involved and taking part to it?

Following the results of the TRAMA Project satisfaction (Fig. 3)

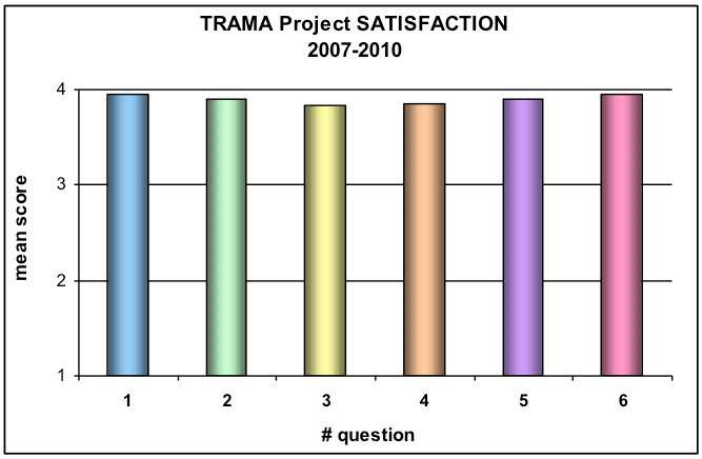


Fig. 3. Results of the TRAMA Project satisfaction questionnaire.

4 Conclusion

In conclusion, all the planned activities of the three year have been conducted and completed and a general satisfaction was present, by the Coordinator of the project and by the Partners.

We think that, thanks to TRAMA project, the knowledge about the use of Motion analysis and about MAL management is increased producing a more efficient use of Motion analysis for clinical applications.

In this way, MAL can improve the support service to clinical centers and so patients will benefit in diagnosis and follow up evaluations.

All the details about the Project program during these three years are summarized in the website `www.biomed.polimi.it/trama/`.

Acknowledgement. This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of Bioengineering Department - Politecnico di Milano (ITALY) and can under no circumstances be regarded as reflecting the position of the European Union.

References

1. Crivellini, M., Galli, M.: The TRAMA Project. Poliscript, Milano, ISBN:9788864930015 (2010) 1–303

Mobile Phones App to Promote Daily Physical Activity: Theoretical Background and Design Process

Gilles Kermarrec[1], Yannick Guillodo[2], Damien Mutambayi[2] and Léo Ballarin[2]

[1] European Research Center for Virtual Reality and Research Center for Education, Learning and Didactics, European University of Brittany, France
[2] Coaching Sport Santé Company, 2920 Brest, France

Abstract. Considering advances from research and technologies concerning physical activity and health, this chapter presents the C2S's Project (Coaching Sport Santé). C2S is a French start-up aiming at designing a device for promoting Daily Physical Activity (DPA). Based on Self-Determination, Self-Esteem Self-Regulation theories, and on the Trans-theoretical Model of Behaviour Change, a mobile phone App was developed including pedometer technology. The app offer a five step-strategy aimed at taking in account a) initial or normal everyday steps counts, b) individual motivational factor, and c) personally-adapted feedbacks.

1 Introduction

Benefits of physical activity for improving health are well established. Regular physical activity is associated with enhanced health and reduced risk of mortality factors, including cardiovascular disease, ischemic stroke, non–insulin-dependent, diabetes, colon cancers, osteoporosis, depression, and fall-related injuries (for an review see [1]). Therefore, a survey of EU countries demonstrated that two thirds of the adult population did not reach recommended levels of physical activity (http://www.who.int/whr/2002/en). In contrary, the prevalence of a sedentary lifestyle has been established in the European Space (e.g. in the 15 Member States of the European Union, [2]). Sedentary lifestyle has been defined according to various criteria: the number of hours that individuals spend sitting down in a typical day, the number of hours expended walking or in other specific physical activities, or how many times a week they participated in an activity that induced sweating [3]. Recently Europeans have been identified as high-risk populations; thus, the European Union's council recommendation of 26 Nov 2013 on promoting health enhancing physical activity called for monitoring of physical activity levels across member states. The Determinants of Diet and Physical Activity (DEDIPAC) European knowledge hub (www.dedipac.eu) organizes a major workshop on physical activity and sedentary behaviour surveillance and assessment in may 2015.

