Personalised wearable cardiac sensor services for pervasive self-care

Asta Krupaviciute

N° d'ordre 2011ISAL0156 Année 2011

Thèse

Personalised wearable cardiac sensor services for pervasive self-care

Conception de services personnalisés pour la capture ubiquitaire de signes vitaux en pSanté

Présentée devant
L'Institut National des Sciences Appliquées de Lyon

Pour obtenir
Le grade de docteur

Formation doctorale : Informatique
École doctorale : Informatique et Mathématiques

Par
Asta KRUPAVICIUTE

Soutenue le 20 décembre 2011 devant la Commission d'examen

Jury MM.

Président	R. BEUSCART	Professeur (Université de Lille 2)
Rapporteur	G. CARRAULT	Professeur (Université de Rennes 1)
Rapporteur	C. CAUVET	Professeur (Université d'Aix-Marseille III)
Co-directeur de thèse	J. FAYN	Ingénieur de recherche HDR (INSERM) (Université de Lyon)
	E. McADAMS	Professeur (INSA de Lyon)
Directeur de thèse	P. RUBEL	Professeur Emérite (INSA de Lyon)
	A. SIMONET	Maître de Conférences HDR (Université Pierre Mendès France, Grenoble)
	C. VERDIER	Professeur (Université Joseph Fourier, Grenoble)

Laboratoire de recherche : MTIC (EA4171) & SFR Santé Lyon Est (US7-UMS3453)

Acknowledgement

I would like to express my deep and sincere gratitude to this thesis supervisors Jocelyne Fayn and Paul Rubel together with Eric McAdams, Christine Verdier and Chris Nugent for their help and guidance during the development and the completion of this research study. Jocelyne Fayn, PhD, has devoted much time and effort from the initial to the final stage and helped me to develop a good understanding of the subject. Her truly scientist intuition has inspired my work and has enriched my growth as a student and a researcher. I gratefully acknowledge Prof Paul Rubel who with patience taught me the subtleties of a new to me cardiovascular disease domain and during all research work was offering me an abundant help and invaluable assistance. I appreciate a lot Prof Eric McAdams's positive attitude and aid. Also I sincerely thank Prof Christine Verdier for giving me a truly help to the progress and smoothness of this research work. Finally, for Prof Chris Nugent willingness to motivate me contributed greatly to this work.

I would like to thank very much all the external members of my thesis committee who have accepted to judge my research work. I especially express my deep gratitude to this thesis' reviewers, Prof Corine Cauvet and Prof Guy Carrault, for their time and dedication and for their very constructive comments. I am greatly honoured that Prof Régis Beuscart, who is a famous expert in eHealth, accepted to be the President of my thesis committee, and I am very thankful to Ana Simonet, PhD, who agreed to provide her insightful comments to my research study.

I wish to express my thank to this research project captain Prof Jean Louis Sourouille and for the WARTHE project which I had an opportunity to join. Deepest gratitude to all members of this project for broadening my view and knowledge.

During this work I have collaborated with many colleagues for whom I have great regard, and I wish to extend my warmest thanks to all those who have helped me with my work in the MTIC, LIG, and NIBEC.

I am grateful to Prof Antanas Mitašiūnas for attracting me to an international academic research and for valuable discussions I had a chance to participate in. His ideals and concepts have had a remarkable influence on my entire career.

I wish to express my warm thanks to Kristijonas Šiaulys who from the beginning supported my choice to work in academic research. A big thank to Hussein Atoui and Ieva Mitašiūnaitė for making my start in France much easier. I also owe a sincere gratitude to Joël Placide who did not let me to give up. His wide knowledge, huge tolerance, logical thinking and explanations have been of great value to me. Finally, I warmly thank to Simon Mitschke and Jeremy Hamilton for their enthusiasm and involvedness in extensive discussions around my work and constructive comments.

I owe my loving thanks to my family, parents Emilija and Vytautas, and my brother Virginijus. Without their overwhelming optimism and steady trust in me along with the encouragement and the understanding it would have been impossible for me to finish this work. My special gratitude is also for my close old friends Gediminas Misius, Vilija Tauraitė, Aurelija Proškutė, Edvinas Bruzgulis and Antanas Matevičius as well as Farnoush Noori, Graham Dale and Adrien Revault d'Allones whom I was lucky to know during this period and who all together stood behind me like a wall nourishing my confidence to step only forward. I remember every detail and each moment you offered me while I was strolling around Europe: picking me at airports, giving a room to sleep, sending me a book or a magazine, a postcard, a birthday gift, medicaments, giving a hug, inviting for a cup of coffee and many other moments. I remember all as it always caused flashes of tremendous happiness inside me and taught me that the distance is not important when there is a bond between people. I feel happy you all let me to experience it.

Conception de services personnalisés pour la capture ubiquitaire de signes vitaux en pSanté

Résumé

L'objectif de cette thèse est de concevoir une architecture de services Web pour la détermination automatique d'un système de capteurs personnalisé, embarqué sous forme de vêtement intelligent dédié au self-care, afin de permettre à tout utilisateur profane d'enregistrer lui-même son propre électrocardiogramme (ECG), à tout moment et n'importe où. Le défi principal réside dans l'orchestration intelligente et dynamique de services métiers en fonction du contexte, pour qu'ils fournissent à l'utilisateur une solution personnalisée optimale tout en maîtrisant la complexité inhérente à la dépendance au contexte des interactions homme-machine, à l'extraction des connaissances des signes vitaux spécifiques à un sujet, et à l'automatisation de la reconfiguration des services. Une solution à ce défi est de créer une intelligence ambiante qui étend la notion d'informatique ubiquitaire et est capable d'offrir à l'instar d'un expert du domaine, une assistance intelligente personnalisée à tout citoyen.

Nous proposons une méthodologie de construction d'une architecture (DM-SOA) orientée-services, dirigée à la fois par les données et par des modèles, pour la production de services métiers intelligents et tenant compte du contexte. Cette architecture permet l'automatisation d'activités sophistiquées et personnalisées qu'un expert mettrait en œuvre pour le traitement d'un cas individuel, à partir de ses connaissances professionnelles. La solution proposée est basée sur une nouvelle approche de modélisation dynamique de processus métiers, et l'implémentation de services reconfigurables automatiquement. Elle consiste à mettre en œuvre un environnement intelligent fondé sur une ontologie de processus métiers des concepts du domaine et de son contexte, et sur une base de règles pour l'orchestration contextuelle des services.

Pour valider le bien-fondé de notre approche, nous avons proposé une ontologie pour l'automatisation de processus d'aide à la décision et nous l'avons évaluée dans le domaine de la cardiologie, en l'appliquant au problème de la sélection la plus adéquate possible d'un système de positionnement d'électrodes, spécifique à chaque individu, et capable de fournir un signal ECG de contenu diagnostique similaire à celui d'un ECG standard à 12 dérivations. Pour répondre aux besoins en situation de self-care tout en préservant la qualité diagnostique des signaux enregistrés à la demande, nous proposons la réalisation d'un nouveau système prototype de capture ECG-3D à trois modalités. Ce prototype a été testé sur huit jeunes adultes volontaires sains (4 hommes et 4 femmes) présentant diverses caractéristiques morphologiques.

Dans ce contexte, l'intelligence ambiante est fournie par un ensemble de services de qualité professionnelle, prêts à l'emploi par des utilisateurs profanes. Ces services peuvent être accessibles au travail, n'importe où, via des moyens classiquement utilisés chaque jour, et fournissent une aide appropriée aux utilisateurs non-compétents. Une telle approche d'intelligence ambiante s'inscrit dans la vision d'une société de l'information ambitionnant de faciliter l'accès à la connaissance et correspond aux objectifs à long terme du programme de recherche Information Society Technologies de l'Union Européenne. Cette approche devrait contribuer à l'amélioration de la santé des individus.

Mots-clés : Intelligence ambiante – SOA – ontologie – modélisation dynamique de processus métiers – orchestration des services Web – Aide à la décision – pSanté – ECG – électrocardiologie quantitative – capteurs biomédicaux – informatique ubiquitaire

Personalised wearable cardiac sensor services for pervasive self-care

Abstract

The aim of the thesis is to design a web services architecture that shall support the automatic determination of a personalised sensor-system, which is embedded in smart garments, and which shall be used in self-care in order to allow a profane user to record himself a personal electrocardiogram (ECG), at anytime and anywhere. The main challenge consists in the intelligent and dynamic orchestration of context-aware business services that supply the user with an optimal personalised solution, while mastering the system's complexity: context dependent user and system interactions, knowledge extraction from subject-specific vital signs, services reconfiguration automation. The solution to this challenge is to create an Ambient Intelligence which goes beyond Ubiquitous Computing and is capable to replace an expert by proposing an Intelligent Assistance to any citizen.

We propose a methodology expressed in terms of Data and Model driven Service Oriented Architecture (DM-SOA), which provides a framework for the production of context-aware intelligent business services. This architecture supports the automation of sophisticated and personalised expert activities, which apply professional knowledge to process an individual case. The proposed solution is based on a new dynamic business process modelling approach and in its implementation via automatically reconfigurable services. It consists in setting-up an intelligent environment based on a business process ontology of context-aware concepts and on related context handling rules for services orchestration. A core-ontology has been designed to support the automation of expert activities related to decision-making. The soundness of the method and of the underlying decision-making techniques has been demonstrated in the cardiology domain for selecting the most adequate subject-specific sensor-system, characterised by its ability to yield for an ECG signal of similar diagnostic content than a standard 12-lead ECG. We propose the design of a new three modalities sensor-system prototype as a response to the need of sensor-systems used on demand in self-care situations and ensuring diagnostic quality signals recordings. This prototype has been tested on a set of healthy volunteers presenting various characteristics in age, sex and morphology.

In this context, Ambient Intelligence is demonstrated as an ensemble of professional quality services ready to use by profane users. These services are accessible pervasively, trough the objects people work with / wear on / use each day, and provide an appropriate guidance to the non-competent users. Such an Ambient Intelligence approach strongly supports the vision of an Information Society which corresponds to the long-term goal of the EU Information Society Technologies Research Programme that aims to ease knowledge access. The future healthcare will benefit from this approach and will significantly improve patient care.

Keywords: Ambient Intelligence – SOA – ontology – dynamic business process modelling - web-service orchestration – decision support system – Personal Health System – ECG – quantitative electrocardiology – biomedical sensors – pervasive computing

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	**CHIMIE DE LYON** **http://www.edchimie-lyon.fr** INSA: R. GOURDON	M. Jean Marc LANCELIN Université de Lyon – Collège Doctoral Bât ESCPE 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04 72 43 13 95 directeur@edchimie-lyon.fr
E.E.A.	**ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE** **http://edeea.ec-lyon.fr** Secrétariat : M.C. HAVGOUDOUKIAN eea@ec-lyon.fr	M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04 72 18 60 97 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr
E2M2	**EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION** **http://e2m2.universite-lyon.fr** INSA : H. CHARLES	Mme Gudrun BORNETTE CNRS UMR 5023 LEHNA Université Claude Bernard Lyon 1 Bât Forel 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cédex Tél : 04 72 43 12 94 e2m2@biomserv.univ-lyon1.fr
EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTE** **http://ww2.ibcp.fr/ediss** Secrétariat : Safia AIT CHALAL INSA : M. LAGARDE	M. Didier REVEL Hôpital Louis Pradel Bâtiment Central 28 Avenue Doyen Lépine 69677 BRON Tél : 04 72 68 49 09 Fax : 04 72 35 49 16 Didier.revel@creatis.uni-lyon1.fr
INFOMATHS	**INFORMATIQUE ET MATHEMATIQUES** **http://infomaths.univ-lyon1.fr**	M. Johannes KELLENDONK Université Claude Bernard Lyon 1 INFOMATHS Bâtiment Braconnier 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04 72 44 82 94 Fax : 04 72 43 16 87 infomaths@univ-lyon1.fr
Matériaux	**MATERIAUX DE LYON** Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry Ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cédex Tél : 04 72 43 83 18 Fax : 04 72 43 85 28 Jean-yves.buffiere@insa-lyon.fr
MEGA	**MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE** Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry mega@insa-lyon.fr	M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04 72 43 71 70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr
ScSo	**ScSo*** Secrétariat : Viviane POLSINELLI INSA : J.Y. TOUSSAINT	M. Lionel OBADIA Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Tél : 04 78 69 72 76 Fax : 04.37.28.04.48 Lionel.Obadia@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Table of contents

FOREWORD ***11***

Chapter 1 – Introduction **13**
1.1. Motivation & Context 15
1.2. Research in the *Ambient Intelligence* domain 17
1.3. Research objectives 18
1.4. Dissertation organisation 19

PART I TOWARDS UBIQUITOUS SELF-CARE IN CARDIOLOGY ***21***

Chapter 2 – State of the art in the electrocardiology domain **23**
2.1. Introduction 25
2.2. Medical background 25
2.3. Electrical activity of the heart: the ECG signal 27
2.4. Information capture 28
2.5. Information processing 37
2.6. Information reconstruction 40
2.7. Information interpretation 43

Chapter 3 – Review on ECG acquisition modalities in self-care **45**
3.1. Introduction 47
3.2. Smart accessories 48
3.3. Smart garments 51
3.4. Sensing possibilities and constraints 53
3.5. Monitoring: *continuous* or *on demand* 56
3.6. Concluding insights 57

Chapter 4 – Personal sensor-system modalities design and evaluation for simplified electrocardiogram recording in self-care **59**
4.1. Introduction 61
4.2. Rationale for personal sensor-system modalities design 61
4.3. Sensor-system modalities design 63
4.4. Study population 65
4.5. Data acquisition 67
4.6. Data analysis method 67
4.7. Results obtained with comparison studies: C1, C2 and C3 73
4.8. Paired t-test results obtained comparing the C1, C2 and C3 studies 82
4.9. Repeatability study results 83
4.10. Discussion 85
4.11. Conclusion 86

PART II AMBIENT INTELLIGENT ASSISTANCE ***89***

Chapter 5 – State of the art **91**
5.1. Introduction 93
5.2. Business process modelling and management 94
5.3. Architectural foundation for services composition 101
5.4. Knowledge engineering & Knowledge management 106
5.5. Context-awareness versus Situation-awareness 110
5.6. Decision-making and support 113
5.7. Summary 116

Chapter 6 – Process knowledge modelling **117**
6.1. Introduction 119

6.2. Expert activities automation ... 120
6.3. Business intelligence view ... 124
6.4. Compositional model ... 126
6.5. Dynamic business process ... 128
6.6. Core-ontology for decision support process knowledge modelling ... 130
6.7. Discussion ... 133

Chapter 7 – Data and Model driven Service Oriented Architecture (DM-SOA) ... 135
7.1. Introduction ... 137
7.2. Engineering decisions ... 137
7.3. General framework ... 138
7.4. Generic DM-SOA for decision support systems ... 141
7.5. Domain-specific DM-SOA for decision support system ... 147

Chapter 8 – Enhanced services orchestration ... 151
8.1. Introduction ... 153
8.2. Use-case scenarios ... 153
8.3. Optimal personal ECG recording system selection: Service Choreography ... 159
8.4. Optimal personal ECG recording system selection: Service Orchestration ... 161
8.5. Conclusion ... 170

CONCLUSIONS AND FUTURE WORK ... 171

Chapter 9 – Final conclusions and perspectives ... 173
9.1. Conclusion ... 175
9.2. Discussion ... 176
9.3. Limitations ... 177
9.4. Future work ... 177

REFERENCES ... 179

TABLE OF FIGURES ... 192

LIST OF TABLES ... 194

ANNEXES ... 195

Annex 2.1 – Principles of the heart's electrical system ... 197

Annex 4.1 – Initial simulation results ... 199

Annex 4.2 – Example of a subject specific data grid ... 203

Annex 4.3 – Repeatability test results ... 204

Foreword

This thesis has been performed within the framework of the European project “Wide Area Research Training in Health Engineering (WARTHE)” of the Marie Curie “Early Stage Research Training” FP6 Programme. The objective of this project was to prepare new generation scientists that are able to integrate knowledge from several domains.

The aim of this thesis is to design a web services architecture that shall support the automatic determination of a personalised wearable cardiac sensor-system, which should be used in pervasive self-care by any profane citizen. The expected societal outcome is to improve early detection of cardiovascular diseases, at anytime and anywhere. Our research work thus deals with multidisciplinary topics, such as sensors and biomedical engineering, signal processing and quantitative electrocardiology, pervasive computing, knowledge engineering, web services composition and ambient intelligence.

Chapter 1 – Introduction

1.1. Motivation & Context 15
1.2. Research in the *Ambient Intelligence* domain 17
1.3. Research objectives 18
1.4. Dissertation organisation 19

1.1. Motivation & Context

Due to the ageing population, the healthcare and the delivery of health services today experience new challenges. People aim at prolonged life time with preserved quality of life and are ready to actively participate in their own health management. A self-care approach allowing the supervision of diseases by the citizen himself and aiming at effective detection and treatment of any problem before it necessitates complex and costly emergency interventions becomes particularly attractive in this context. These current demographic and cultural changes demand the traditional medicine, which is mainly focused on disease treatment in centralised healthcare units by professionals, to make necessary reforms to get along with new expectations of society and to ensure the quality of service in these changed conditions.

This shift towards the preventive, personalised and citizen-centred healthcare also has a strong economical basis. Healthcare costs, especially in chronic disease (mainly cardiovascular disease, but also cancer, diabetes and obesity) management are steadily increasing: the latest reports chronic disease healthcare costs in United States of America estimate that over 75% of the total national healthcare expenditures is spent for treatment of ongoing illnesses and long-term care services *[Orszag'08], [CDC'11]*. Currently, the expectations and so the investments of the healthcare community are set on the design and development of new utilities and services that would be accessible to everyone and at any time and that would have a low cost and so could deliver healthcare services cheaper and in more responsive ways than services that are now obtained through large centralised institutions.