Considering these advertising, and advances from research and technologies, this chapter presents **the C2S's Project (Coaching Sport Santé).** C2S is a French start-up aiming at designing tools for promoting Daily Physical Activity (DPA). Especially, an interdisciplinary team of exercise psychologist, doctor and health experts, managers and engineers collaborated in designing a physical activity smartphone application.

Based on Self-Determination, Self-Esteem Self-Regulation theories, and on the Trans-theoretical Model of Behaviour Change, a device including a mobile phone App was developed to promote Daily Physical Activity (DPA).

The purpose of this chapter is to summarize reviews about physical activity interventions, to provide a framework for increasing DPA, and to present our App's design process and its outcomes.

2 Changing for Daily Physical Activity: Advances from Research

Recommendations in order to benefit from regular physical activity are well-known: 30 minutes of moderate-intensity activity on 5 or more days per week, or 20 minutes of vigorous-intensity activity on 3 or more days per week. Activities that expend 3- to 6-fold the energy expenditure of sitting at rest (3 to 6 metabolic equivalents or METs, 1 MET=3.5 ml O2•kg-1•min-1) are defined as moderate (walking), those that expend more as vigorous, and less as light intensity (running) [4]. Strategies and interventions to promote DPA are problematic. Several reviews have been conducted in this area and provided us with guidance for our project. The conclusions can be summarized into three points.

2.1 Several Types of Interventions for the Promotion of Physical Activity

Kahn et al., [5] produced a systematic review of interventions for the promotion of physical activity. Three categories of interventions have been distinguished:

- Informational approaches to change knowledge and attitudes about the benefits; especially in the self-regulation theory, knowledge about PA and health are a key component of the behavioural change mechanism. Knowledge help people to identify new goals and goals lead to behavioural strategies.
- Behavioural approaches to teach people the skills necessary for both successful adoption and maintenance of behaviour change. Especially studies including goal setting, self-monitoring, self-assessment, specific feedbacks showed that behavioural change in DPA could be achieved.
- Environmental and policy approaches to change the structure of physical and organizational environments to provide safe, attractive, and convenient places for physical activity. Especially, interpersonal setting is often thought to have potential for motivation (cooperation and competition) and social support are effective in increasing PA level.

Behavioural approaches and individually adapted health behaviour change programs consist of the most successful way. There is strong evidence that this kind of face-to-face PA programs are effective in increasing level of physical activity. Thus, face-to-face interventions are considered to be the optimal means for changing health-related behaviour.

2.2 Findings of Research on PA Interventions

Hillsdon et al., [6] summarized the evidence from sixteen systematic reviews and meta-analyses. Their 'review of reviews' provided a summary of the findings of research on interventions aiming at promoting physical activity for adults.

Evidences suggested that short-term change is achievable, and that use of a motivational and behaviour change theory will help. For instance, intervention could be based on various social cognitive theory of motivation (e. g., self-efficacy theory, self-regulation theory) or on the trans-theoretical model of change. Nevertheless, in many studies evidences were not consistent, or the research method could be criticized because PA was assessed thanks to declarative and self-administered questionnaires, such as the "International Physical Activity Questionnaire".

Moreover, the authors pointed out there are no consistent evidence for changes in workplace settings despite the fact that the importance of promoting physical activity through organisations is frequently pointed out. Especially the workplace, while targeted extensively in North America, has shown inconsistent involvement in physical activity promotion especially in European Space. Nevertheless, the workplace can offer large numbers of individuals and larger companies use to offer an infrastructure to support health promotion initiatives. Considering that adults spend about one quarter of their time at their place of work during their working lives, walking may be the best way to increase DPA, so that we suggested that pedometer technology could be relevant for the C2S project.

2.3 Pedometers: Furnishing a Realistic Measure for DPA

Biddle and Mutrie [7] produced a synthesis of the literature in which the use of pedometers is considered an efficient motivational tool. Using pedometers is not new, and studies showed they are accurate to count steps and assess PA [8]. Then, researchers have pointed out the effect of pedometers on motivation and PA [9]. Even the presence of pedometers alone could increase walking steps, and feedbacks from pedometers seem to be relevant information in order to involve motivation and DPA [10]. Therefore, Biddle and Mutries [7] pointed that in other studies walking steps did not increase significantly. In one of their studies they demonstrated that pedometers provided a short-term effect, but that this effect was not evident in the long term.

Thus, all over the world, campaigns promoted the idea that 10,000 steps a day are required for health, and pedometers seems to be a reliable technological support for assessing DPA and providing feedbacks to people. Therefore, reported an overview of 32 studies Biddle and Mutries pointed that 10,000 steps could be too low or too high objectives for some people (active or sedentary individuals), so that the 10,000 steps goal might lead to reduce motivation, especially if people do not feel able to reach this goal.

In accordance with these authors, we suggested that pedometers measures could involve PA programs' efficiency, because they allow to promote more adapted or personalized step-goals.

2.4 Conclusions

Research advances are particularly relevant to design intervention in DPA. There was strong evidence that motivational support through social cognitive motivation framework, self-regulation theory, or trans-theoretical model for behavioural change, should increase DPA program effectiveness. Interventions should target behaviour change by personalized and adapted interventions. Therefore, Interventions aiming at increasing DPA are still problematic because:

a) Information about PA and health are relevant but few healthcare professionals are trained to promote PA's effects
b) PA program Intervention is efficient only if it is individually adapted or when people is face-to-face
c) In most of countries it seems to be hard to change DPA in workplace settings
d) There was insufficient evidence that technology-based support interventions effectively increased physical activity

Finally, the C2S project tried to take in account these advances. The global project is described on the web, and many informational resources on PA and health are provided (www. agircontrelasedentarite.org). The C2S Project included a promotion initiative, called "Challenging sedentary lifestyle", gathering together many companies or workplaces settings. Previous research showed that promoting DPA should consider the workplace setting. Workplace could be an interpersonal setting which has potential for DPA changes, only if companies are members or partners of the project. In 2015, 20 local companies (in the west of France) and 318 employees participated voluntarily in this ride.

The global project and the challenge need efficient technological support. A mobile phone App (www.bouge-application.fr) was designed in order to deliver an individually adapted program. The design process was theoretically based on motivational and behavioural change frameworks. These frameworks and their implications for the program will be presented in the second section.

The mobile phone application includes pedometer technology and may offer efficient strategies aiming at taking into account a) initial or normal everyday steps counts, b) individual motivational factor, and c) personally-adapted feedbacks. The design process will be presented in the third section.

3 Enhancing Daily Physical Activity: Impact of Psychological Factors

The development of exercise psychology has led to the proliferation of theories, primary tested in social and health psychology. Thus psychological factors of physical activity have been studied extensively and helped us understand why people are motivated or not ("amotivated") and why they adopt a physically active lifestyle or not (sedentary). The study of human motivation has been central to exercise psychology. Vallerand and colleagues [11] offered an operational definition of motivation, considering some the following behavioural indicators: the initiation, the direction, the persistence, the intensity, and continued motivation. These components of a motivated

behaviour could be influenced by social and cognitive factors.

More precisely, considering this perspective and previous empirical studies we chose to take in account many models of motivation and behaviour in an exercise setting. They were supposed to lead to principles for DPA program and to guide the App's design process. Thus each model we chose can be considered as a blend of theoretical and practical support for our strategy and device. Consequently, our intervention strategy consisted of five steps (see figure 1), including diagnosing, initiating, monitoring, maintaining and evaluating. Each step has to be theoretically based, in order to insure the all-strategy reliability.