The self-care approach is crucial within the cardiac domain, where late detection increases the likelihood of patient disability or of premature death. A worrying statistic that 60% of citizens, who suffered from a stroke, die or become dependent on others for help with their everyday living indicates prevention as a high priority strategy *[Mackey'04]*. The worst factor is that almost two thirds of cardiac deaths occur outside of hospitals, because people misinterpret first symptoms and delay the visit to specialised healthcare units *[TFR'98]*. New healthcare scenarios and new tools helping to reduce the time between the first symptoms and the decision to treat are needed.

Indeed, as related in *[Fayn'10]*, "*the relationship between mortality and time delay to treatment in case of acute coronary syndromes has been widely reported for several years [Luca'04], [Luepker'05], [McGinn'05], [Garvey'06]. In addition, according to World Health Organization (WHO), heart diseases remain the main cause of mortality in industrialized countries. Enabling a common use of very popular, advanced Personal Health Systems (PHS) is the first issue to be addressed for improving patient's health support where he/she is, at any time, at work, at home or during a trip. Their large scale dissemination is thus promoted by WHO eHealth resolutions for an overall development of information and communication technologies for health [Healy'06], [Dzenowagis'05]*."

Motivating scenario

As stated above, an ideal prevention scenario for cardiovascular disease management is foreseen to be based on non-invasive personal health monitoring systems with easy access for everybody, at anytime or location. A citizen there takes an active role, which is supported by physician or pharmacist roles and makes constant check-ups when he (her) thinks they are needed. This scenario is seen to be set in an intelligent self-care environment, which could be achieved by the use of a personal smart device (Personal Cardio Assistant) capable to record and analyse a patient specific electrocardiogram (ECG). The ECG recording remains indeed the most important, independent, objective and non-invasive source of information for the clinical evaluation of patients with chest pain*[Lee'00]*, *[Garvey'06]*, and/or heart rhythm abnormalities *[Zipes'06]*. The personal device needs to have at least two principal elements: 1) the intelligent agent (monitor), which might be sold as a separate device, might be embedded into a wallet, a mobile phone, a watch, or even installed into an ordinary personal computer and serves for signal recording, storing, processing and analysis, and 2) a sensor-system, which must be easily, though reliably and conveniently placed on an individual in their appropriate location by a non-specialist citizen himself, in order to capture the heart’s electrical signal. A Personal Cardio Assistant shall contain the functionality which allows recording anywhere and at anytime a diagnostic quality ECG and analysing it “on-board” by comparing it to a reference ECG, if possible. In addition, it shall be capable of making a reliable recommendation from a digital artificial intelligence-based diagnosis, which would enable any citizen to take timely and correct decisions upon his (her) health status. In a case of urgency, the assistant shall put the citizen in a contact with healthcare providers.

In the late 1990s, a concept of Personal Health Systems (PHS) was introduced to describe a new type of health care delivery process, which puts the patient at the centre of it and enables him to take timely decisions with the help of new tools that ensure pervasive self-care. Starting from then new tools capable to record one or several biosignals (pulse, body temperature, respiration rate, electrocardiogram and others) in home-based conditions were developed. These novel systems were aiming at convenient and robust patients monitoring in their homes while they are performing their daily activities, trying to not interfere significantly with their comfort or lifestyles. Quickly wearable sensor technologies took there a pivotal role. Current research projects are focusing on four groups of wearable monitoring platforms: (i) “Holter-type” systems with standard sensor design and locations, (ii) Body-worn sensor patches, (iii) Body-worn bands and harnesses and (iv) “Smart garments” for long-term application *[McAdams'11a]*. However, all these already established approaches have various limitations, for an on demand use of the corresponding sensor-systems by every citizen, especially in self-care situations. For instance they need a constant professional supervision or a prolonged continuous wearing and so they are not ensuring the independent and occasional use of such devices by a profane.

In case of occasional monitoring, a potential clinical need is to have at the user’s disposal a small, portable system, maybe even integrated into a mobile phone or a watch, which goes together with a special sensor-system, a kind of a bib or an apron, that is easily placed on the thorax to capture the heart electrical signal, easily transported and wear on and off in any place. The system shall be compliant with the body surface map configuration and shall enable a profane user to make a quick and discrete recording of his vital signs. It also shall provide an initial diagnosis leading the user to make a rapid and evidence-based decision upon his health status. This direct feedback or the possibility to send information to a remote monitoring station in case of an emergency was a key target of the European IST EPI-MEDICS project *[Fayn'10]*. The project outcome is a prototype of a Personal ECG Monitor (PEM), which includes a reduced-electrodes set that allows 10 seconds recording of three pseudo-orthogonal ECG leads I, II and V2 according to the Mason-Likar system. It also embeds algorithms allowing the reconstruction of standard 12-lead ECGs and a risk stratification based on signals analysis. However, as it has been widely reported, the diagnostic accuracy of the ECG and thus the success of the PEM and of similar systems depends on the user ability to apply sensors in the correct anatomical positions. This task requires a specific knowledge and time to find accurate electrodes' locations, and according to several studies remains problematic even for skilful professionals *[Rajaganeshan'08], [McCann'07], [Sejersen'06]*.

The signal pattern sensitivity related to sensors placement is the main issue this thesis addresses. Sensors displacement errors have a direct impact on the medical diagnosis, especially if the diagnostic procedure is automated *[Schijvenaars'97]*. It is crucial to propose personalised and automatic electrode placement solutions in order to overcome the inter-subject and intra-subject variability and to ensure the records quality, thus an accurate and faithful diagnosis. Current prototypes of devices intended to fulfil Personal Health Systems scenarios would greatly benefit of this achievement, which would give them a strong input to pass beyond the prototype phase towards clinical trials.

1.2. Research in the *Ambient Intelligence* domain

Placing sensors together with input/output modules at different thorax locations to record a diagnostic quality ECG is a key responsibility of the sensors-system. An accurate recording is an initial step to a correct interpretation and to a trustworthy diagnosis. However a digital evaluation of an ECG signal leading to digital profane user guidance in order to obtain a required quality ECG is an open multidisciplinary research area of an *Ambient Intelligence*.

Ambient Intelligence represents a vision of a future interactive environment equipped with smart electronics, usually embedded in every-day objects. It, actually, stands for a network of hidden intelligent interfaces that recognise our presence and mould our environment to our immediate needs. *Ambient Intelligence* demands electronics to sense and to adapt to our context, to create a natural interaction with us and even to be proactive if it is necessary. This context-aware environment aims to enhance our productivity, creativity, expressiveness, healthcare and well-being.

Ambient Intelligence steps beyond *Ubiquitous Computing* because additionally it includes *Intelligent Assistance*. Figuratively, we could see *Ambient Intelligence* as the human brain, which has two distinct cerebral hemispheres: *Ubiquitous Computing* and *Intelligent Assistance*:

- *Ubiquitous Computing* makes microprocessors pervasive by integrating them into our environment: furniture, clothes, toys, etc. Usually, these improved devices then facilitate the performance of some routine based scenarios.
- *Intelligent Assistance* aims to deliver a professional support in cases where sophisticated and personalised solutions are needed. It performs intelligent business services that correspond to expert activities based on the application of professional knowledge to process an individual case. It is applied in some complex domains such as preventive medicine or e-learning where reliable and self-adaptive solutions are expected by users.

Like the brain that needs both sides to operate fully, the *Ambient Intelligence* shall integrate the wisdom of each related though distinct part to provide an ensemble of professional quality services ready to use by profane users. The number of research projects on the *Ubiquitous Computing* part shall be joined with works on *Intelligent Assistance*. This connector role, helping to make a meaningful whole, is taken by the *Ambient Intelligence*. The research in the *Ambient Intelligence* domain shall lead towards the contextual information management and data that are important about the case processing. It shall propose how to derive different data from a case, how to transform these data and integrate them with contextual information while using available and suitable data processing and analysis methods in order to achieve the required result. As knowledge about the case is usually heterogeneous, voluminous, poorly structured and steadily evolving, its automated use in practice is extremely difficult. The lack of the global picture, of the theoretical and methodological power in order to consolidate all related domains, is the principal research area this thesis is addressing. In the context of this thesis, the concrete goal of *Ambient Intelligence* is to propose a packet of services, which would support the automatic determination of a personalised sensor-system. These services shall be accessible pervasively, through the objects people wear on, and an appropriate guidance for the exploitation of these services by non-competent users should be set.

The *Ambient Intelligence* strongly supports a vision of an information society, which corresponds to a long-term goal of EU Information Society Technologies Research programme. The aim is to create a society, which easily accesses the knowledge and the data, which is educated to use modern tools and various information sources and which is open for the newest scientific achievements.

1.3. Research objectives

The principal objective of this thesis is to create a methodology for an intelligent assistant development and to apply this methodology to solve self-care issues in the representative cardiologic domain. The specific goal from the cardiac domain is to provide any non-specialist user with a reliable sensor-system and related decision-making services. The selected sensor-system shall acquire the best possible personal ECG in order to enable a correct signal's interpretation that is a trustworthy diagnosis. To reach this goal, a large set of participating business services and all the possible errors sources need to be taken in account. These services identification, evaluation and orchestration as well as solutions for errors elimination need to be found out. The investigation for a solution that is able to handle dynamic contextual information, is capable to adapt to changing circumstances, and to provide accumulated knowledge is needed to be done. Thus, the main challenge is to merge parameters obtained from various resources together with personal data and sophisticated knowledge from the cardiovascular domain in order to provide personalised assistance.

For explicitness, we have divided the objectives of this research into two interconnected levels: *conceptual* and *application*, where each of them resolves into smaller research bits.

Research objective at the *conceptual* level

Propose a methodology, in software architecture terms, that would permit to develop ambient intelligent software that adapts to the user-specific context

- Propose an approach that allows automating sophisticated and personalised expert activities, which apply professional knowledge to process an individual case.
- Propose/Design the required tools that need to be used in such expert activities automation following the proposed approach.

- Create a generic architectural framework that allows developing domain- and user-specific ambient intelligence.

Research objective at the application level

Design a web services architecture that supports the automatic determination of a personalised sensor-system, which is embedded in smart garments, and which shall be used in self-care in order to allow a profane user to record himself a personal electrocardiogram (ECG) at anytime and anyplace.

- Elucidate important points of consistent signal processing and data analysis techniques selection and enhance already available calculations in order to meet home-based self-care requirements.
- Review the advantages and disadvantages of current wearable sensor-systems in order to identify requirements for personalised sensor-system design.
- Design an adaptable to the human morphology personal sensor-system.
- Assess a new design proposal on a set of volunteers presenting various characteristics in age, sex and morphology with respect of the computerised ECG analysis.
- Specify the general architecture to provide an intelligent business service for the selection of the most optimal and best suited light equipment providing the highest diagnostic accuracy ECG for the given individual.
- Demonstrate the orchestration of web services that implements the intelligent business service: “Optimal personal ECG recording system selection”.

1.4. Dissertation organisation

This dissertation manuscript is organised in two parts: **Part I** and **Part II**, which are enveloped with the **General introduction** and the **Conclusions and future work** chapters.

Part I: The first part consists in three chapters:

- Chapter 2 introduces the cardiovascular domain, biosignal processing and analysis techniques. It underlines current research studies that are related to self-care.
- Chapter 3 overviews current wearable sensor-systems used to obtain the electrical heart’s signal. It marks their advantages and disadvantages in self-care.
- In Chapter 4 we propose a new 3 sizes sensor-system prototype to be used in self-care. We then describe data acquisition and used data analysis methods and test this sensor-system prototype on a set of different morphology, age and sex volunteers.

All over Part I, we emphasise the heterogeneity and complex relations of contextual information and we propose ways to shape the context in order to handle this heterogeneity while seeking to solve a concrete sensor-system selection problem.

Part II: This part consists in four chapters:

- Chapter 5 overviews current information technologies that have an impact on creating a cognitive agent, which works as an intelligent assistant.

- Chapter 6 defines an expert activity that applies professional knowledge to solve an individual case. In this part an approach of how such activities can be automated is presented and tools that are needed to implement this approach are described.
- Chapter 7 describes our proposed general Data and Model driven Service Oriented Architecture (DM-SOA), which enables the construction of an intelligent assistant and then specifies this architecture to a cardiology domain to create an intelligent agent – Personal Cardio Assistant – that selects an optimal personal ECG recording system for each individual.
- Chapter 8 describes new self-care scenarios in cardiology domain and presents the implementation of a Personal Cardio Assistant via the complex orchestration of participating services.

The second part describes our proposed methodology for *Ambient Intelligence* software development expressed in terms of Data and Model driven Service Oriented Architecture (DM-SOA), which provides a framework for the production of context-aware intelligent business services.

Part I

Towards ubiquitous self-care in cardiology

Chapter 2 – State of the art in the electrocardiology domain

2.1. Introduction 25
2.2. Medical background 25
2.3. Electrical activity of the heart: the ECG signal 27
2.4. Information capture 28
2.4.1. The standard 12-lead ECG recording system 28
2.4.2. The Mason-Likar ECG recording system 30
2.4.3. Reduced lead set ECG recording systems for self-care 30
2.4.4. The vectorcardiography recording system 32
2.4.5. Discussion about ECG recording systems 33
2.4.6. Discussion about ECG information capture reliability 34
2.5. Information processing 37
2.5.1. ECG signal disturbances correction 37
2.5.2. ECG waves recognition, typification and delineation 37
2.5.3. ECG features extraction 39
2.5.4. Discussion 39
2.6. Information reconstruction 40
2.7. Information interpretation 43

2.1. Introduction

Ubiquitous computing in eHealth is often mentioned in the context of improving health-care with its promise of better patient care and lower costs. Certainly, ubiquitous systems are the cornerstone of telemedicine, vital signs monitoring, and ambient assisted living, which according to technology forecasts will be common lifestyle elements in the future society. This prospected vision is strongly supported by various European IST (Information Society Technologies) programs, which in years 2006 to 2013 set a special focus on early illness detection and encourage non-invasive health monitoring systems with flexible access for everybody, at any time and any location. IST approach is crucial within the domain of cardiac-care, where the ischemic heart disease is classified worldwide to be the first cause leading to death and is among top six causes of burden of disease *[Mathers'08]*. This thesis aims to contribute to the early detection of cardiac ischemic events by putting in practice decision support systems based on ubiquitous computing, which provide the citizen with an intelligent and self-adaptive support. The following sections are dedicated to overview the medical background, the available and applicable ECG recording and processing techniques and to introduce to current results of related research studies.

2.2. Medical background

Ischemia is a medical condition where a blood flow restriction occurs, generally due to various factors related to blood vessels and causes the damage or the dysfunction of the insufficiently irrigated body tissue. Cardiac ischemia is caused by the insufficiency of blood flow and so the lack of oxygen supply to the heart muscle *[Macfarlane'10]*. Coronary Artery Disease (CAD) causing cardiac ischemia is the most common type of heart disease. CAD occurs when the arteries that supply blood to the heart muscle (the coronary arteries) become hardened and narrowed due to build-up of material called plaque on their inner walls. The plaque, known as atherosclerosis, increases inside of coronary arteries and starts blocking the blood flow so reducing the much-needed oxygen supply to the heart muscle. Over time CAD weakens the heart muscle and contributes to heart failure, which may reveal itself through angina, various arrhythmias or even Acute Myocardial Infarction (AMI), commonly known as a heart attack. Angina (angina pectoris) is a medical term for a chest pain, which happens when the heart muscle getting not enough oxygen produces a by-product called lactic acid that builds up in the muscle and causes the pain. This pain is usually described by a bunch of symptoms, such as a discomfort, heaviness, pressure, fullness, aching, burning, or painful feeling in the chest area (but may also be felt in the shoulders, arms, neck, throat, jaw, or back), and unfortunately, is often mistaken with indigestion so disturbing acute infarction diagnosis and thus an early treatment. Symptoms of heart failure such as shortness of breath, cough, swelling of feet/abdomen, weight gain, and others most often begin slowly and are hardly noticeable without a careful examination. Some people have ischemic episodes without knowing about them. This silent ischemia can lead to heart attacks that occur without warning, and sometimes even without a pain. In fact, the heart failure does not mean that the heart has stopped or is about to stop as it might be thought. It means that the heart is failing to pump blood the way that it should to meet body needs, including needs of itself. When the heart does not get enough blood to operate properly and to respond to the body demands of more oxygen and nutrition, especially when a body strain appears, the electrical system responsible for the regular and coordinated contraction of the heart muscle might be affected as well and might lead to an irregular heart rhythm – arrhythmia. This can be a reversible situation – remove the strain and things return to normal, or irreversible – the damage permanently affects the electrical heart system causing cardiac Arrhythmias. The electrical impulses may happen too fast, too slow, or erratically, causing the heart to beat too fast, too slow, too early or irregularly and so disturbing the effective blood pumping. Most arrhythmias that cause short occasional abnormal awareness of heart beat (palpitations) even being merely annoying are more or less harmless. Others, though, lasting longer or occurring more often, may have a serious impact on the body as due to the abnormal heart activity the blood flow is significantly reduced (risking to form thrombus as well) and results in continuous oxygen and nutrition insufficiency (or even block), which leads to the damage of organs. This becomes life threatening if the brain, the heart or lungs are affected. That is exactly how a heart attack happens. A blood clot develops at the site of plaque in a coronary artery and suddenly cuts off most or all blood supply to that part of the heart muscle. Cells in the heart muscle begin to die if they do not receive enough oxygen-rich blood, causing a permanent damage to the heart muscle. If this situation is left untreated for a sufficient period of time it can cause a death -of heart muscle tissue (myocardium).