3.1 Diagnosing Attitude towards Behaviour Change: The Trans-Theoretical Model

The first model we took in account was the trans-theoretical model (TTM) of behaviour change. The TTM is not a model of motivation, but it as been classified as a stage-based behavioural model [7]. The TTM was developed as a comprehensive theory of behaviour change and was initially applied to smoking cessation [12]. The TTM has been applied to physical activity, it could be considered as a precious tool aiming at diagnosing the attitudes towards PA. The stages are [13]:

- Pre-contemplation: no intention to start physical activity
- Contemplation: considering starting physical activity
- Preparation: beginning a limited program of exercise
- Action: engaging in regular physical activity for less than six months
- Maintenance: engaging in regular physical activity for more than six months.

Studies [13] showed that TTM is a modest predictor of exercise.

Therefore, these stages are useful to diagnose if individuals are ready or not to accept the program; they should lead designers or practitioners to adapt their program.

In the TTM, the processes of change are the strategies used to progress along the stages of change. The processes are divided into cognitive (thinking) and behavioural (doing) strategies so that self-regulation theories should constitute a complementary resource for modelling changes in DPA.

3.2 Initiating and Monitoring Behavioural Change: The Role of Self-regulation Components

Aiming at understanding the initiation and the monitoring of behaviour, early attempts in exercise psychology favoured theories of perceived control. One of the most popular is the self-regulation theory advocated by Flavell [14]. Thus, goals, strategies, metacognition and knowledge are considered as components within the self-regulation process. Goals are considered as internal specific or general representations of a desired state: people could try to involve daily steps, or to be more active; sometimes they want to please a parent or the doctor, or to take care of themselves. Researchers interested in self–regulation showed that goals depend on prior knowledge about the concerned domain, and on metacognition.

Knowledge about PA and health (sometimes researchers called them beliefs) de-

termines motivation and behaviour towards PA. If someone knows that walking is the most reliable way to well being, he might adopt goal to change his sedentary behaviour. Then, self-regulation theory describes a relationship between goals and behavioural strategies: people whose goals are to enhance their DPA would adopt strategies to walk everyday. Strategies are means or solutions that people imagine, or that coaching should offer in order to reach a goal (e.g., walking during you're phoning; walking to get the next bus stop). These strategies are supposed to help people to monitor and involve DPA.

Metacognition is simply defined as cognition about self [14]. People engaging in metacognition will internally design knowledge about their own capabilities, and their own skills. If someone knows that he is not able to walk more than 5,000 steps a day, he spontaneously might not adopt a goal up to 10,000 steps a day. In contrary, if a PA program is addressed to him, and seems too difficult, too high, he simply would give up. Thus, metacognition is a precious component when designers want to select appropriate goals and strategies in exercise settings [15]. In this perspective pedometers should be helpful tools for people to know their own real DPA, and lead them to adopt a relevant goal.

Complementary, when one wants to initiate PA, goal setting theory, and the self-determination theory are also well-known resources. Research shows that motivation for physical activity is likely to be more robust if environment offers choices and self-determination rather than external control. This conducted us to consider that the program should offer alternative goals. Individuals should be invited to choose the best goals for themselves, or the most motivating one.

3.3 Maintaining Active Behaviour: Interest of Achievement Motivation

Because goals are personal representations, people are usually motivated through various types of goals. According to the Achievement Motivation Theory goals and behaviour could be referred to mastery – oriented or performance – oriented elements. Closely related to the issue of Goal Achievement Motivation Theory, the climate, or the relationships, within the exercise environment [16]. Perceptions of the motivational climate within a workplace or a training group can be classified as "mastery" or "performance". A mastery climate is one in which the participants perceive that self-improvement is the most important. A "performance" climate is one where participants are often compared with each other or with normatively superior performance (e.g. 10,000 steps a day).

A meta-analysis of climate studies across all physical activity settings quantified the links between climates and outcomes [16]. The overall effects from fourteen studies involving over 4,000 participants showed a large effect for mastery climate on positive outcomes and a moderate effect for performance climate on negative outcomes.

Because feedbacks in a device are important elements of a perceived climate, this line of research provides an important rationale for designing PA setting. Mastery oriented feedbacks or performance oriented feedbacks should be addressed to participant depending to their personal motivation orientation or to their physical self-esteem.

3.4 Evaluating Outcomes of DPA Intervention

Contemporary self-esteem theory proposes that our global view of ourselves ("global self- esteem") is underpinned by perceptions of specific domains of our lives, such as social, academic and physical domains. Based on this approach and on Fox previous work [17], Ninot et al., [18] has developed an operational measure of physical self-perceptions and its self-perception subdomains of sport competence, perceived strength, physical condition, and attractive body. Self – esteem theory proposed that everyday events are likely to affect more specific perceptions of self, such as the belief that one can walk 10,000 steps a day, which may eventually contribute to enhanced self-perceptions of physical condition or even physical self-worth. Self-perceptions could be important psychological constructs guiding general motivated behaviour, when people have to initiate PA. Self-perceptions can also be viewed, such as consequences or outcomes of a PA program.

Finally both of Physical – Self Esteem Scale and TTM of behaviour change furnished guides to implement the evaluation stage of the five-step intervention strategy.

3.5 Conclusions

The C2S Project aimed at implementing a technological solution, more specifically a mobile phone application, based on advances from research in exercise psychology.

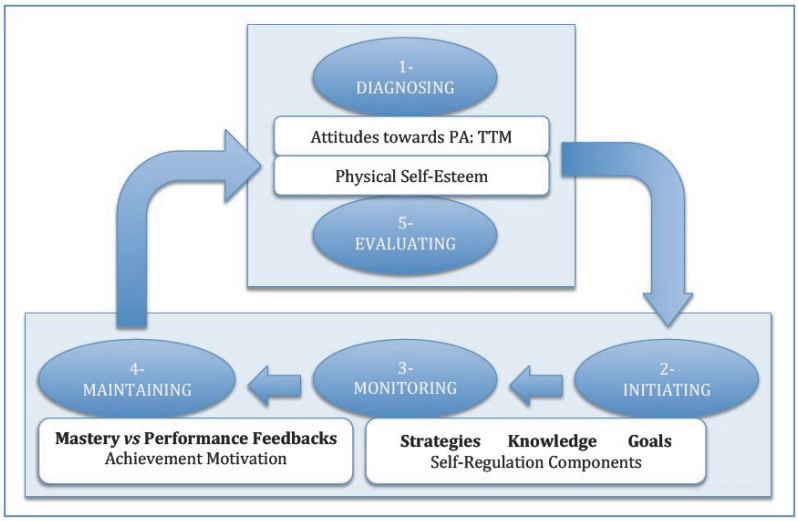


Fig. 1. The Five-Steps Strategy: a rationale for the App design process.

The program was considered as a set of personalized walking goals, behavioural strategies and knowledge about PA and health, and daily-individualized feedbacks. The question for the designer of a DPA program is how to deliver these information or artefacts to individuals? The whole strategy and its backgrounds are summarized in figure 1. These rational have been presented to the designers at the beginning of the design process.

4 Delivering a Personalized Physical Activity Program: A Design Process

Despite an explosion of mobile phone applications concerning PA, few have been based on theoretically derived constructs in order to promote health behaviours and reduce sedentary behaviour [19].

During the period from September 2013 – May 2014, the interdisciplinary team undertook initial app design, programming, and iterative user testing. Following these activities, initial app was developed and feedbacks from potential users were obtained. The Android smartphone platform was used. The smartphone battery life sufficiently to allow continuous accelerometer data capture throughout the day. The data collected via the smartphone's built-in accelerometer were transmitted to the project's local servers each evening for data storage and to allow researchers to monitor the quality of data while the design progressed. Feedbacks from users and ergonomic analysis lead to involve the design. During the period from September 2014 – January 2015, a second version of the app, called MOVE ("BOUGE" in French language), was developed, and was able to be commercialized. MOVE delivered an 8-weeks program aiming at involving DPA

The 5-steps intervention strategy was implemented within the design of the App.