A worrying statistic that 60% of citizens, who suffered from a stroke, die or become dependent on others for help with their everyday living, indicates prevention as a high priority strategy *[Mackey'04]*. Normally, patients would have great chances to survive if the treatment would be delivered on time – *"the earlier treatment is given the better"*, though in western countries *"almost two thirds of cardiac deaths occur outside the hospital" [Rubel'05]*. Studies have shown that the most crucial element affecting the survival of patients having a heart attack is how quickly the arteries of the heart are re-opened using a thrombolytic medication or applying the surgery called Percutaneous Coronary Intervention (PCI), known as coronary angioplasty. In order to respond immediately, the symptoms shall be recognised and adequate actions have to be taken as early as possible. Citizens are usually taught to recognise myocardial infarction, however in real situations symptoms are often interpreted incorrectly and only a small amelioration shortly after the course has been reported *[Rubel'05]*. Then, the objective of eHealth and especially of pHealth (personalised Health), is to equip citizens with more efficient and responsive tools that would assist in decision making enabling citizens themselves to constantly undertake the primary health provision so reducing the time between the detection of first ischemia symptoms and the appropriate treatment.

Typically, patients with acute chest pain and/or heart rhythm abnormalities potentially of ischemic origin presenting at emergency departments situations are assessed with three principal tools: *"the history of the event, the 12-lead electrocardiogram (ECG), and cardiac enzymes and other serum makers of myocardial injury" [Pollehn'02]*. From these three tools only the ECG is suitable to sustain daily non-invasive self-care routine performed by the citizen himself, as it is a non-obtrusive source of instant and objective information *[Fayn'10]*. The morphology of ECG recording (even of 10 seconds long) contains considerable diagnostic information about the heart's activity and so a signal analysis can provide evaluation that supports the decision making. *"ST-segment deviation will evolve in ECG leads oriented along the axis of the current of epicardial injury after an acute coronary occlusion occurs, and thus, ST-segment analysis has become an important noninvasive tool for diagnosing acute myocardial infarction and for deciding on and evaluating the efficacy of therapy." [Sejersten'07]*. Then, a potentially helpful tool would be a small smart personal device *[Rubel'05]* having the capability to record discretely anywhere and at anytime a diagnostic quality ECG and fitting into our life as naturally as blood pressure monitors have recently done. This device should have an embedded intelligence that is capable of analysing the ECG "on-board" and of comparing it with a reference ECG. Likewise, it should be capable of making a diagnostic recommendation and, when needed, of putting a patient in a contact with the appropriate medical service. The objective of this thesis is to support the success of this challenging but potentially very rewarding aim by elucidating important points of consistent techniques selection and enhancing already available calculations in order to meet home-care requirements. Indeed, the use of ECG in such personalised tools shall be based on mature reflections of the whole system, which can be constructed from a variety of different, though, related methods of ECG recording and processing.

2.3. Electrical activity of the heart: the ECG signal

The body needs for oxygen-rich blood are met with the rhythmic and sequential contraction of atria and ventricles performed by the cardiac conduction system in order to pump the blood throughout the body. The graphical representation reflecting this continuous repetitive electrical heart activity is called an electrocardiogram or an ECG for short (Fig.I.2_1). The electricity in biological systems appears due to electrochemical mechanisms represented by movements of ions (sodium, potassium, and calcium) and induces the contraction of heart muscle in order to function as a blood pump. *"Cardiac and muscle cells are able to change the permeability of their surface membranes to ions, and thereby to generate a stereotypical depolarization-repolarization sequence known as an action potential." [Macfarlane'89].*

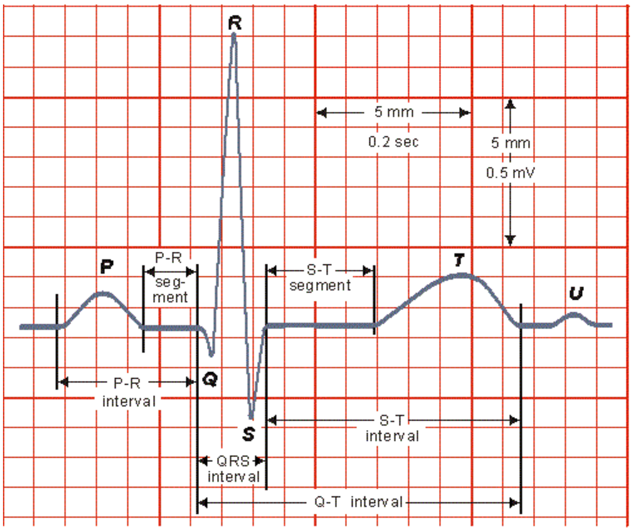


Fig.I.2_1 ECG beat corresponding to one electrical cycle

The flow of the electric currents is detected with a pair of electrodes placed on the skin, where amplified potential differences that are caused by the depolarisation and the repolarisation of the heart muscle produce over time one electrocardiographic lead (the simplest form of an ECG). A typical ECG tracing of a cardiac cycle, or heartbeat, consists of a P wave, a QRS complex, and a T wave, corresponding respectively to the atrial depolarisation, and to the ventricular depolarisation and repolarisation. The normal durations are about 80-90 ms for the P wave, 90-100 ms for QRS, and 300 ms for the ST-T phase.

An accurate recording of these tiny electrical changes is an initial step towards a correct signal interpretation that is to a trustworthy diagnosis. Many factors, though, need to be considered in order to obtain a diagnostic quality ECG. The amplitude and form of the recorded wave may be influenced by the myocardial mass, the distance between the electrode and the myocardium, the thickness and properties of the intervening tissues, and the net vector of depolarisation. This becomes crucial in case of computerised ECG diagnosis, which could lead to automated user guidance. The digital management of these influences in order to obtain an accurate ECG is complicated and currently remains an open multidisciplinary research area. We categorise the concerned influences in four pretty large groups: information capture, information processing, information reconstruction and information interpretation, which we'll overview in the following sections.

2.4. Information capture

Usually, more than two electrodes are used to record an ECG in practice. Their various combinations result into a number of pairs, known as leads. Leads can be bipolar and unipolar. *"A bipolar lead measures the potential difference between two points – hence the term bipolar". [Macfarlane'89]* A unipolar lead represents the potential variation at a single point according to a so called "central terminal", which is a sum of potentials in several points and is usually constant during the cardiac cycle. Different types of ECGs can be referred to the number of leads that are recorded, for example 3-lead, 5-lead, 12-lead or even 80 (and more)-lead ECG.

2.4.1. The standard 12-lead ECG recording system

Classical lead theory, developed by Einthoven and Goldberger during 1889-1942, consist in twelve leads: three main leads (I, II, III), the so called limb leads, three augmented limb leads (aVR, aVL, aVF) that are derived from the same three electrodes as leads I, II, and III, and the precordial leads V1, V2, V3, V4, V5 and V6 recorded from six electrodes on the chest. All twelve leads are obtained using ten electrodes in total, from which four are placed on the limbs and six on the thorax (Fig.I.2_2a), resulting in 8 recorded leads: I, II, V1 to V6. The 4 additional leads (III, aVR, aVL, aVF) of the standard 12-lead ECG, which is the electrocardiogram usually analyzed by the physicians, are then derived from the recorded leads by a set of linear combinations:

$$III = II - I$$
$$aVR = -(I + II)/2$$
$$aVL = I - (II/2)$$
$$aVF = II - (I/2)$$

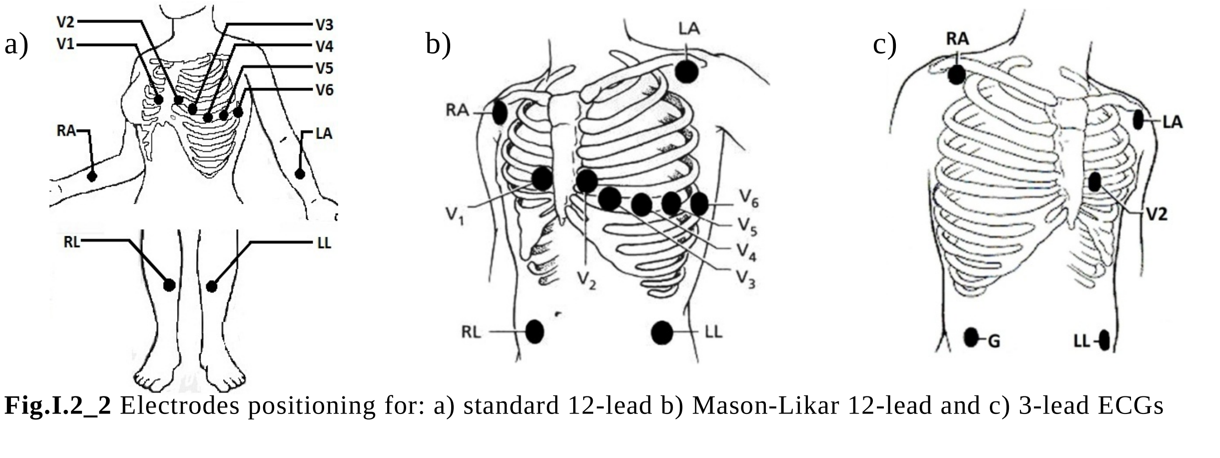


Fig.I.2_2 Electrodes positioning for: a) standard 12-lead b) Mason-Likar 12-lead and c) 3-lead ECGs

The positioning of the three main leads I, II and III of the 12-lead ECG using three electrodes placed on the right arm (RA), left arm (LA) and left leg (LL) was firstly introduced by Einthoven and therefore is called the Einthoven's triangle. Lead I is bipolar and obtained placing the negative electrode on the right arm (RA) and the positive electrode on the left arm (LA). Lead II is bipolar too with the negative electrode in position RA and the positive electrode on the left leg (LL). The whole set of six limb leads and augmented limb leads form the basis of an hexaxial reference system, which is used to calculate the heart's electrical axis in the frontal (vertical) plane. The other six leads of the 12-lead ECG are called precordial leads and as they are close to the heart, they do not require augmentation. Precordial leads are unipolar since the measurement is performed between one positive electrode on the chest (V1, V2, V3, V4, V5, or V6) and the so called Wilson central terminal *[Wilson'34]*, which sums the potentials of electrodes at the right and left arms and at the left leg and is assumed to be almost constant. The obtained leads are named according to the name of the electrode used: V1 is referred to be a right precordial lead and V2, V3, V4, V5, and V6 are referred to be left precordial leads. V1 and V2 electrodes are put at the **4th** intercostal space just right (for V1) and just left (for V2) of the sternum. Then **V4** is placed at the palpable apex or at the 5th intercostal space in the midclavicular line. V3 is halfway between V2 and **V4**. V5 and V6 are in the same horizontal plane as **V4**, though V5 is in the anterior axillary line and V6 in the mid axillary line. These remaining six precordial leads view the heart's electrical activity progressively in the horizontal plane. The QRS complex should be negative in lead V1 and positive in lead V6, so the gradual transition from negative to positive between leads V2 and V4 can be seen. Finally, the last from the whole set of 10 electrodes is placed by convention on the right leg (RL), although, in theory it can be placed anywhere on the body. It was formelly called ground electrode and it is nowadays used by amplifiers to reduce common mode voltages. This defined placement of 10 electrodes allowing the recording of the described 12 leads is now a worldwide used standard as it provides an anatomical perspective of the heart from a vantage point helping to identify the area of several heart diseases, such as acute coronary injury. The ECG obtained using this arrangement is labelled as a standard 12-lead ECG.

The standard 12-lead ECG is commonly recorded in hospitals under the supervision of skilful nurses, cardiac technicians or heart specialists, who place electrodes in their correct anatomical locations and with the use of a medical equipment called electrocardiograph that simultaneously records all leads. The requirements for electrocardiographs used in hospitals are described in several international standards (i.e. Medical electrical equipment BS EN 60601-2-51:2003 *[BS'03]*). Most electrocardiographs are designed to record during 10 sec. the standard 12-lead ECG from resting, supine patients. The signal recording from moving or shaking patients may produce erroneous 12-lead interpretation results. Therefore, medical staff is asked to ensure that the patient remains motionless and relaxed during the recording. Initially, ECGs were printed on specially lined paper and afterwards analysed by cardiologists. Now the digital format is in use and facilitates not only the storage of the ECGs, but also permits initial ECG pre-processing via a set of algorithms designed and programmed for that purpose. The international ISO standard for digital rest ECG data exchange is the SCP-ECG protocol, which includes not only the digital signals, but also, if available, the measurements and the interpretation results, and all the ECG recording metadata that are relevant for its interpretation *[ISO'05]*.

We'll enlarge the knowledge on signal processing steps and available techniques in the following section (section 2.4). This embedded into the electrocardiograph or external software can, actually, even assist cardiologists by giving automated signal interpretation, which we'll talk more about in 2.6 section. Currently, the printed copy of digitally saved ECG is made only for convenience.

2.4.2. The Mason-Likar ECG recording system

The same electrocardiograph used in hospitals is also employed to register another type of 12-lead ECG, which is called an exercise or stress ECG due to conditions it is recorded in. The patient is asked to do some activity under strict medical control, typically, to ride a special bicycle or to do some simple exercises, and a 12-lead ECG is recorded at the end of different workload stages. Working muscles need more oxygen to keep exercising and so the heart starts beating faster to pump more oxygen-rich blood. This cardiac stimulation can show abnormal blood flow, or changes in the heart's muscle tissue and so its analysis complements the interpretation of the standard 12-lead ECG. The difference between the standard ECG and the stress ECG recordings is not only the patient activity level, but also electrodes placement. The four limbs electrodes (RA, LA, RL, LL) are replaced from limbs on the torso aiming to reduce possible signal distortion caused by the motion of arms and legs, which generates signal artefacts and for muscle tremor. This modified electrodes placement is known under the name of Mason-Likar as it was suggested by R. E. Mason and I. Likar in 1966. The right and left arm electrodes are placed according to their sides at a point in the infraclavicular fossa medial to the border of the deltoid muscle, 2 cm below the lower border of the clavicle. The left and right legs' electrodes are placed at the iliac crest of the appropriate side (Fig.I.2_2.b). The ECG obtained using the Mason-Likar lead system is not really equivalent to the standard ECG as it tends to shift rightward the QRS axis, to reduce the R wave amplitude in lead I and aVL, and to significantly increase the R wave amplitude in leads II, III and aVF *[Papouchado'87]*.

2.4.3. Reduced lead set ECG recording systems for self-care

Recently, the use of reduced lead set ECGs started to gain importance as there is a growing demand for developing personalised and non-hospital-based care systems. These so called self-care systems are intended to be used by patients themselves either continuously, or on demand during the patients' daily life, for example, at home, at work, or on holidays. Obviously, these special conditions generate new challenges, such as: simple electrodes placement allowing patients themselves to put on the electrodes-system, motion artefacts management and convenient wearing for prolonged time if needed, what in total demands to reduce the number of electrodes. The classical approach of multiple wires connected to sensors spread across the body would limit the patient's activity and would reduce the level of comfort. So self-care systems, naturally, are designed to have fewer electrodes, what allows them to be placed and managed easier by the patient himself, to be less obtrusive and to not interfere significantly with human movement – so if needed to be worn continuously. Current fabrics or prototypes of such electrode-systems contain 2 to 5 active electrodes and may record 1- to 4-lead ECGs, typically, 1-lead ECGs or 3-lead ECGs. The quantitative quality of the obtained signal is usually satisfactory, though the diagnostic information content and the successful interpretation of the ECG remain main problems when seeking to employ these systems in practice. The electrode-systems capable to record an 1-lead ECG are, thus, rather targeted to sportsmen or military to evaluate their training efficiency according to the changes in their heart rhythm than to be used in a medical context, even if in some cases these systems might be very helpful in detecting occasional arrhythmias.

In reality, the only representation cardiologists are trained to use for an accurate heart analysis and the estimation of the medical condition is the standard 12-lead ECG. This results that the reduced lead ECG should be transformed into a 12-lead ECG to be interpreted. Indeed, mathematical signal reconstruction methods may be applied in order to derive a so called synthesised or reconstructed 12-lead ECG from a couple of recorded leads. We'll discuss in more details the information reconstruction methods in section 2.5.

ECGs having 3 leads were demonstrated by various scientific studies as bearing sufficient spatio-temporal information for the 12-lead ECG reconstruction and as containing enough diagnostic information to detect transmural ischemia, infarction and various types of arrhythmias *[Rubel'05], [Fayn'07a], [Fayn'11], [Green'07]*. The 3-lead ECG recording can be obtained with at least 4 active electrodes, though their configuration on the body varies depending to the scientific studies. *"Ever since the beginnings of electrocardiography there has been controversy about how many electrodes are needed to faithfully capture the electrical activity of the heart and where on the body surface they should be placed." [Kors'02]*.

Today, we can distinguish between two main approaches for designing self-care electrode locations: a) optimised 12-lead ECG reconstruction and b) ease at electrodes placement, which contributes to better ECGs reproducibility. The first approach builds new knowledge upon already existent fundamental knowledge elaborated and actively used for more than 100 years. Electrode-systems that follow this approach keep electrodes in their standard locations, thus keeping the same notion of leads as in a standard 12-lead ECG, and try to remove the ones which are less important in signal reconstruction and interpretation. The research studies related to this approach evaluate different methods of 12-lead ECG mathematical reconstruction from the reduced lead set and assess the diagnostic performance of different leads subsets. A leader among this approach followers is the European EPI-MEDICS project that was carried out between 2000 and 2005 *[Rubel'05]*. The project performed a comparison study of 10 different subsets of 3 leads, where the information reconstruction and diagnostic accuracy for acute myocardial infarction and for transmural ischemia were assessed *[Simon-Chautemps'04]*. Two subsets composed from lead I, lead II and lead V2 or V3 resulted to contain nearly as much diagnostic information as the standard 12-lead ECG and luckily were also best in the 12-lead ECG reconstruction. This project has set the priority for the subset of leads I, II and V2 as it is easier to place the V2 electrode in its correct position – 4^{th} intercostals space just left to the sternum, than to locate the V3 electrode. The importance of the V2 lead for a reduced lead set design was also confirmed by other ECG reproducibility studies: *"we consider electrodes V3-V6, and in particular V4 and V5, most likely to be misplaced, especially in women and the obese" [Kors'02]*. Practically the same conclusion came from another European project, called WEALTHY (2002-2005) *[Paradiso'05]*, which has dwelled on a 4 leads subset, adding lead V5 to the set of I, II, V2 leads. However, in daily situations, where the patient records the ECG by himself or wears the electrode system all day long, the body movement interferes with the recording and can distort the signal. Therefore, the EPI-MEDICS and WEALTHY projects as well as other studies working with reduced lead set ECGs prefer the torso placement for all needed electrodes. This attitude does not affect the precordial electrodes (V2 for instance), but it requires to take off limbs' electrodes, what naturally attempts to use the Mason-Likar positioning for the recording of leads I and II (Fig.I.2_2.c).

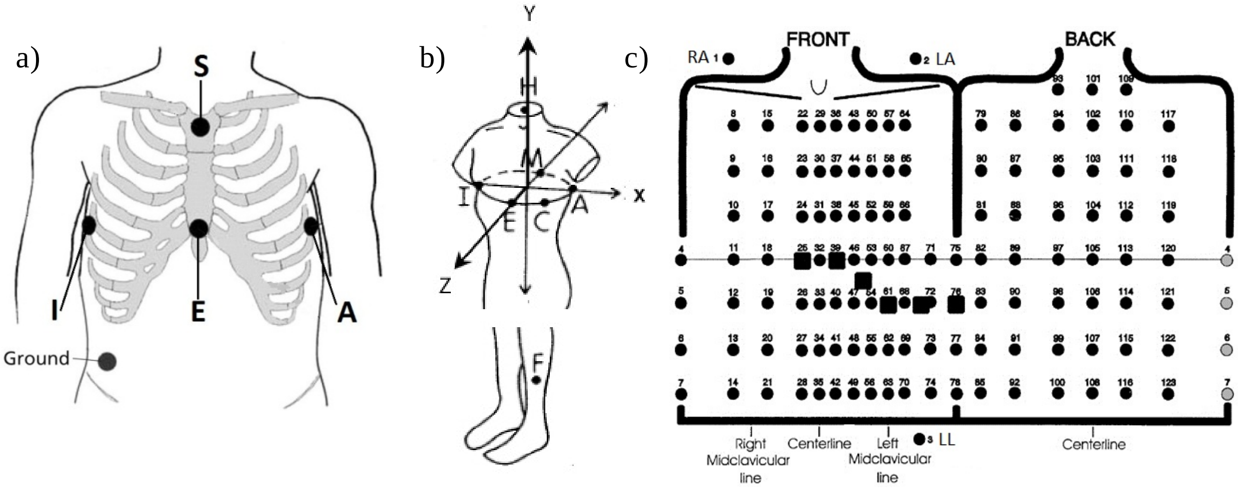


Fig.I.2_3 Electrodes positioning for: a) EASI 3-lead ECG b) Frank XYZ-lead VCG *[Macfarlane'89]* c) BSPM

The system called EASI (Fig.I.2_3.a) represents the best the second approach of reduced lead set ECG recording, where the easier electrodes placement is set as a priority. This electrode-system has years of research behind. The EASI placement scheme follows clear anatomical body marks, where each letter stands for one active electrode: *"the E electrode on the lower extreme of the sternum, the A and I electrodes in the left and right midaxillary lines, respectively, at the same transverse level as E, and the S electrode on the sterna manubrium" [Sejersen'06]*. This system has clear advantages of easy (therefore pretty accurate) and rapid electrodes application on body, which is essential for unskilful users or in stressful situations for paramedics. It also allows avoiding the need to determine intercostals spaces and putting electrodes under/on the breast, what in women with large breasts deflects the correct electrodes' placement *[Macfarlane'03]*, *[Rautaharju'98]*. Moreover, accurate electrodes placement has a significant importance in ECGs comparison over the time, therefore clear anatomical sites greatly reduce the possible error related to electrodes displacement. However, the 3-lead ECG obtained from these non-classical electrodes' locations and the standard 12-lead ECG do not really possess a common methodical background.

2.4.4. The vectorcardiography recording system

The basis for vectorcardiography (VCG) is the modelling of the electrical heart's activity as a cardiac dipole that was first introduced by D. Waller in 1887 and thus the possibility to represent its behaviour as a vector. The dipole is represented by a vector whose components measure changes in the dipole's magnitude and orientation. Ideally, the heart vector H can be constructed from its projections onto three orthogonal axes: Hx, Hy and Hz. There, a number of different lead systems were proposed till the understanding came *"that while electrode systems might appear to be geometrically orthogonal, the equivalent lead-vector orientations were far from orthogonal. Such systems then became known as "uncorrected" orthogonal-lead systems." [Macfarlane'89]*. In 1956, E. Frank was first to publish details upon a truly corrected orthogonal lead system, which took in account distortions caused by the boundary and internal inhomogeneities of the body. His proposed recording system consists of 7 electrodes: A, C, E, I and M placed at a level corresponding to the electrical centre of the heart, H placed on the back of the neck and F used as the ground electrode placed on the left leg (Fig.I.2_3.b). The Frank's electrode network and the associated resistors matrix used to equalise the vectors' lengths are capable to provide three orthogonal and normalised leads: X, Y and Z. Actually, the total information content of the VCG is not greater than the content of the12-lead ECG and is more or less similar to the content provided by leads V2, V6 and aVF in the 12-lead system, though it is provided in an orthogonal (corrected) form, which allows displaying the dipole's behaviour in the rectangular coordinate system using an oscilloscope or a plotter. This display is called a spatial vectorcardiogram and it allows to much easier observe the direction of the heart vector and to analyse the progress of the activation, which is especially important in initial and terminal parts and may have clinical importance. The three vector's components can be also displayed in function of time, what is called a scalar vectorcardiogram. The vectorcardiography (VCG) clinical investigation method is claimed to be *"less specific but more sensitive than the standard 12-lead electrocardiogram (ECG)" [Scherer'92]*.

The Frank system, at present, is the most common VCG system used in clinical practise, throughout the world. However, probably due to a little more complicated leads recording system (placing M electrode on the back of the patient might be inconvenient for elderly or patients after an accident or a surgery) or because of additional physical modelling knowledge needed to interpret it (what means supplementary hours in the medical educational program). Nowadays, the number of recorded VCG's represents less than 5% of all recorded electrocardiograms. This, however, could be changed if electrodes application following the EASI scheme became popular.

2.4.5. Discussion about ECG recording systems

There are many recent research works based on the EASI electrode-system, which contribute to the research field by designing complex mathematical methods for standard 12-lead ECG synthesis from the 3-lead EASI ECG *[Feild'02]*, *[Horacek'08]*. Their results show that EASI derived 12-lead ECGs are comparable with standard 12-lead ECG, even in cases of wide complex tachycardias *[Horacek'00]* and also of myocardial ischemia *[Sejersten'07]*, *[Sejersen'06]*, *[Horacek'02]*, though a list of differences go together with these affirmations. The American Heart Association in their recent recommendations for the standardisation and interpretation of ECGs already includes EASI ECGs as known alternative to standard ECGs and acknowledge its advantages of placement, but underlines that *"although synthesized ECGs that use the EASI lead system may be demonstrably adequate for some purposes, such as monitoring rhythm, they cannot be considered equivalent to standard 12-lead recordings or recommended at present as an alternative for routine use" [Mirvis'07]*. Additionally, experiments on body-surface mapping demonstrated that *"no single dipole can be "equivalent" in the sense of accurately reproducing the body surface observations" [Taccardi'63]*, what staggered the fundamental idea of VCG modelling, and so the one of EASI. These findings encouraged the design of multi-dipole models that, in principle, divide the heart muscle into regions, where each is imagined as a dipole. However, the main weakness of multi or single dipole models comes from their inability to measure their properties experimentally. Potential difference measurements are made at a distance from the source, on the body surface. Then, in order to find a mathematical way to express the epicardial excitation, measurements on the whole thorax surface were started to be made. Electrode-systems from 80 to 242 electrodes called Body Surface Potential Maps (BSPM) are used for that purpose (Fig.I.2_3.c) and, at present, they are the only ones capable to collect all the electrical information at the surface of the body *[Kors'02]*. Many technical difficulties were met at first, though, and only recently, with the help of advanced technologies, the sequential recording and the processing of large number of leads were automated, so encouraging resuming the topic studies in a number or research centres. The data obtained from BSPM are displayed in various manners that include the use of colours, contours or shading between contours, and nonlinear scaling. There, the aim is usually to detect patterns of abnormal spread of excitation over the thorax and to connect them with clinical diagnosis. These efforts demonstrated that many details about underlying events can be recognised from BSPM. *"It is known that body surface potential maps (BSPMs) contain diagnostic information not easily retrieved from the standard 12-lead electrocardiogram (ECG)." [Kors'02]*. However, BSPM correct interpretation and classification present a number of difficulties and has not succeeded in getting accepted by clinicians. Besides that, the measurement of surface potentials has the conceptual disadvantage of being simply a reflection of a surface quantity, rather than a model of what really happens within the muscle. There is still a need for conducting basic research jointly with broader clinical tests in order to find a mathematical model explaining the heart's activity and, finally, to fill the knowledge gap and to establish (if possible) vectorcardiography based methods, such as EASI, which would allow then to reliably reconstruct the 12-lead ECG for cases, where it is demanded like home-care. The European Commission is currently setting this aim in the scope of the 7th Framework Programme (2007-2013) and supports the development of patient-specific computer models and their application in personalised and predictive healthcare by connecting various Virtual Physiological Human projects *[Ayache'05]*, *[Hernandez'09]*.

Hereof, in this thesis context we give more credits to the first approach, which was chosen by the EPI-MEDICS project and which designed a reduced lead set aimed at providing the most reliable 12-lead ECG reconstruction and acute infarction detection, than to the one based on easy electrodes positioning. However, the advantage of the second approach – the easy electrodes placement contributing to the ECGs' reproducibility, needs to be taken into account when we talk of a final self-care system. Actually, the related literature underlines that the ECG signal morphology is very sensitive to electrodes positioning *[Sejersten'07]*, *[Lateef'03]*, *[August'58]*, and that not only the serial comparison performed between ECGs recorded distally may lead to serious errors, but also an ECG recorded from proximal electrodes' positions can lack of sensitivity/specificity for various cardiac pathologies. The greatest advantages of the 3-lead ECG system proposed by Rubel within the EPI-MEDICS project, are that, being pseudo-orthogonal, it allows to retrieve almost all the spatiotemporal information of the electrical activity of the heart and, by recording the 3 standard leads (I, II, V2), which is a subset of the 8 recorded leads of the 12-lead ECG, it is already well-known for ECG interpretation by the physicians and for a reliable electrodes positioning by nurses and recording technicians.

2.4.6. Discussion about ECG information capture reliability

Misplacement errors even among specialists are still common, even though a clear description exists on the 10 electrodes' placement for the standard 12-lead ECG *[Rautaharju'58]*. A proper electrodes' positioning requires good anatomical knowledge and time to find accurate electrodes' locations on the patient's body. Recently, R. Rajaganeshan in six hospitals (mainly in Great London) identified that nurses, general practitioners (GPs) and cardiologists involved in recording ECGs are significantly less trusted to place the chest electrodes in the correct positions than cardiac technicians *[Rajaganeshan'08]*. Surprisingly, cardiologists performed the worst and nurses together with GPs were much worse than cardiac technicians. K. McCann in his study in one emergency medicine department in Australia detected that the mean displacement of the electrodes was 13.5 mm mean in the vertical direction and of 16.5 mm in the horizontal direction *[McCann'07]*. In paired measurements these numbers gave more than 25 mm mean displacement. M. Sejersten performed a study on the direction of misplacement. *"The directions of the misplacement were divided into 5 groups: (1) above, (2) below, (3) correct placement, (4) left, and (5) right." [Sejersen'06]*. She and her co-workers demonstrated statistics that for 20 patients just 9% of all 10 electrodes for the standard 12-lead ECG recording were placed correctly, 42% were misplaced slightly below, 23% slightly above and 13% to the right and left. They saw a tendency to misplace the V1 and V2 electrodes slightly above the correct location in 60% of the cases and to misplace the V3-V6 electrodes slightly below the correct location in 59% of the cases. Finally, they found *"that routine placement of precordial electrodes for pre-hospital ECG recording by paramedics is associated with mean 37-mm misplacement most often downwards" [Sejersen'06]*. This error is not greater for female than for male and does not really correlate with body mass index, even though it was expected so. That the misplacement's distance and direction are, indeed, worrying was also shown by Hadzievski and his team while they were testing a novel portable ECG recording system in 2004 *[Hadzievski'04]*. This research team made several experiments on deciding the proper positioning of their device and found that device positioning is crucial. During the experiment, device was misplaced for 1 cm, 1.5 cm, 2 cm and 3 cm with the respect to the original position in both horizontal and vertical directions and the corresponding ECGs were recorded. The results showed that 1-1.5 cm displacements had no visible errors compared to the original position ECG, while 2 cm displacement produced some errors in T wave amplitude and, finally, 3 cm displacement produced significant errors in QRS complexes. The results were even more sensitive to the variation of the device orientation angle. The device lower corner was moved 1 cm and 2 cm to the left and to the right. This misplacement of 1 cm produced some noticeable errors in T wave amplitudes, while the misplacement of 2 cm caused significant errors in QRS complexes. QRS complex could be affected by alternatively placed electrodes (closer to the heart that it is done for standard 12-lead ECG recording). These differences could be at the origin of false-negative and false-positive infarction determination. Evidences from these and several more studies *[Wenger'96]*, *[Bartosik'95]*, *[Donnelly'06]* confirm that clinical electrode's mal-positioning of 20-30 mm influence the morphology of the signal and so can change its interpretation. Appearing or disappearing waveform features, that is changes in signal's patterns: the number of waves, the number of peaks in the wave, a sign of the peak, and ranges of amplitudes, also the ratio between amplitudes peaks or a peak and a baseline may happen due to the difference between the needed and actual recording systems and can affect the diagnostic accuracy. *"Placing precordial electrodes accurately in the anatomically defined positions is essential for securing a correct 12-lead ECG interpretation, because waveform duration and amplitude measurements can be significantly altered if they are not." [Sejersten'07]*.

The quality of recorded body's electrical signal using electrodes is highly influenced not only by correct electrodes locations, but also by heterogeneities of the impedance of the torso tissue. That refers to the impedance of skin, muscles, bones, fat, blood vessels, nerves, that is to the morphology of person. The differences between two ECGs obtained a couple of days', months' or years' interval may, actually, be caused by changes in the person's morphology: fat level, skin preparation, body temperature, and not by changes of the pathology of the heart muscle. In Fig.I.2_4 we have summarised all the possible morphological sources making influence on the ECGs signal's quality and as a consequence, on the interpretation, and categorised them into primary and secondary sources.

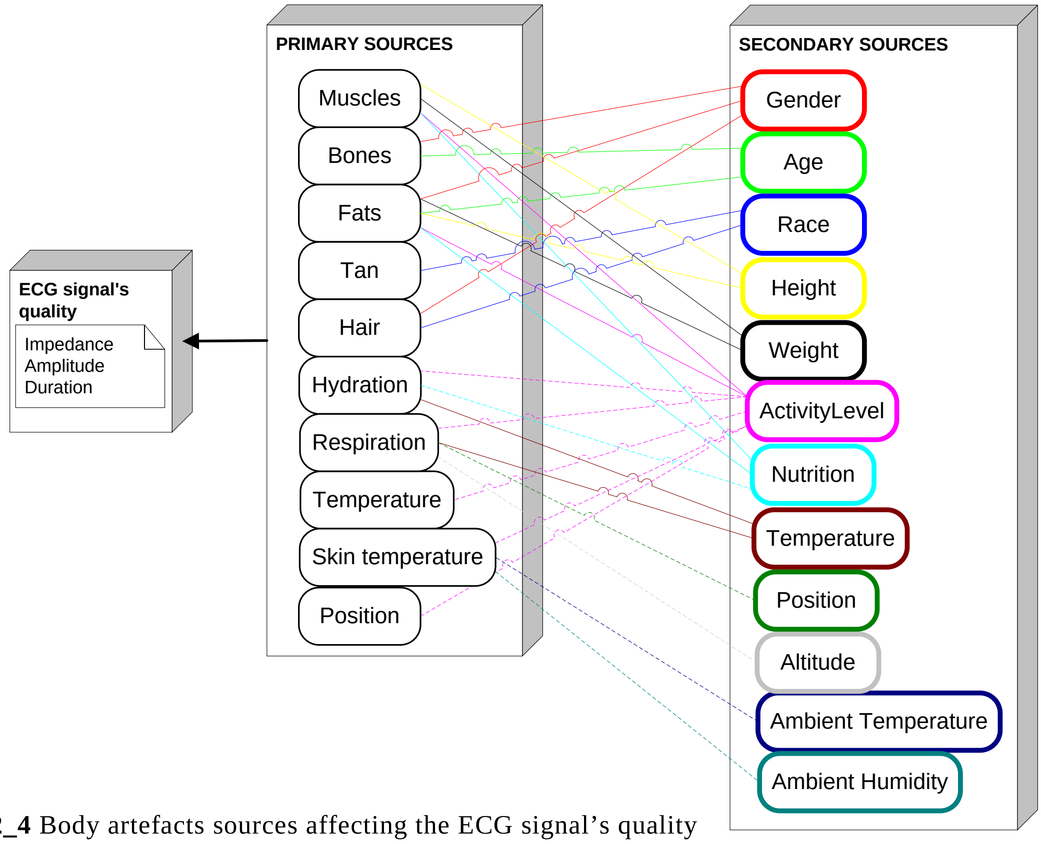


Fig.I.2_4 Body artefacts sources affecting the ECG signal's quality

In general, it is known that in average the percent of fat in a man is roughly 10 percent lower than in a woman, consequently, the impedance is naturally lower for men than for women. Though, fat's percent for the same person varies with time for many reasons: the age of person, heredity, or diet make people gain and lose weight. All these characteristics influence also a person's skin. The accurate measurement of the heart's electrical signal is often dependent on the interface between the body tissue and the electrodes. The impedance of the skin is generally the largest component in the overall "contact" impedance *[McAdams'96]*. A naturally darker skin colour (also a tanned skin), hairiness, lack of humidity increase the impedance, while a sweaty or a creamy skin makes it lower. This is the reason why in hospitals the state of skin is carefully assessed by a specialist and if needed it is individually prepared before electrodes are placed. That preparation could include a hair shave, a degreasing with the help of alcohol or other soak wipe, a gel application on the skin and even a light abrasion for better conductivity *[McAdams'94]*. Furthermore, the respiration, body dehydration and body temperature affect the impedance in an even less predictable way. The person's posture/position can also alter the signal's amplitudes and axes *[Mirvis'07]*. Historically, the supine position is used, though additional calculations and predictions must be defined for laying, sitting or standing positions.