4.1 Diagnosing

The initial session is used to provide instruction on the general use of the App, and to collect data including age, size, weight and gender. Especially, attitudes towards PA and Physical Self-Esteem are diagnosed. The user has to answer the Physical-Self Inventory (PSI-6), a six-item questionnaire developed and validated by Ninot et al., [18]. The PSI-6 is a short version of a previously validated questionnaire, the PSI-25, adapted from the Physical Self-Perception Profile [18]. PSI-6 contains one item for global self-esteem (GSE), one item for physical self-worth (PSW), and one item for each of the four sub-domains identified by Fox and Corbin [17]: physical condition (PC), sport competence (SC), attractive body (AB) and physical strength (PS). This questionnaire was proven to reproduce the factorial structure of the corresponding multi-items inventories [18] and to possess the same hierarchical properties. Each item is a simple declarative statement, to which participants was invited to respond using an analogic visual scale.

Attitude towards PA and behaviour change is measured using the stage of exercise behaviour change scale, adapted from Cardinal [20]. Users are asked to place themselves in one the five stages. During this first week, descriptions of the physical activity recommendations for health and sedentary lifestyle risks are available on a single screen.

The first week of the 12-weeks program is used as a baseline or a testing period for delivering an adapted program. Users are requested to continue with their normal physical activity and sedentary behaviours during the baseline week. The main screen of the app provides the user's current daily number of steps.

At the end of this initial week, the program can be personalized and users receive goals and advices.

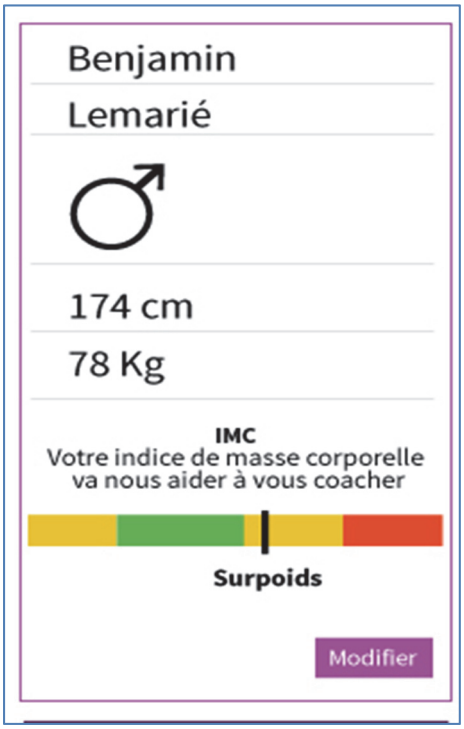


Fig. 2. User's profile on a screen.

4.2 Initiating and Monitoring

At the beginning of each week, users receive specific goal setting, which emphasized walking steps increase. For each week a distal goal is assigned depending both based on the score or number of walking steps the user reached at the end of previous week (e.g., 30 000 steps a week), and on his psychological profile (i.e., Physical Self-Esteem score). Depending on previous data, 5, 10, 15, 20% enhancement in weekly walking steps was used as references points. Participants were provided with three goal options of varying difficulty (e.g., 33 000 steps, or 34 500 steps, or 36 000 steps). These choice options were given based on the self-determination and goal-setting theory principles.

Whenever he wants, the user can see on the same screen his just-in-time score, the score for previous days, and the target at the end of the week.

In addition to having access to some "help" as part of this app, users participants can edit a set of behavioural strategies. On a specific screen, written solutions for increase DPA are listed and the user is invited to choose some of them. Twice a day brief health information and knowledge about benefits of PA (e.g., 1 minutes for PA = 10 minutes for life) are displayed.