To the best of our knowledge there are no rules or predefined methods able to detect the influence of electrode misplacement on the ECG signal's diagnostic content, except for the interchange of the left and right arm electrode positions. In addition, inter-subjects' and intra-subjects' variability is a source of context-sensitive information, which can make an impact on the ECG signal morphology and interpretation. In hospitals, a medical personnel is greatly encouraged to take routine and up to date training. While, home-based care systems call for a systematic approach to different human's body contexts, their classification and evaluation, that is for personalised procedures / algorithms capable to manage this complex data and to allow the recording of diagnostic quality ECG by the patient himself. New medical devices and services are required to evaluate particular citizen's conditions anywhere and anytime and to automate the adjustment of the sensor-system according to the specificities of the subject's morphology. In home-based care systems placing sensors together with input/output modules at different thorax locations to record a diagnostic quality ECG is a key responsibility of the sensors system and services that comes together with it. Therefore different wearable sensor systems embedded in some garment or accessories are presented to replace a number of separate electrodes and ensure comfortable usage of the medical device. Though, due to a large number of nonconformity sources and a lack of clarification and standardisation a creation of a personal sensor system remains a big challenge.

2.5. Information processing

An electrical heart signal captured via electrodes placed on the skin and adequately sampled (minimum 500 samples/second, as recommended by AHA and SCP-ECG) through appropriate analog-to-digital converters (resolution: minimum 5 microvolts, 11 to 16 bits), should follow a rigorous procedure of signal analysis in order to obtain useful physiological and clinical information out of it, which will lead to diagnostic ECG classification. This further analysis is performed in a series of steps, where each step is based on a distinct methodological background expressed via diverse data processing algorithms. These steps include: (1) ECG signal filtering, (2) ECG waves recognition, (3) ECG signal typification, (4) typical/median beat determination, (5) waveform delineation, (6) base line correction, (7) global measurements computation and (8) lead by lead measurements computation.

2.5.1. ECG signal disturbances correction

Indeed, different kind of disturbances, such as power-line interference, baseline wandering caused by electrode polarisation drift, muscle movement or respiration caused artefacts, sudden baseline shifts, spikes, or amplitude saturation may appear during the recording and so may influence the quality of the recorded data. The responsibility to manage the level of these disturbances or to eliminate them if possible is committed to the signal quality control module. This signal processing step may result with a demand to record again the ECG signals, or in correcting some of above listed artefacts. The corrections are mostly done using digital filtering methods *[Mortara'77]*. Low-frequency noise, usually caused by respiration, brings the deviation of the signal above and below the baseline, while muscle tremor or radiated and capacitively coupled electromagnetic interference result in a high-frequency noise *[Zywietz'90a]*, *[Zywietz'90b]*. In 2007, the American Heart Association published recommendations in agreement with the ANSI/AAMI standard *[ANSI'01]*, which continue to advise a low-frequency cut-off of 0.05Hz for routine filters with a possibility to extend the cut-off till 0.67Hz for linear digital filters with zero phase distortion, and a high-frequency cut-off of at least 150Hz with extension to 250Hz for children' ECGs *[Mirvis'07]*.

2.5.2. ECG waves recognition, typification and delineation

The ECG waves' recognition, typically, steps after the signal is filtered. It aims to detect all beats (P-QRS-T complexes) in the signal. For instance, a typical 10 seconds ECG record obtained in hospitals on the patient with a steady rhythm would contain 9-10 beats. Different strategies might be chosen or combined using techniques of QRS complex detection and P wave detection, having in mind that if a QRS happens, the associated T wave showing the mandatory repolarisation of the ventricles happens as well. Almost all programs nowadays use so-called multiple lead algorithms *[Kors'86]*. These algorithms transform simultaneous leads into a detection function, which enhances waves in comparison to other parts of the signal and so helps to increase the detection rate. One of the most commonly used transformations is the computation of the spatial velocity of the VCG or of a similar derived function for the 12-lead electrocardiogram *[Tighiouart'03]*, *[Rubel'86]*. There, then, different thresholds and heuristic rules are applied to detect the QRS complexes and the P waves *[Swenne'73]*, *[Klusmeier'76]*. The detection of the P waves is one of the most difficult parts in automated ECG analysis. Problems may arise due to low amplitude, variable morphology and varying time location of P waves. Also, P waves may be hidden in T waves, so can be invisible in the ECG. The failure in detecting P waves might largely influence rhythm interpretation programs *[Willems'72]*.

When the signal is “cut” in beats, a classification, based on morphological appearance of beats, is performed. The aim of this classification is to obtain the beat, which represents the whole signal the best. This classification is also called wave typing or “typification”. All beats are compared with each other and the ones which resemble among them are classified together. Different strategies can be used to compare the beats. Wolf and Rautaharju *[Wolf'72]* apply the clustering procedure based on 3 measurements (QRS duration, the preceding RR-interval and maximal amplitude). The AVA program *[Willems'72]* selects 8 measurements (amplitude of R and S waves in leads X, Y and Z, QRS duration and spatial magnitude) and performs a comparison with corresponding median values. Certainly, a simple correlation and RMS (Root Mean Square) error can be applied as well. The dominant class (the one which has most beats in it or the one selected with the help of some heuristics if there are a couple of classes that have an equal number of beats), then, shall be identified.

The dominant class is the one of a particular interest as the representative beat will be selected out of it by many programs. Though, here again different signal processing computer programs can use different strategies. Some will make the median from all beats in the dominant class and that “fake” beat will stand as the representative beat. The Lyon program selects the original beat that is most similar to the median beat *[Arnaud'00]*.

Finally, when the typical beat is selected (or purposely constructed), it is processed further in order to delineate the separate waves in the signal: waves P, QRS and T. The QRS onset and offset, as well as P onset and T offset are usually detected during the two previous steps when the beat itself was identified. The Lyon program detects in addition the T onset *[Arnaud'00]*. Here, these limits might be revised and corrected if needed and, additionally, P offset is found. Most programs apply some defined threshold on the basis of the chosen method: amplitude differences or detection functions crossings *[Stallmann'61]*, *[Willems'86]*. Others perform different kinds of template matching *[Tighiouart'03]*, *[Rubel'86]*, *[Swenne'73]*. The selected beat, also, might be slightly shifted upwards or downwards, or might have a slope according to the baseline. These alterations often have amplitudes of a few microvolts and durations of only a few milliseconds, but in risk assessment and the prediction of sudden cardiac death they play an important role. Then, their detection requires a high accuracy of the baseline estimation and thus signal correction if needed is performed.

2.5.3. ECG features extraction

Once the beginning and end of the various ECG waveforms have been identified and signal baseline is corrected, a computer program aims to extract clinically useful ECG features and parameters. Basic parameters are time intervals, wave durations and amplitudes of the various deflections. Ratios, such as Q/R and R/S, gradients, integrals and angles can be derived from these parameters. When simultaneously recorded orthogonal leads are available, spatial or planar vectors magnitudes and directions, polar vectors, instantaneous vectors and many other measurements can be obtained. Pipberger at al. *[Pipberger'61]* introduced the technique of time normalisation by dividing the QRS complex and the ST-T segment in 8 equal parts. In this way, measurements from QRS complexes with different durations can be compared. Less common parameters such as the relative area of the spatial QRS loop in each octant of the three-dimensional space or the QT dispersion are also measured in some systems, as well as many different measures of wave shape. Advanced mathematical techniques like Fourier analysis, polynomial fitting, Karhunen-Loeve expansion and time-scale or time-frequency analysis such as wavelet decomposition have been applied by some investigators *[Tighiouart'03]*, *[Laguna'96]*, *[Garcia'98]*.

Frequently, over 300 measurements are made in each ECG by various computer-processing systems. The diagnostic relevance of all these measurements however has not been demonstrated. Ultimately, only a limited subset is used in the classification programs. For pragmatic reasons, even a more restricted number is printed out in the final computer report. Nevertheless, computer ECG processing is making an important contribution to electrocardiography by enabling large numbers of different measurements to be obtained and compared in large population samples in order to identify the most efficient diagnostic parameters and criteria. In addition to the printout of the measurements, several processing systems display vector loops as well as the averaged P-QRS-T complex or a selected beat with the wave onsets and offsets derived by the program. The cardiologist or technician thereby can perform a visual quality control of the measurement results.

2.5.4. Discussion

The start of computer-based ECGs' processing manner goes back to 1957. Since then, different techniques have been developed by university research groups and by industry in Europe, America and Asia. Different computer programs were obviously based on different heuristics and different mathematical algorithms and so resulted with slightly different waves' measurements. In America, first attempts to standardise the ECG signal processing and interpretation procedure started in 1978, when the American College of Cardiology produced a collection of reports on optimal electrocardiography *[Horan'78]*. The most recent American Heart Association (AHA) recommendations for the standardisation of leads and general technical requirements of ECG instruments were published in 2007 and replaced the previous AHA recommendations published in 1975. In Europe, in 1978-1989, the European Commission (EC) funded an international project called Common Standards for Quantitative Electrocardiography (CSE), which aimed to harmonise the acquisition, encoding and storing of digital ECG data, to reduce the wide variation in waves' measurement and to improve the assessment and the classification of ECG *[Willems'90]*, *[Morlet'00]*. This project overlooked 15 different computer programs *[Willems'90]* and compared their results with the clinical truth and with the ones obtained by 7 cardiologists coming from six different EC Member States. The database for the assessment of the ECG and VCG interpretation programs was made from 1220 15-lead ECGs *"corresponding to a wide variety of morphologies" [Willems'91]*. 10 programs were assessing standard 12-lead ECG and the rest 5 the X, Y and Z leads from a VCG. The results obtained from the CSE study indicate that *"the best computer programs proved to be almost as accurate as the best of seven cardiologists in classifying seven main disease categories, i.e. normal, left, right and biventricular hypertrophy, anterior, inferior and combined myocardial infarction" [Willems'91]*. However, the results obtained so far also indicate that the classification accuracy of several programs and also of physicians can still be improved. The main limiting factor of unary interpretation of the ECG is inter-subject variability. The importance of the development of quantitative test procedures and reference libraries for assessing the precision and accuracy of ECG computer programs has been emphasized on several occasions *[Fayn'07]*. And even a number of recommendations aiming to standardise the automation of ECG signal processing appeared in the past several decades, thought the adoption of a unique standard for accurate signal analysis is still awaited. In fact, prior to the CSE study no one had proceeded with some action in this direction, and we are still lacking databases (not only the one used by CSE) for the assessment of serial ECG analysis programs. This absence might be one of the reasons why digital home-based self-care systems remain in the prototype phase for more than decade.

In the scope of this thesis we used the Lyon Program *[Arnaud'90]*, which has been used in the EPI-MEDICS project as well.

2.6. Information reconstruction

The information reconstruction is a common step in cardiology permitting to synthesise desired leads out of sets of predictor leads. The scientific interest to derive leads out of others appeared due to recording or diagnostic limitations. For example, as most cardiologists are only able to accurately interpret the standard 12-lead ECG, a synthesis of the standard 12-lead ECG was deemed necessary in case only the Frank's orthogonal three leads system (VCG) was recorded. Indeed, the possibility to use both systems – VCG and ECG – to each patient in order to exploit the diagnostic advantages of each has long been considered *[Hu'73]*. However, the routine of registering them both was not really convenient in the clinical practice. And so, an alternative of establishing a relation between ECG and VCG data sets in order to derive mathematically one out of other was taken *[Dower'84], [Rubel'91]*. The reconstruction, also, may aim at obtaining Body Surface Potential Maps (BSPM) from standard 12-lead ECG or VCG, or a standard 12-lead ECG from a reduced lead set ECG. The latter has a great importance in ensuring the practicality of a self-care approach. This sets the constraint of an accurate 12-lead ECG reconstruction out of a reduced ECG lead set in such systems *[Nelwan'04]*. It is desirable that the reconstructed 12-lead ECG contains the same diagnostic information as would contain an original 12-lead ECG. There, then, differences in the selection of the predictor leads and of the reconstruction methods seeking to get a maximal performance exists.

The derivation of one lead out of others has been considered already by Einthoven. Einthoven's law, based on Kirchhoff's law that the sum of the voltage gains and drops in a closed circuit is equal to zero, says that the potential in lead II is equal to the sum of the potentials in leads I and III at any instant in the cardiac cycle *[Macfarlane'89]*. Therefore any of them can be calculated out of the other two. As already indicated in section 2.4.1, the augmented limb leads: aVR, aVL, and aVF can be mathematically derived as well. If VR, VL and VF represent the potential differences between the Wilson central terminal (as an average potential of RA, LA and LL) and RA, LA, and LL respectively, and if, for example, aVL represents the difference of potentials between the left arm and the corresponding modified Goldberger terminal, the average potential between RA and LL, then aVL = 1.5 VL and can be calculated as aVL = (I - II)/2 *[Mirvis'07]*. The remaining six precordial leads are independent, as the exploring electrodes are not connected in a closed electrical loop, and so none of them can be calculated precisely from other information in the ECG. Therefore, the standard 12-lead ECG actually contains 8 independent tracks of information: 2 leads from the Einthoven triangle and 6 precordial leads, while the remaining 4 leads can be calculated out of 2 limb leads. Systems targeted to self-care, typically, are recording 3 to 5 leads and so shall, then, synthesise the remaining of 5 to 3 leads to obtain the whole of 8 independent leads allowing a faithful reconstruction of the 12-lead ECG. Generally, the reconstruction is performed following linear or non-linear approach.

The linear technique is "*based on the assumption that the electrical system heart-torso is linear and quasi-stationary, and on the approximation that the electrical activity of the heart can be represented as a stationary current dipole (dipole approximation).*" *[Hadzievski'04]*. That then anticipates the relationships among leads, which can be expressed mathematically in linear manner. Typically, in linear information reconstruction methods a transformation matrix is used. This matrix consists of coefficients obtained by multiple regression analysis *[Atoui'06]*, *[Scherer'89]*, *[Nelwan'04]*. The two operational ways to calculate these coefficients are based on the data source: a) out of a database of a couple of hundreds or thousands ECGs selected for that purpose or b) out of the standard 12-lead ECG of the patient himself, which was recorded earlier. The first way gives a coefficients matrix that reconstructs the best all ECGs in the database and so tends to reconstruct well any given new ECG, which for, this matrix is named a generic transformation matrix. "*...global coefficients work reasonably well for all individuals. Also the reconstruction error was not influenced by factors, such as age, sex, height, or weight.*" *[Bartosik'95]*. The second matrix is called a patient-specific transformation matrix. This matrix is applied to that specific patient newly recorded data. Several studies have demonstrated that the reconstruction accuracy using a patient-specific matrix is usually a little higher than the accuracy of 12-lead reconstruction using a generic matrix *[Atoui'04], [Atoui'10], [Nelwan'04]*. Nevertheless, a generic matrix based reconstruction remains particularly attractive in situations where no standard 12-lead ECG could be recorded prior to the recording of a reduced ECG lead set and so no patient-specific transformation could be computed. Also, reconstructions based on generic transformation matrices are less sensitive to noise artefacts or to small electrode displacements since their coefficients were computed from a very large set of ECGs of different origins.

Thereof, the uncertainty about the reality of the electrogenesis of the cardiac signal being a purely linear process *[Gulrajani'88]*, *[Modre'06]*, triggered the development of non-linear signal reconstruction methods. Best known representatives of this approach are methods based on Artificial Neural Nets (ANN). In this case, a neural net is constructed on the basis of leads' segments extracted out of the ECG and synthesises the required leads out of the input leads. Actually, the non-linear approach applies the same principal of general or patient-specific transformation as linear methods do. A generic neural net is build on the data of many purposely selected ECGs of different origin, while a patient-specific neural net is constructed out of the previously recorded personal standard 12-lead ECG. Non-linear methods are not as transparent and easy to master as linear methods and their testing is often the hardest part. Despite these difficulties, the non-linear approach is now getting more and more credits in the scientific projects for its higher performance. Typically, the performance is assessed measuring the deviation between waveforms of a derived 12-lead ECG and the standard 12-lead ECG using different criteria of goodness. Measures for goodness of fit include an RMS (Root Mean Square) error, relative error, similarity coefficients for a part or the whole ECG beat, and correlation results. Using these measurement, the EPI-MEDICS project team in the years 2004-2010 demonstrated that the technique based on an ANN approach surpasses the multiple linear regression technique in the reconstruction of the 12-lead ECG using both generic and patient specific matrixes *[Atoui'04], [Atoui'10]*. The linear methods yield high accuracy, when electrodes remain in the same position, though they falter when the reconstruction of ECGs shall be made at several days or weeks intervals. This condition of ECG recording may add additional noise and signal distortions related to small electrodes misplacements and cumber an accurate leads synthesis. Non-linear methods though face this challenge much better and are considered as the better choice for serial ECG analysis.