Thus, goals, knowledge and strategies are supposed to stimulate self-regulation behaviour aiming at initiating and monitoring DPA.

Fig. 3a and 3b. User's daily current screens.

4.3 Maintaining

It was attempted that personalized feedback delivered by the app aim maintain motivation and PA. Balance sheets are provided in the middle of the fourth, the eighth and the eleventh week of the program. An email is sent to the user including graphs and encouragements or advices. Feedbacks' content depends on scores and on psychological profile. Thus they can be mastery – oriented (e.g., Congratulations! An increase of 15% over the previous weeks! Go on! You walk for your health!), or performance – oriented (Congratulations! You've reached 50 000 steps a week! You're now considered as an active person!).

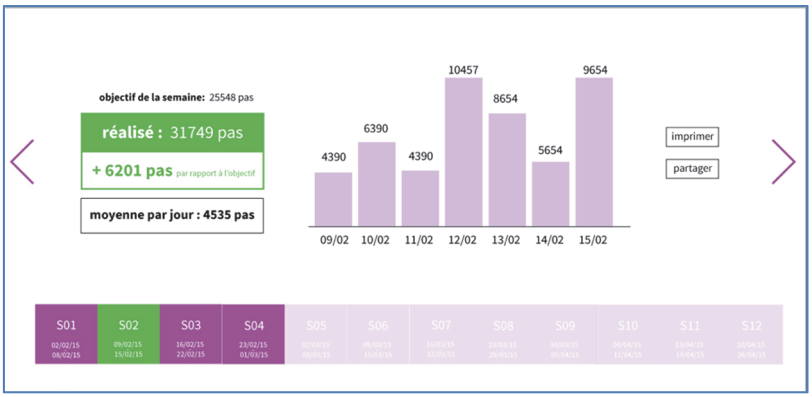


Fig. 4. User's balance sheet.

4.4 Evaluating

In line with previous research in exercise psychology, pedometers are reliable for furnishing a realistic and objective measure about everyday walking steps. Such an on-going evaluation is supposed to impact motivation and behaviour control. Users perceived the gap between the goal they have chosen and their own DPA. Consequently, self-regulation mechanisms consist in imagining solutions, imitating pairs' behaviour, seeking for advices. The app helps people in taking in charge their own behaviour.

Finally, as soon as the program is closed, users receive a final balance sheet and advices for future. This constitutes a milestone in our 5-steps strategy; when someone is able to evaluate his own progress, it was hypothesized that his own self-perception would enhance. Thus at the end of the program, users are asked to answer the Physical Self-Esteem. If users observe increases concerning both of DPA and Physical Self-Esteem, they would get confidence in the device effectiveness. These psychological effects would favour continued motivation for DPA.

5 Perspectives

Benefits of Physical Activity are attempted for adults in the European Space. The C2S Project including a web resource, a ride for companies called "Challenging sedentary lifestyle", and the MOVE App, is a medical, technologic and scientific project. Thus, an empirical study (200 participants in experimental group *vs* 100 participants in control group) has been conducted to assess the effect of the program on DPA. Results could have important implications for advancing the field of PA sciences, and will be precious to involve the design of the App. Moreover collected data on daily PA or behaviour changes in a workplace setting will be stored and should be useful for health institutions.