The essential part of ECG reconstruction is not only the waveform itself, but the sufficiency of diagnostic information seeking to diagnose a cardiovascular disease. This issue is hard to address clearly as there are no unambiguous qualitative ECG criteria that would allow the assessment. Scientific teams are usually taking three tools to measure differences between an original and a reconstructed ECG: the waveform comparison with the standard 12-lead ECG, heuristics about the model, method and the data used in the reconstruction, and statistical diagnosis verification. The weak point of the waveform comparison is that it is often made while the transformation matrix or neural net are constructed in order to judge and to demonstrate the efficiency of the method itself, but not in real conditions, during the clinical practice. This vulnerability has already been noticed by researchers *[Nelwan'08], [Atoui'10]* and the aspect of a dynamic choice or, better said, a note of personalisation based on comparison techniques performed just after the ECG is recorded is taking first steps into the domain. We stress this aspect in this thesis as it is a key to home-based self-care systems. Use of heuristics is one of the fundamental tools. They overwhelm electrode placement, input leads and leads that need to be synthesised. For example, the reconstruction of a 12-lead ECG out of 3 leads recorded with the EASI system is backed with the reconstruction of standard12-lead ECG from 3-lead VCG, which has been studied for years and has a number of scientific proofs *[Martinez'07]*. Though, following this approach, all 8 independent ECG leads need to be synthesised as none of the recorded ones are part of the standard 12-lead ECG. Meanwhile, the system proposed by the EPI-MEDICS or the WEALTHY projects records 3 (or 4 in the WEALTHY project case) standard ECG leads and so the reconstruction shall be made only for the 5 (4) other independent leads. Different studies provide different bits of facts so it’s hard to select one model or to settle down to one technique. The mathematical and fundamental backgrounds are so complex that the health-care scenario shall be first taken into account and then the final judgement of the best fit shall be made. Researches, though, often reserve this part for the clinical testing *[Martinez'07], [Schijvenaars'96]*, which is again a complex, thorough and time consuming work that everyone agrees is as a must task. *“However, the issue of how much diagnostic information is retained in derived lead sets should be addressed properly on the statistical basis.” [Horacek'08]*.

2.7. Information interpretation

Diagnostic classification is the final step of ECG analysis. This classification can be performed on: a) the current 12-lead ECG – unary analysis, b) several current 12-lead ECGs – serial analysis, and, also c) a current and one or a couple of previous 12-lead ECGs – serial analysis. Indeed, in quantitative Electrocardiology, it has been well demonstrated that performing a serial analysis by comparing a current ECG to a previous patient's reference ECG allows to significantly increase the diagnostic accuracy *[Rubel'86a]*, *[Fayn'07]*, *[Fayn'11]*. The rationale is to overcome the inter-subject variability which is quite important. Also the human interpretation of the ECG could benefit from the availability of a reference ECG, anywhere and anytime, as claimed by the ACC/AHA/ACP-ASIM Task Force that highly recommends, to electrocardiogram readers, to compare the current ECG with previous recordings to enhance the accuracy of some diagnoses, in particular of acute myocardial ischemia *[Kadish'01]*. Likewise, a more accurate interpretation of ST-T changes can be obtained by means of serial analysis, as stated by Pollenh and al.: *"Obtaining serial ECGs is perhaps the most powerful tool available to helping distinguishing from among the causes of ST segment changes. A comparison with a prior ECG tracing is, of course, invaluable." [Pollehn'02]*.

According to the CSE proposed types, diagnostical ECG classification is based on three types of statements: A, B and C. ECG analysis results with a type A statements, when referring to diseases such as myocardial infarction and hypertrophy. B type statements refer to rhythm diagnoses, and conduction defects. And, finally, C type statements indicate ST-T anomalies and axes deviations.

This diagnostic classification step is usually entrusted to cardiologists, though diagnostic algorithms may be applied upon the processed ECG signal as well. *"Diagnostic algorithms may be heuristic (experience-based rules that are deterministic) or statistical (probabilistic) in structure." [Mirvis'07]*. Heuristic diagnostic algorithms try to ground cardiologists' way of interpreting ECG by incorporating discrete measurement thresholds into a decision tree or criteria of boolean combinations *[Bonner'72]*. While statistical algorithms overwhelm techniques based on Bayesian logic *[Cornfield'73]*, discriminant analysis *[Romhilt'68]*, neural nets *[Bortolan'93]*, or Bayesian networks *[Atoui'06a]*. Statistical algorithms are likely to provide more reproducible results than heuristic methods, even though *"they still may result in discrete thresholds for diagnostic statements" [Mirvis'07]*. Statistical methods depend a lot on a database on which they were tuned and tested. The database shall contain a sufficient number of cases with varying degrees of abnormalities ranging from mild to severe cases. Results from the CSE second assessment study *[Willems'91]* on the diagnostic performance of computer programs and of cardiologists against ECG-independent validated data showed that programs which used statistical methodology for classification, had higher ($P<0.01$) total accuracy than those which used deterministic methodology when the clinical evidence was used as reference. However, when the combined cardiologist results were used as reference, the study demonstrated that some deterministic programs showed greater correspondence with the readings of the cardiologists. "*Overall, the percentage of ECGs correctly classified by the computer programs (median 91.3%)*" against the CSE diagnostic "*was lower than that for the cardiologists (median 96.0%)*" *[Mirvis'07]*, *[Willems'91]*. When comparing against the 1220 cases including cases of ventricular hypertrophy and myocardial infarction, the overall median total accuracy level was again 6.6% lower for the computer programs than for cardiologists (69.7% vs. 76.3%).

“However, the performance of the best programs nearly matched of the most accurate cardiologists.” [Willems'91].

Despite the reported lower rate in correct classification, computer assistance is able to improve the diagnostic performance of less expert readers and would be an indispensable component in home-based self-care systems. However, here, the interpretation stability becomes a key point and shall be ensured as it is the main factor for user acceptance. This requirement is not easy to fulfil, though. In fact, a given diagnostic algorithm might perform differently when applied to the same signal, which has undergone a processing pathway with different signal analysis and processing techniques. Methodological standards and well defined deviation ranges should minimise differences. One more challenge is related to interpretation stability in serial ECG interpretation, where multiple interpretations are performed on different ECGs from the same individual *[Fayn'88], [Rubel'86], [Schijvenaars'97], [Schijvenaars'01], [Atoui'10]*. There, the interpretation variability is often caused by intra-individual ECG variability. Non-pathological intra-individual variability can be related to two types of problems: technical (e.g. electrode positioning, patient posture, mains interference) or biological (e.g. respiration, non-cardiac muscle activity, circadian, day-to-day and month-to-month variations). The most prominent sources of this variability are electrode placement variations and respiration *[Schijvenaars'01]*. The respiration *“results in variations within a single ECG”*, while electrodes misplacement results *“in variations between ECGs” [Schijvenaars'01]*. Electrode placement variations appear to have a bad effect on automatic ECG classification, but classification by a human expert is less affected. That is, if a cardiologist lets himself be less influenced by the effect of electrode shifts, the computer based diagnostic classification algorithms appear to be pretty sensitive and can generate differences in interpretation, not originating from changes in heart condition *[Schijvenaars'97]*. Variations in electrode positioning, especially in the vertical direction, are revealed as non negligible in a significant number of cases. Therefore, selecting the correct electrode positions and keeping them steady during all continuous recordings remains a mandatory condition if current diagnostic programs are planned to be used.

In conclusion, the primary objective of the present review was essentially aimed to draw the scene about the complexity of the networks of algorithms and methods that an ECG signal should pass in order to result with some diagnosis. There is a huge amount of possible pathways. A critical issue is to avoid pathways that might produce false positive or false negative diagnoses.

In order to develop optimally designed, patient-acceptable, clinically-viable monitoring systems we shall also be aware of the potential clinical applications and used technologies with their technical constraints and possibilities. The overview on the existing wearable sensor-systems aimed at self-care applications is presented in the following chapter.

Chapter 3 – Review on ECG acquisition modalities in self-care

3.1. Introduction 47

3.2. Smart accessories 48

3.3. Smart garments 51

3.4. Sensing possibilities and constraints.. 53

3.5. Monitoring: *continuous* or *on demand* 56

3.6. Concluding insights 57

3.1. Introduction

Sensing and associated technologies have dramatically changed diagnostic medicine in the last couple of centuries. “*During this period the physician’s senses* [used to evaluate patient condition, for example touching the forehead with the hand to evaluate the body temperature] *were progressively replaced by sensors, first by those that measured/recorded the same parameters in a more reliable fashion (such as temperature and pressure) and later by sensors that detected phenomena indiscernible to human senses (such as the Electrocardiogram and X-rays).*” *[McAdams'11a]*. However, current developers are asked not only to bring new technologies in to the scene and propose them to end-users, but also to identify their application scenarios and if needed to create the required infrastructure so that these scenarios would be practicable, thus encouraging citizens to develop new habits.

The development of such habits is crucial in self-care for cardiac events detection. Elderly people or people at higher risk could potentially reduce the fatal rate of critical conditions and benefit from a pervasive healthcare, which enables a patient-centred everyday care. Pervasive healthcare is the “*healthcare to anyone, anytime, and anywhere by removing locational, time and other restraints while increasing both the coverage and the quality of healthcare” [Varshney'07]*. Novel systems then are called to be pervasive, natural, convenient and easy to use in order to be widely accepted and so to fulfil their primal goal of citizens’ prevention. The development of wearable sensor technologies in the around 15-20 last years seem to be taking there a pivotal role.

Wearable sensor-systems (Fig.I.3_1) are seen as a potential and very practical tool to record physiological parameters of human. They are foreseen to be used at home, during emergencies, in military or aeronautics as well as for leisure or sport applications *[Cabrol'05]*, *[Jafari'05]*. Wearable technologies tend to propose different sites on the body surface for data recording, to prolong the recording time or to increase the recording frequency so providing overall more data, which then combined with ubiquitous computing for data processing aggregate the knowledge upon the case and allow timely decision-making. The enhanced situation-awareness that these systems can provide while the user is performing his daily activities, without interfering significantly with the comfort or user lifestyle, makes them one of the most attractive choices for self-care.

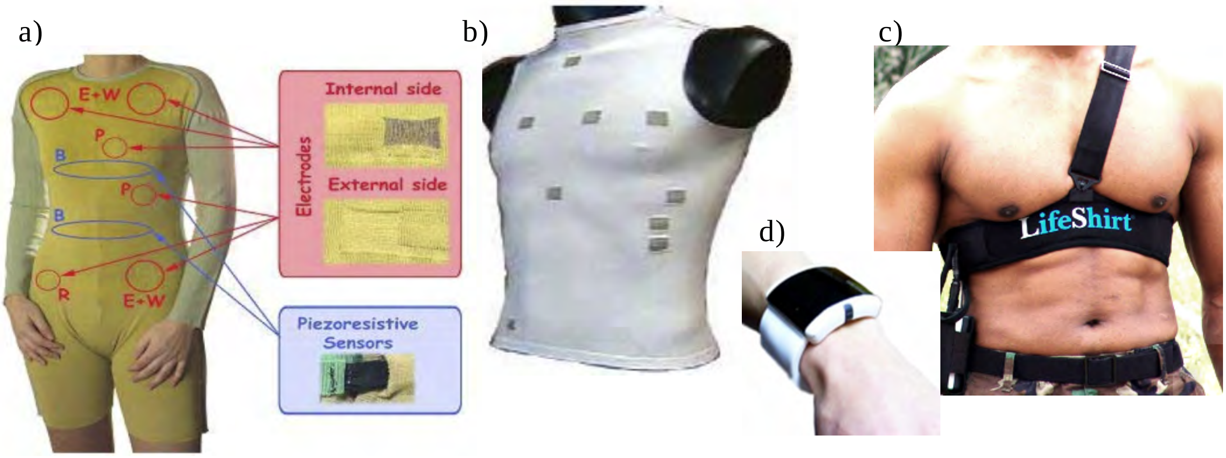


Fig.I.3_1 Examples of wearable sensor-systems: a) Wealthy suit, b) Philips HeartCycle t-shirt, c) VivoMetrics Vivo Responder belt, d) Exmovere BT2 watch (Images a) and b) are extracted from *[Goodson'06]*, *[Luprano'06]* and c) and d) reprinted from *[McAdams'11]* with permission of the authors)

The technology itself is logically split into two units: the *sensor-system* and the *monitor* for data processing. Those two units are interconnected with the third one – *data transferring unit*, which is responsible for signal transfer from sensors to the monitor.

The sensor-system aims at *"removing the task of placing the sensors by a professional as well as offering a natural interface with the body with accurate, reproductive positioning of the sensors" [Lymberis'08]*. Typically, the backing used to integrate sensors defines two large groups of sensor-systems: smart accessories (watches, belts, gloves, patches) and smart garments (clothes: t-shirts, vests, suits). However, the *"adherence, characteristics of the fabric, cut, positioning of sensors are factors to be considered in order to guarantee a good signal quality, freedom in the movement and comfort to active people monitored during daily life, sport and work activities" [Rienzo'07]*.

The monitor: It contains intelligent applications for user guidance, data processing, data analysis and a real-time feedback to the user. *"Medical monitoring applications differ from other wearable applications in their I/O requirements, sensors, reliability, privacy issues, and user interface." [Martin'00]*. The monitor can be attached to the sensor-system - a pocket type device (PDA), or can be installed into the mobile phone, personal computer and accessed remotely.

The transfer: It ensures the recorded data transmission to the monitor. Low power components permitting cabled or wireless data transmission via Bluetooth or Zigbee protocols are used *[Zaunseder'07], [Kurlkarni'07]*.

To compose these three elements as a whole, which provides an increased personalised functionality, facilitates communication and embeds the decision support in order to make users autonomy available, requires a multidisciplinary research approach. *"Techniques for on-body sensing, context awareness, user friendliness, power autonomy, intelligent data processing and interaction with professional medical services are among the challenges concerned." [Lymberis'06]*.

Starting from 1996 when the Georgia Institute of Technology in the United Stated of America (USA) got a funding from the US Department of Navy in order to develop a wearable "intelligent fabric" that could be used to monitor vital signs during combat conditions *[GTWM'96], [Moulton'98]*, the era of wearable sensor-systems has begun. *"The Wearable Motherboard project and related work at Georgia Tech has led to the creation of a system for monitoring a user's health, including heartbeat and respiration as well as the location of a bullet wound." [Martin'04]*. By now we count around 15 to 20 different wearable sensor-system projects that resulted with wearable sensor-system prototypes or products capable to record the ECG. We will review them in the two next sections: smart accessories and smart garments, seeking to highlight for each of them the target audience and the underlying technologies and to present their functionality. Then, we aim at discussing those products/prototypes appropriateness with respect to the self-care scenarios.

3.2. Smart accessories

Watches

AMON (2001.01-2002.12), Zurich, Switzerland is an Advanced care and alert portable telemedical MONitor developed by *Art of technology AG* under the EU FP5 IST project. *"The system comprises two separate parts: a wrist-worn unit (WWU) and a stationary unit* *at telemedicine center (TMC)." [Kurlkarni'07]*. It is aimed at high-risk cardiac/respiratory patients. It measures blood pressure, SpO2, one lead ECG and is capable of recognising the user activity. Its main functionality is to continuously collect and evaluate multiple vital signs, to detect emergencies and to connect to a telemedicine centre in order to send the data for evaluation. The watch has three integrated electrodes devoted for ECG recording (RA, LA, and RL). In order to obtain a 1-lead ECG the RL electrode needs to be in contact with the user skin (on the back of the watch), while the RA electrode is situated on the top of the watch and needs to be touched with the index of the user's right hand. At first, the developers used typical Ag/AgCl electrodes without any past or gel. However, after three months' testing, due to oxidation no of ECG signal could be measured, and in following prototypes gold based electrodes started to be used.

Telepath (2008-present), New York, USA is a psychological state monitor developed by *Exmovere Holdings, Inc.* (former *Exmocare*). The watch can measure the heart rate, heart rate variability, skin conductance, skin temperature and a relative movement, process these data via the emotional interpretation engine and can send a report upon the wearer's emotional and physiological state. It has a functionality of various alerts based on data interpretation, may connect over the GSM network for remote monitoring around the world, has point-to-point wireless ZigBee communication and also can remotely track the location of a wearer. A rich onboard OLED (organic light-emitting diode) based user interface allows to use the device without a personal computer. The management of a fire-fighters crew is a typical application for this device. Each fire-fighter is supposed to wear the watch so that the data about their emotional or physical condition as well as their location into the burning building/place can be transferred to a central monitor. The crew manager can then observe real-time data and take timely decisions. Besides the emergency/rescue crews, any citizen, clinical researcher or academics and smaller home health care providers are the target auditorium of this product.

Belts

Vivo Responder (2000-present), California, USA is a lightweight chest strap designed by *VivoMetrics, Inc.*, to be worn by first responders or hazmat workers (during fire-fighting and hazmat training, fire and hazmat emergencies, industrial clean-ups with heavy, hot, protective gear, biohazard occupational work). It is intended to be set into under layers of protective clothing that are used by these users and is capable of recording a 1-lead ECG.

Ergometry belt (2005-2008), Petach Tikva, Israel is an easy to fit electrode system designed by *Tapuz Medical Technologies* for recording diagnostic 12-lead ECGs. It is a stretchable belt that fits to anatomically accurate electrode-body locations (for electrodes V1-V6 positioning) for men and women. It is announced to be ideal for hospitals, clinics and homecare facilities in order to avoid positioning errors and to save time on placing the electrodes. The belt can be connected to almost all ECG devices via banana plugs or studs. The manufacturers indicate that there is no need to shave patient and for women to remove the bra during examination. It is claimed that the innovation enables the correct electrodes (Ag/AgCl coating) connection, irrespective of the patient's chest dimension, hence, the level of staff expertise, for ECG examinations, is reduced. The belt is reused as it can be disinfected with any standard disinfectant used for electrodes or washed with antiseptic soap and water.

BioHarness (2007-present), Auckland, New Zealand is a fabric-based, dry contact strap designed by *Zephyr* for sports, health and wellness, academic research, or military applications. The strap contains 2 textile electrodes that allow the monitoring of the heart rate (R-R) based on a 1-lead ECG. It also measures the breathing rate and depth, skin temperature and activity/posture via a 3D accelerometer. The strap is unobtrusive due to fabric-based electrodes, is comfortable over long periods, and can be washed. The monitor that goes along with the strap is performing a comparative analysis and can work at real-time or off-line.

Chest Strap (2006-present), Philadelphia, USA is a fabric based chest strap designed by *Numetrex* (created by *Textronics*) for sportsmen. It is used to measure the precise heart rate (R-R) on the basis of a 1-lead ECG. The manufacturer indicates that it is washable and that up to 80-100 washes, no changes of the electrode characteristic can be detected, afterwards, small accuracy changes could appear. The monitor, usually integrated into the watch that goes with this strap, proposes a 5 levels training program based on the heart rate and advices which level needs to be worked out to keep fit.

Patches

LifeGuard (2004), California, USA is a sensor pad created by NASA (Ames Research Center Astrobionics), US Military (Matick Soldier Center) and Stanford University (Medical Center, National Biocomputation Center). It is designed for soldiers that are undergoing physiological stress tests or in field, climbers, astronauts during re-entry, and human research on centrifuges *[Mundt'05]*. The pad contains several electrodes, which are stuck to the plastic pad in order to keep their correct position. The user shall follow a rigorous procedure in order to obtain the best results: shave the chest in the area of the sensors, use an enclosed alcohol rub to prepare the skin surface, align the sensor pad with the midline of the body. Then, he shall top the plastic tab into the neck, peel off the sticky white backing, fix the electrodes to the body, and, finally, peel off the clear plastic alignment template. The pad is intended to be used for 1-2 days continuous monitoring. Four out of all the electrodes are used for ECG recording and allow getting a 2-lead ECG (leads II and V2). Besides the ECG, the device is capable to measure a respiratory rate, has a pulse oximeter, a blood pressure monitor, and a temperature probe. The personal monitor is wired to the electrodes and sends the recorded data to the base station for data analysis.

Wireless Vital Signs Monitor (2006-present), Belfast, Northern Ireland, UK is a patch and a small attachable monitor that have been designed by *Intelesens, Ltd* (former *ST&D*) for comfortable patch wearing for long lengths of time (it can be even kept while the person is taking a shower). The patch contains 3 closely set electrodes that are glued in the middle of the chest. They are capable of providing a 1-lead ECG. The monitor is set to detect various arrhythmias and to send alerts to the medical center.

Gloves

Glove-Sleeve (1999-2007), Raanana, Israel is a patented design (by *David Daniel*, *Levy Irving*, *Liber Serge*) of a glove and sleeve that contains all 10 electrodes for 12-lead ECG with Mason-Likar positioning recording. The electrodes are localised throughout the length of the inner side of the glove-sleeve (hand plus forearm) and are positioned on the torso when a user places his left arm with this glove-sleeve garment on the chest. The invention aims at sensor applications with no prior knowledge of proper sensor placement. The glove-sleeve garment shall assure the proper placement of all sensors. The targeted applications compound human subjects that are to be monitored in hospitals, clinics, or doctor's offices, for remote locations (home environment, work place, recreational activity) or unnatural environment (underwater, outer spaces). Designers expect the glove-sleeve to record not only the ECG, but also blood pressure, blood oxygen saturation, skin resistance, and motion.

PhysioGlove (2009-present), Evanston, USA is a glove-sleeve manufactured by *Commwell, Inc.* All technical aspects resemble the Glove-Sleeve design we have described above.

Handheld devices

Alive Heart Monitor (2007-present), Queensland, Australia is a wireless health monitoring system for screening, diagnosis and management of chronic diseases, and for consumer health and fitness designed by *Alive Technologies*. The system consists in a small device with 2 wired electrodes that are placed on the torso using standard Ag/AgCl electrodes in order to acquire a 1-lead ECG. This device is considered as a flexible alternative in which patients with atrial fibrillation, heart failure, during cardiac rehabilitation and fitness monitoring undertake supervised walking while their speed, location, ECG-based heart rate and ECG are monitored over the Web by an exercise physiologist.

Handheld repositionable ECG detector (2007-present), The Netherlands, USA is a patented design (by *Koninkl Philips electronics NV, Cross Brett, Lyster Thomas, Fong Shannon, Solosko Thomas, Gehman Stacy, Hansen Kim J, Herleikson Earl, and Hugh Steven*) of an ECG monitoring device that allows clinicians to determine the optimal patch and monitor locations. It is a small, handheld, repositionable device which allows the physicians to query several locations and decide which ones to monitor prior to adhesively attaching the patch which indicates exact device location on a particular body. It is targeted to allow the patient to independently take an ECG for 30 seconds or a minute at a time, and then put the device away until the next measurement time. The primal design contains 4 electrodes so it should be able to record a 3-lead ECG. The monitoring system is not described.

3.3. Smart garments

T-shirts

SmartShirt (1996-2001), Georgia, Maryland, USA is a pioneer of the smart shirts. It was developed by the *Georgia Institute of Technology* and *Sensatex, Inc.* under funding of the US Navy. The project developed a prototype of a washable, light, breathable, cotton fabric, which incorporats a patented conductive fiber sensor-system designed specifically to capture the required biometric information. The shirt is capable of measuring heart and respiration rates, body temperature, and to record a 3-lead ECG. The acquired data are sent to a PDA type device or to a base station using Bluetooth or Zigbee protocols.

LifeShirt® (2000-present), California, USA is a lightweight, machine washable, comfortable, easy-to-use shirt with embedded sensors designed by *VivoMetrics, Inc.*. The shirt is made of a hand-washable, reusable stretch-material into which are sewn an array of physiologic sensors to monitor 30+ vital signs. LifeShirt® can be used during normal everyday activities at work, home, play and during sleep – but not while swimming or bathing. The project started hoping to diagnose sleep disorders. However, later on it developed it to meet other clinical, academic, corporate and preclinical research targets and to be used within sport and fitness activities. It records a 3-lead ECG at Mason-Likar positioning. It is produced in 15 different sizes.

MyHeart (Heart Cycle) t-shirt (2004-2013), Aachen, Germany (plus 34 partners all over the Europe) is a textile sensor system integrated into the t-shirt that was designed by *Philips* and *EU IST FP6 MyHeart* and *EU IST FP7 Heart Cycle* projects. It is aimed at citizens to fight cardio-vascular diseases by preventive lifestyle and early diagnosis. *"Four electrodes to record a three-channel ECG and all electronics arranged on a partially flexible printed circuit board are integrated." [Zaunseder'07]*. The electrodes are places following the EASI positioning and allow the recording of X, Y, Z leads. The recorded data are sent to the doctor for evaluation.

Quasar ECG shirt (2004-2007), San Diego, USA is a design of a smart t-shirt made by *Quasar* using capacitive bio-electrodes capable of through-clothing measurement of a 1-lead ECG. The project targets at fire-fighters and "tender" patient populations (elderly, allergic, paediatrics and neonates) monitoring. The t-shirt is, also, suitable for long term studies of rare events detection. The integrated sensors operate without direct electrical connection to the subject. Therefore, ECG measurements can be made very fast as no skin preparation or a clothes removing is required. This project currently does not discuss data analysis issue.

CardioShirt (2006-present), Philadelphia, USA is a shirt with two woven/knitted textile electrodes designed by *Numetrex, Inc.*. The garment is targeted at sportsmen as 2 electrodes provide a 1-lead ECG and so an accurate R-R interval, which serves to set training program. CardioShirt is sold in several different sizes, though specific electrodes' placement via augmenting size of clothes is not indicated.

Suit

WEALTHY suit (2002-2005), Italy (France, Germany and Swizerland) is a suit with integrated textile electrodes (conductive fibres are woven with stretchable yarns) that has been designed within the *EC IST FP6 project WEALTHY*. The garment contains 6 ECG electrodes, 4 impedance electrodes, 9 insulated connections and 2 embedded temperature sensors in order to provide 3-lead ECGs (II, V2 and V5), respiration measurements, core and skin temperature, and body position (accelerometer). The garment is aimed at monitoring patients during rehabilitation, professional workers during risk activities, users in everyday tasks or during physical exercises. *"Due to the good quality of recorded signals, the ECG can be adequately employed to study non invasively and in behaving conditions more complex functional indexes related to the sympatho-vagal balance, such as low frequency and high frequency components derived by spectral analysis of RR interval variability, respiratory sinus arrhythmia and area under T wave of the ECG." [Paradiso'05]*.

Vest

SensVest (2001-2004), Greece, Germany, Italy, Austria and England, has been designed by the *EU IST FP5 project Lab of tomorrow* seeking to enhance the learning experience. The vest was intended to be used by schoolchildren all day long to capture their activity level, heart rate and temperature changes. The SensVest project passed several phases to measure the heart rate. In the phase where the ECG was recorded, two electrodes were attached to the chest and one electrode to the hip. All electrodes were held with adhesive pads. The project claimed that this way caused difficulties in practice, because of the need to position electrodes correctly and to prepare the skin. Also, electrodes became loose when the body got sweaty or affected by arm movements, and children complained of a discomfort when electrodes were removed. The SensVest was produced in two sizes: one for women and one for men *[Knight'05]*.

3.4. Sensing possibilities and constraints

Every project on wearable sensor-systems claims of a positive user feedback and a set of possible applications. However, to get a profound understanding, we should picture the possible monitoring scenarios and be aware of sensing possibilities and constraints.

Home-based personal medical care started with portable monitors and a set of wired sensors that were mounted on the patient body by professionals and let for some prolonged time (over the night for example) in order to get more data and so to get a more precise diagnosis. *"Traditionally, personal medical monitoring systems such as Holter monitors (a portable/wearable ECG device) can collect data for up to 24 hours. The recorded data is subsequently retrieved and analyzed by a clinician." [Kurlkarni'07]*. However, this type of monitoring had several limitations: (a) data could not be analysed on real time: first data collection, then off-line processing, (b) a rigid device: many wires and adhesive electrodes, therefore used most often while sleeping to avoid motion artefacts (c) nonexistent support for massive data collection and knowledge discovery (e) unknown contextual data to get reliable interpretation *[Kurlkarni'07]*. These limitations one by one were eliminated or lightened in terms of *"miniaturisation, seamless integration, functionality, comfort, data processing and communication"* thanks to the significant technological advancements that happened in the last 10-15 years *[Lymberis'08]*.

Textile-based unobtrusive sensors: *"In recent years, existing technological advances have been made in development of flexible electronics. These technologies offer the opportunity to weave computation, communication and storage into the fabrics of the every clothing that we wear, therefore creating intelligent fabric." [Jafari'05]*.

Commercial stainless steel threads were twisted around a standard continuous viscose textile yarn in order to obtain a conductive fabric electrode *[Scilingo'05]*, *[Wijesiriwardana'04]*. A comparison of these new textile-based, flexible electrodes with standard ECG clinical electrodes was continuously performed by various research studies *[Scilingo'05]*, *[Pola'07]*, *[Mestrovic'07]*. Signals coming from both types of electrodes were evaluated upon most significant parameters extracted from the ECG waveform morphology (P-waves, QRS complexes, ST segments, and T waves), performing a spectral and other analysis. Different knitting techniques: a) fabric electrodes realised by flat-knitting technology (Wealthy project) + hydrogel membranes (Intelesense, Inc.) and b) fabric electrodes realised by seamless knitting technology (My Heart project) without hydrogel membranes, were evaluated *[Pacelli'06]*. The evaluation results overall proved that both signals are of similar quality. *"The impedance value recorded from fabric electrodes is similar to that obtained from RedDot* [by 3M™] *electrodes in the same frequency range." [Scilingo'05]*. Fabric electrode response did not show any significant polarisation during the time and, moreover, the magnitude squared coherence function calculated on ECG signals with fabric and classical electrodes showed *"a cross-spectrum factor greater than 0.095" [Scilingo'05]*. These recent achievements greatly encourage the use of knitting technologies in order to provide *"elastic, adherent, comfortable garments with these* [close to skin] *inherent properties" [Pacelli'06]* that are needed to detect vital signs and user movements. Fibres are giving a comfortable, easily wearable garment and optimise *"the positioning of the sensitive regions, thus making the monitoring more accurate and reproducible" [Scilingo'05]*.

Textile-based electrodes are unobtrusive as they are not adhesive or sucked to the skin. However, the micro-movements of these electrodes with respect to the skin results in the deterioration of the signal-to-noise ratio. This might explain why at first it was popular to simply insert standard adhesive clinical electrodes into seams and pockets of the garment, thus forming a hybrid tool. So far, this issue for textile-based electrodes *"has been neutralized by the interposition of hydrogel membranes, but further efforts in modifying the structure of the yarn in the sensitive spot are necessary to stabilize the skin-electrode contact" [Scilingo'05]*. This solution, though, is problematic in long-term monitoring as with time the gel will dry out. This is probably the main reason why all current smart garments or accessories that use textile-electrodes are intended for applications, where sweating due to stress (fire-figters) or a physical active (sportsmen) is one of the major environmental components. As the sweat is a natural gel, it ensures lower impedance and helps electrodes to stick better to the skin so reducing micro-shifts of the textile electrodes and therefore improving the quality of the signal.

Non-contact unobtrusive sensors: Typical ECG sensors as well as textile-based sensors require conductive gel to ensure low-impedance electrical contact between the sensor and skin. The skin irritation caused by standard clinical sensors due to adhesive electrode attachment and salty gel as well as the micro-movement artefacts coming from electrodes based on conductive fibres encouraged research on non-contact sensors. *"To address these problems, a new class of miniature, ultra low noise, capacitive sensor that does not require direct contact to the skin, and has comparable performance to gold standard ECG electrodes, has been developed." [Park'06]*. Non-contact electric field sensors are being developed and tested by several research groups and have the potential to revolutionise biopotential recording systems *[Park'06], [Sullivan'07], [Matthews'07]*. These sensors operate in 3mm distance from the skin. *"Experiments coupling the sensor to human scalp through hair and to chest through clothing produce clear EEG and ECG recorded signals." [Sullivan'07]*. *"For each physiological variable, a comparison with conventional wet electrode technology has demonstrated that QUASAR's biosensors provide data of similar quality*." [*Matthews'07]*. These sensors have a clear advantage of being truly unobtrusive and of a fast use: no clothes removal or skin preparation. However, the complete elimination or management of motion artefacts issue remains unsolved.

Motion tolerance: *"Motion artefact is a major limitation in most practical implementations of wearable health monitoring devices." [Such'07]*. Quantification of the motion artefacts in order to design properly the ECG electrode is one of the recent research areas. Various clinical exercises are performed and evaluated. A motion artefact *"can increase and decrease depending on how hard the underlying muscle works, the amplitude and speed of these movements (and hence the amount of deformation of the skin) must be replicated as accurately as possible" [Kearney'07]*. Posture assessment and classification as well as thermal stress effect are studied *[Pawar'07], [Rienzo'07]*. However, there currently does not exist a structured and consistent approach to address this problem, *"nor exists a reference database or accepted model design to this field of research, making it almost impossible to objectively assess the performance of motion artefact reducing algorithms" [Such'07]*. Contact potential related problems due to material or movement, contact impedance related problems due to tissue impedance, signal attenuation, signal distortion, patient insulated from ground, patient grounded or other issues happening due to electrostatic interference, defibrillation, cable artefacts, or a magnetic field make the solution of those problems complex and hardly achievable.

Power consumption: The duration of monitoring, wearable sensor-system autonomy allowing users at being more independent, and even at the device price are directly related to power consumption issues. There are two conceptual nodes that require power: the recording itself and, in particular, data transfer. Data recording with body sensor networks, differently than wearable sensor-systems, requires only a single centralised power source, while wearable sensor-systems are usually having several distributed batteries. *"The sensors could either be directly connected to wearable electronics using dedicated tiny conductive wires knitted like normal textile yarns, or local processing nodes could be designed to acquire the signals from different sensors." [Luprano'06a]*. Meantime data transferring is a key technological constraint for making wearable sensor-systems mobile systems *[Jafari'05]*. Optimisations related to sensor-system connection to the monitor and data transfer protocols are being made: *"We optimized the Bluetooth parameters and operation modes to reduce the power consumption as much as possible." [Borromeo'07]*.

Recording sites: One of possible motion artefacts management techniques is a well chosen electrodes placement. The areas with greater signal amplitude, lower electrical noise are prioritised there. It is also desirable to avoid sensitive places: body hair and sensitive skin. *"Sometimes patients have wounds or bandages that preclude access to the proper body locations." [Brodnick'00]*. One of the first studies on body locations *"for comfortable and unobtrusive wearability"* was performed by Francis Gemperle in 1998 *[Gemperle'98]*. Only specific sites on the body were recommended for wearable sensor-systems by this study, as they not interfere with rigorous requirements of comfortable attachment, device accessibility, weight, thermal aspects, long time use and even aesthetics. Klug and Muhlhauser were checking another aspect of wearable devices - *"how well different interaction devices/techniques integrate with specific realworld scenarios"* – by taking in account that workers (i.e. fire-fighters) shall first of all deal with their primal task and not with additional device control *[Klug'07]*. They found out that indeed conflicts *"between task performance and device requirements"* are possible and need to be well thought of in the device design phase. In cardiology, Finlay et al. performed a study that *"aims to assess the effect of various practical constraints that may be encountered when choosing electrocardiographic recording sites for wearable health systems falling within the category of smart shirts for cardiac monitoring and analysis" [Finlay'08]*. Observations upon 10 selected recording sites concluded that *"where possible, as much anterior and precordial territory should be explored and systems limiting recording zones to the posterior and lateral surfaces should only be utilized where absolutely necessary" [Finlay'08]*. However, authors confessed that they were using dry electrodes and so they could cause micro-moves and insert additional noise. A study including more precise measurements is suggested to be done.

Electrodes' placement: The recording site selection has one more critical issue – a correct and repeatable electrodes' placement on different morphology subjects. *"Accurate placement of the electrodes onto the patient's body surface is required to record a useful electrocardiogram (ECG). The attachment points are defined and accepted with the medical community but routine correct placement is difficult to achieve." [Brodnick'00]*. In 1991, Hermann et al. *[Herman'91]* invented a sliding ruler that is placed on the patient's chest and indicates where electrodes should be placed. However, the test results revealed that displacement of chest (precordial) electrodes is still present, while only 2 cm displacement already resulted in significant electrocardiographic interpretation changes. Brodnick et al. proposed one more method to find out a correct and repeatable placement. His team suggested to calculate the angles between the main leads of a reference 12-lead ECG and to compare them with the angles of the newly obtained ECGs. However, *"knowing the lead II vector does not by itself give enough information to identify where the RA electrode and the LL electrode are" [Brodnick'00]*. Thus, a repeatable position of well places electrode remains a challenge.