References

1. O'Donovan, G. , Blazevich, A., Boreham, C., Cooper, A.H., Crank, H., Ekelund, U., Fox, K.R., Gately, P., Giles-Corti, B., Gill, J.M.R., Hamer, M., McDermott, I., Murphy, M., Mutrie, N., Reilly, J.J., Saxton, J.M. and Stamatakis, E. : The ABC of Physical Activity for Health: A consensus Statement from the British Association of Sport and Exercise Sciences. J. of Sports Sc. 28, 6 (2010) 573-591
2. Blair SN. : Physical Inactivity: the Biggest Public Health Problem of the 21st Century. Br. J. Sport. Med. 43 (2009) 1–2
3. Varo, J.J., Martinez-Gonzalez, M-A.: Current Challenge in the Research About Physical Activity and Sedentary Lifestyle. Rev Esp Cardiol. 60 (2007) 231-233
4. Williams, P.T. and Thompson, P.D.: Walking vs Running for Hypertension Cholesterol, and Diabete Risk Reduction. Art. Thomb. Vasc. Biol. 33, 5, (2013) 1085-1091.
5. Kahn, E. B., Ramsey, L. T., Brownson, R. C., Heath, G. W., Howze, E. H., Powell, K. E., Stone, E. J., Rajab, M. W. and Corso, P.: The Effectiveness of Interventions to Increase Physical Activity: A Systematic Review. Am. J. of Pre. Med. 22 (2002) 73–107.3.
6. Hillsdon, M., Foster, C., Naidoo, B. and Crombie, H.: A Review of the Evidence on the

Effectiveness of Public Health Interventions for Increasing Physical Activity Amongst Adults: A Review of Reviews. London: Health Development Agency (2003).
7. Biddle, S. J. H. & Mutrie, N.: Psychology of Physical Activity. Determinants, Well-being and Intervention. 2nd Edition, Routledge, New-York and London (2008).
8. Crouter, S., Schneider, P., Karabulut, M. And Basset, J.: Validity of 10 Electronic Pedometers for Measuring Steps, Distance, and Energy Cost. Med. Sc. In Sp. and Ex. 35, 8 (2003) 1455–60.
9. Rooney, B., Smalley, K., Larson, J. and Havens, S.: Is knowing Enough? Increasing Physical Activity by Wearing a Pedometer. Wis. Med. J. 102, 4 (2003) 31–6
10. Stovitz, S., VanWormer, J., Center, B. and Bremer, K.: Pedometers as a Means to Increase Ambulatory Activity For Patients Seen at a Family Medicine Clinic. J. Am. B. of Fam. Prac. 18 (2005) 335–43
11. Vallerand, E.L., Pelletier, R.J., and Ryan, L.G.: Motivation and Education: the Self-Determination Perspective. The Ed. Psy. 26 (1991) 325-346
12. Prochaska, J. O. and DiClemente, C. C.: Transtheoretical Therapy: Toward a More Integrative Model of Change. Psychotherapy: Th. Res. and Prac. 19 (1982) 276–88
13. Callaghan, P. Khalil, E. and Morres, I.: A Prosective Evaluation of th Transtheoretical Model of Change Appled to Exercise in Young People. Int. J. of Nurs. St. 47 (2010) 3-12
14. Flavell, J. H. Cognitive monitoring. In: Dickson, W. P. (Ed.): Children's oral communication skills. New York: Academic Press (1981)
15. Kermarrec, G., Todorovitch, J. and Fleming, D.: Investigation of the Self - Regulation Components Students Employ in the Physical Education Setting. J. of Teach. in Phys. Ed. 2, 23 (2004) 23-142
16. Ntoumanis, N. and Biddle, S. J. H.: A Review of Motivational Climate in Physical Activity. J. of Sp. Sc. 17 (1999) 643–65
17. Fox, K.H. and Corbin, C.B.: The Physical Self-Perception Profile: Development and preliminary validation. J. of Sp. and Ex. Psy. 11, (1989) 408-430
18. Ninot, G., Fortes, M., and Delignières, D.: Validation of a Shortened Assessment of Physical Self in Adults. Perc. And Mot. Skills 103 (2006) 531-542.
19. King, A.C., and al.: Harnessing Different Motivational Frames via Mobile Phones to Promote Daily Physical Activity and Reduce Sedentary Behaviour in Aging Adults. Plos One 8, 4 (2013) 1-8
20. Cardinal, B.J.: Construct Validity of Stages of Change for Exercise Behaviour. Am. J. of Health Prom. 12, 1 (1997) 68-74

Author Index