3.5. Monitoring: *continuous* or *on demand*

Technical sensor-system requirements and the performance-related, reliability and availability issues play a great role in the possible health-care scenarios. Almost all our discussed wearable sensor-systems underline in unison the vision of a continuous monitoring. This vision is based on the fact that the more data are collected about the case, the more precise diagnosis should be made. However, the type of required and situation-aware monitoring and the type of best suited wearable sensor-system for the selected type of monitoring shall be discussed prudently. We could distinguish two main monitoring approaches intended for self-care applications: *continuous* monitoring and *on demand* monitoring.

Continuous monitoring: It is a type of constant monitoring, where the signal(s) is recorded non-stop or in automatically set intervals.

On demand monitoring: It is an occasional monitoring, where the signal is recorded randomly and the quantity of signals may vary.

In order to identify underlying differences, in table I.3.1 we in brief overview several monitoring steps that are common to both approaches.

Step	Continuous monitoring	On demand monitoring
Correct electrodes placement	*Remains an issue*	*Remains an issue*
Repeatable electrodes placement	*During the period of time when the sensor-system is placed on the body and not moved, a signal recorded continuously or signals recorded at some in advance set intervals are of the same origin as the same electrodes positioning was ensured. The repeated sensor-system placement for a new period remains an issue*	*Remains an issue*
Data storage	*Typically, data are continuously or at some set intervals transferred to the central monitor station*	*The device itself may be used to store data*
Data analysis / Alerts	*Comparison techniques are mostly applied. Any significant change in comparison with previous signals or previous intervals of the signal can*	*Sophisticated techniques need to be applied in order to analyse the recorded signal. Comparison techniques might*

	cause an alert [Callens'08]	*be used as a supplementary tool if there are previously recorded (and comparable) signals stored in the device*
Recommendations	*Usually are given by a professional, who analysed the data via a central monitoring station*	*Automated, based on accumulated knowledge and sophisticated techniques*

Table I.3.1 A brief overview on *continuous* and on *on demand* monitoring types via 5 common monitoring steps

3.6. Concluding insights

Having a broad picture on existing wearable sensor-system prototypes / products, used techniques and intended scenarios, we hereafter discuss one after other each type of smart sensor-system with respect to ECG recording in self-care for Myocardial Infarction prevention and we summarise the appropriateness of smart accessories and garments for cardiac self-care. However, at first it is needed to be said that all prototypes/devices capable of recording only one lead ECG are not suitable to detect Myocardial Infarction related risk as a diagnosis of coronary artery diseases requires spatio-temporal information about the heart. We would also state that *on demand* monitoring suits better our prevention objectives, while *continuous* monitoring is rather useful in rare, though repetitive events detection, or specialised cases. We shall keep in mind that *"Design considerations for e-textile applications encompass a range of areas, including the physical environment, sensor behaviour, human body size and motion, the shape of the cloth, garment and fabrics manufacturability, and standard computer engineering areas such as networking, power consumption, and software execution." [Martin'04]*.

Smart accessories

Watch: Does not suit as a sensor-system – records only 1-lead ECG, but could be used as a monitor if combined with a sensor-system capable to record 3-lead ECGs.

Belt: Does not suit as a sensor-system – records only 1-lead ECG and cannot be used as a monitor.

Adhesive patch: Its usability depends on the model. Patches that allow recording only one lead ECG are not suitable. Patches that are capable to record 3-lead ECGs can be considered for *continuous* monitoring with the condition that the electrodes are placed in the suitable for wearability (especially prolonged) recording sites. *On demand* monitoring would be problematic as each placement and removal of this sensor-system would cause skin irritation.

Glove: The design of the glove that is capable to record 3-lead ECG is an attractive option for *on demand* applications. However, the issue of placement repeatability needs to be considered and some additional solutions need to be provided. *Continuous* monitoring with this sensor-system design shall not be recommended.

Handheld device: The prototype and the functionality of the *Handheld repositionable ECG detector* can be partly considered as an option corresponding to the cardiac self-care which this thesis is targeting at. The strong point of this approach is the phase where the doctor evaluates the signal's quality while repositioning the device and chooses a good, patient-specific recording site, which is then somehow marked. Marking is the step that needs to be developed further. A piece of patch or a marker can be used only for a short period of time. The disadvantage of this approach is that the device is intended to be placed in the middle of the chest (otherwise it would not be a hand-held device), while this recording site is not recommended for wearable sensor-systems if we follow Gemperle et al. recommendations. This scenario also would require a visit to a professional (in order to get a correct electrodes' positioning), which can be a limit as many potential users would find it too complicated.

Smart garments

Suit/ T-shirt / Vest: Textile-based electrodes integrated into the garment are a good option and are suitable with this thesis objective. Differently than it is now widely accepted, the currently designed/developed smart garments would better suit for the *on demand* monitoring than for *continuous* monitoring. *Continuous* monitoring using this type of garment is meaningful only for users that maintain some activity, which allows perspiration in order to ensure a low level of noise (reduced micro-moving) and lower impedance. Though, this would not be necessary for short and calm signal recording. The issue that this approach should deal with is the repeatability of electrodes placement. Garments integrating non-contact electrodes would probably suit less as they would be less fixed on the body and so less stable. Besides, they will have the same problem of electrodes' places reproducibility. However, the non-contact type of electrodes might be a good option for continuous monitoring, if motion tolerance is tested.

In conclusion, wearable sensor-systems in the last decade gained a lot of interest as systems capable to increase the knowledge about a specific user in a specific situation. However, their usability scenarios are still limited, require a rigorous employment procedure and are only targeted at a specific audience in order to avoid some technological issues. This shall be changed if we aim at enabling a true pervasive health-care accessible to a profane user. We believe that our review on already finished or currently undergoing projects provides this connecting element, which helps to see the shortcomings of current prototypes/products in a light of user requirements, needed applications and technological possibilities. We shall state that currently there is no a sensor-system and, in general, a monitoring device that would enable a non competent user to record on demand by his own his personal diagnostic quality ECG in order to make a timely decision upon his health.

Chapter 4 – Personal sensor-system modalities design and evaluation for simplified electrocardiogram recording in self-care

4.1. Introduction 61

4.2. Rationale for personal sensor-system modalities design 61

4.3. Sensor-system modalities design 63

4.4. Study population 65

4.5. Data acquisition 67

4.6. Data analysis method 67

4.7. Results obtained with comparison studies: C1, C2 and C3 73

4.8. Paired t-test results obtained comparing the C1, C2 and C3 studies 82

4.9. Repeatability study results 83

4.10. Discussion 85

4.11. Conclusion 86

4.1. Introduction

In this chapter our goal is to propose a design of a sensor-system that would meet the following requirements that are aimed at supporting the scenario of self-care in cardiology:

a) Keep electrodes in their standard locations, so that we can analyse the obtained signals using the current knowledge background and already applicable algorithms in order to process the signal and to provide a digital interpretation.
b) Record a 3-lead (I, II and V2) ECG in order to detect cardiovascular diseases
c) Provide a sensor system that can be put on and off very easily, so that the user can discretely check his health status anywhere and at anytime.
d) Shall be compatible to the morphology of each sensor-system user in order to obtain a diagnostic quality ECG signal.

Hereinafter, we first describe the rationale for designing a sensor-system with different modalities. We then propose a sensor-system that is appropriate for self-care, we evaluate our proposal by performing a series of experiments on eight healthy volunteers and we present the results of 280 ECG comparisons in terms of correlation, RMS values, as well as changes in diagnosis probability and of selected ECG measurements.

4.2. Rationale for personal sensor-system modalities design

Previously we reviewed a number of scientific studies that in different stages of ECG signal processing and analysis thoughtfully state how the accuracy in positioning sensors in their correct anatomical positions is crucial for the correct computer based diagnosis and thus the success of home-based self-care systems. However, none of our overviewed potential sensor-systems takes in account morphological differences of each individual in order to ensure the required accuracy. Probably due to the previewed sensor-systems integration into garments an idea that comes by default is that it's enough to provide different sizes of a smart garment to solve this problem. Hereof, we first decided to make a mini-simulation on real data taken from 2 subjects to check this very natural hypothesis.

We thus have tested the two following statements: (1) one electrode displacement can significantly distort ECG signal, but (2) simultaneous displacement of a set of electrodes does not introduce any bias when serial analysis is performed. This second statement also includes the assumption that in case the signal is altered it could be possible to calculate a coefficient to correct disturbances.

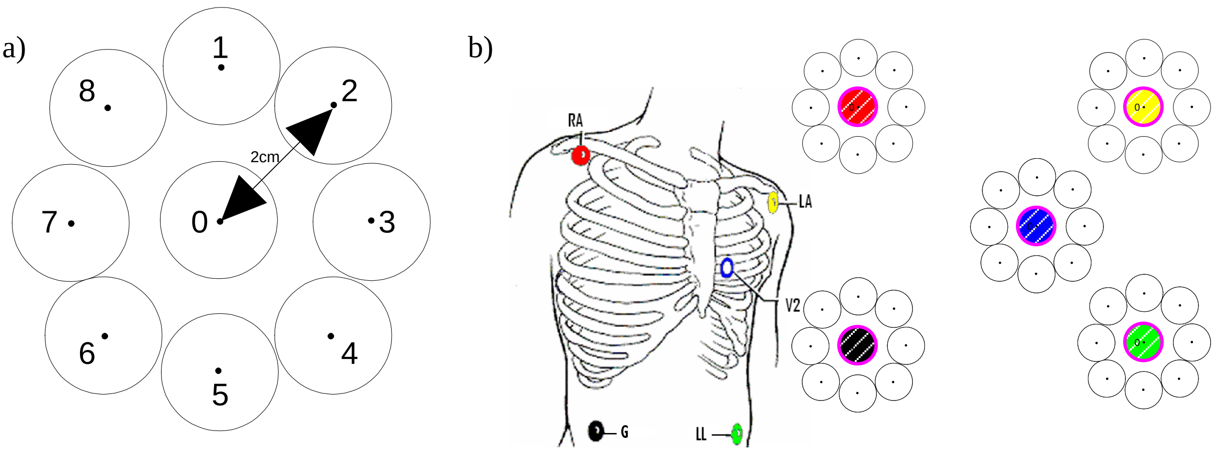


Fig.I.4_1 a) a circle type element b) electrode displacement map

In order to check these statements we first defined the electrode displacement step and direction. For this purpose we have proposed a not absolute coordinates system (the “0” position depends on the morphology of each individual) describing the place of an electrode on the body. The main element of this coordinates system is a circle of 9 possible electrode places in 2cm distance from the centre (Fig.I.4_1a). The position named “0” (zero) is called a reference position and it is placed in the anatomically correct electrode location. There can be eight possible displacements for each displacement step of 2 cm to directions: up, down, left, right, 45°up-left, 45°up-right, 45°down-left, 45°down-right from the reference position. One, two or even the whole ensemble of electrodes can be moved according to this coordinates system. All recorded ECGs were expressively labelled using this coordinates system.

To simulate self-care in real situations all recorded ECGs used the Mason-Likar positioning, but only the electrodes defined by the EPI-MEDICS project as most convenient to use in home-based applications – RA, LA, LL, V2 and G – were moved (Fig.I.4_1b). Two principal schemes related to our statements were assessed in this simulation: (1) moving 1 electrode at a time or (2) moving a set of 4 electrodes at a time. We used a standard electrocardiograph simultaneously recording all leads with standard (Ag/AgCl) electrodes. Each time when the electrode needed to be moved / replaced, a new electrode was taken. Each recorded ECG was labelled with a five numbers’ code corresponding to a number of an electrode in the chain: red, yellow, blue, green, and black (RA, LA, V2, LL, G).

Twenty two 12-lead ECGs were recorded for each of 2 voluntary subjects following the hereafter listed protocol:

[1]: One standard 12-lead ECG.

[2]: One 12-lead ECG with Mason-Likar positioning: 4 limbs electrodes were replaced on the torso, while precordial leads electrodes remained untouched. This positioning according to our new labelling was a reference positioning named as 00000_ECG.

[3-14]: Twelve 12-lead ECGs with Mason-Likar positioning while moving only one electrode at a time. Red electrode (RA) position, yellow electrode (LA) position and blue electrode (V2) position has been changed one by one to places: 2, 4, 6 and 8. ECGs: 20000, 40000, 60000, 80000, 02000, 04000, 06000, 08000, 00200, 00400, 00600, 00800, were recorded.

[15-22]: Eight 12-lead ECGs with Mason-Likar positioning while moving the whole set of 4 electrodes (RA, LA, LL and G) up-down, left-right, and diagonally, while keeping the V2 electrode in “0” position (Fig.I.4_2). ECGs: 46082, 82046, 64028, 28064, 37073, 73037, 11055, 55011, were recorded.

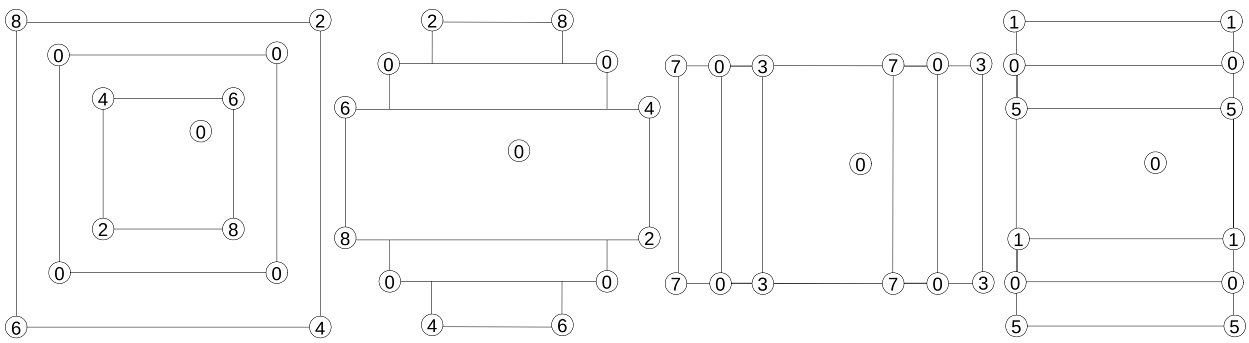


Fig.I.4_2 Electrodes’ positioning while moving a set of 4 electrodes

We used the Lyon program to calculate global and lead by lead parameters of each recorded 12-lead ECG and we compared these parameters with the normal limits of the 12-lead ECG in white Caucasian males of age 18-29, as our two subjects fitted this range. The obtained results (annex 4.1) demonstrate that none of our 2 statements are correct. All recorded ECGs are different in their own way whatever the number of electrodes moved. In addition, a linear shift of the electrodes position did not provide linear changes in recorded data. Moreover, the imposed electrodes mal-position, even in those ECGs were the electrodes' positioning was closest to the heart, put a risk that normal limits would be exceeded in some leads.

The same conclusion was reached when we compared the global and the lead by lead parameters of the synthesised 12-lead ECGs with the normal limits of the 12-lead ECG. Synthesised ECGs were obtained taking three displaced leads: I, II and V2 from recorded ECGs and using the Lyon program by applying generic and patient-specific matrixes in order to reconstruct 12-lead ECGs. We used both: a standard and a reference (Mason-Likar positioning) 12-lead ECG to calculate patient-specific matrixes. This analysis results confirmed that all synthesised ECGs were slightly different and we could not observe any linear changes that would permit to correct or to improve the signal's data and so would positively affect the diagnostic information encoded in this data.

In conclusion, this limited to 2 subjects simulation demonstrates that electrode displacement provides non-repeatable and therefore not correctable signal changes. Hereof, the human's body surface complexity shall be taken into account while designing a sensor-system, as only then a recorded ECG would suit for the digital diagnostical evaluation.

4.3. Sensor-system modalities design

Differences in subject's morphologies lead us to design a sensor-system that has a couple of versions (modalities) and so could be better adaptable to different figure users. We also took in account that this sensor-system should be used out of hospital, that is in different daily life conditions, such as at home, at work or while travelling. This called us to follow the experience of the EPI_MEDICS project and to choose the same reduced set of electrodes consisting in four active electrodes: RA – right arm, LA – left arm, LL – left leg and V2 all placed on the torso (as the Mason-Likar positioning defines). The fifth electrode, called ground electrode and used to remove interferences can be placed anywhere on the body, though, usually, it is placed 2 cm above the right pelvic bone. Finally, as the sensor-system is intended to be used by a profane user with no additional help from a specialist, its application and positioning shall be taken in account. Hereof in our system we have chosen a V2 electrode to be a reference point of the system at the approximate centre of it and to be the only point where user intervention is needed. That is, any user shall be able to find by himself (or with the help of an assisting person in case of elderly people) this central point as it is defined – 4th inter-costal space, 1-2cm to the left from the sternum while he is in a rest position.

Our proposed sensor-system (Fig.I.4_3) is based on the following hypotheses:

1. Three home-care sensor-system modalities: S, M and L, which respectively roughly correspond to the clothes sizes Small, Medium and Large (Table I.4.1), could be systematically provided to any user of a personal ECG recorder without significant added cost.

2. The user can then select the size modality which is the most convenient for him, provided that he will always use the same modality in order to preserve serial analysis results integrity.
3. The sensor-system must be easy to wear for self-care.

Our sensor-system template is made from three semi-rigid strips of different lengths joined together at one point where the electrode V2 is fixed. These strips among themselves keep fixed angles and each of them has three electrodes fixed at the end. These - let's call them tail-end – electrodes taken one by one from each strip can form triangles of different modalities. This sensor-system design can be easily expanded by prolonging strips and adding more electrodes at their ends. This would enable to form more triangles if for instance XL (extra large) and XXL (extra extra large) sizes are needed. In this thesis we used a sensor-system version of only three sizes' for experimental purpose.

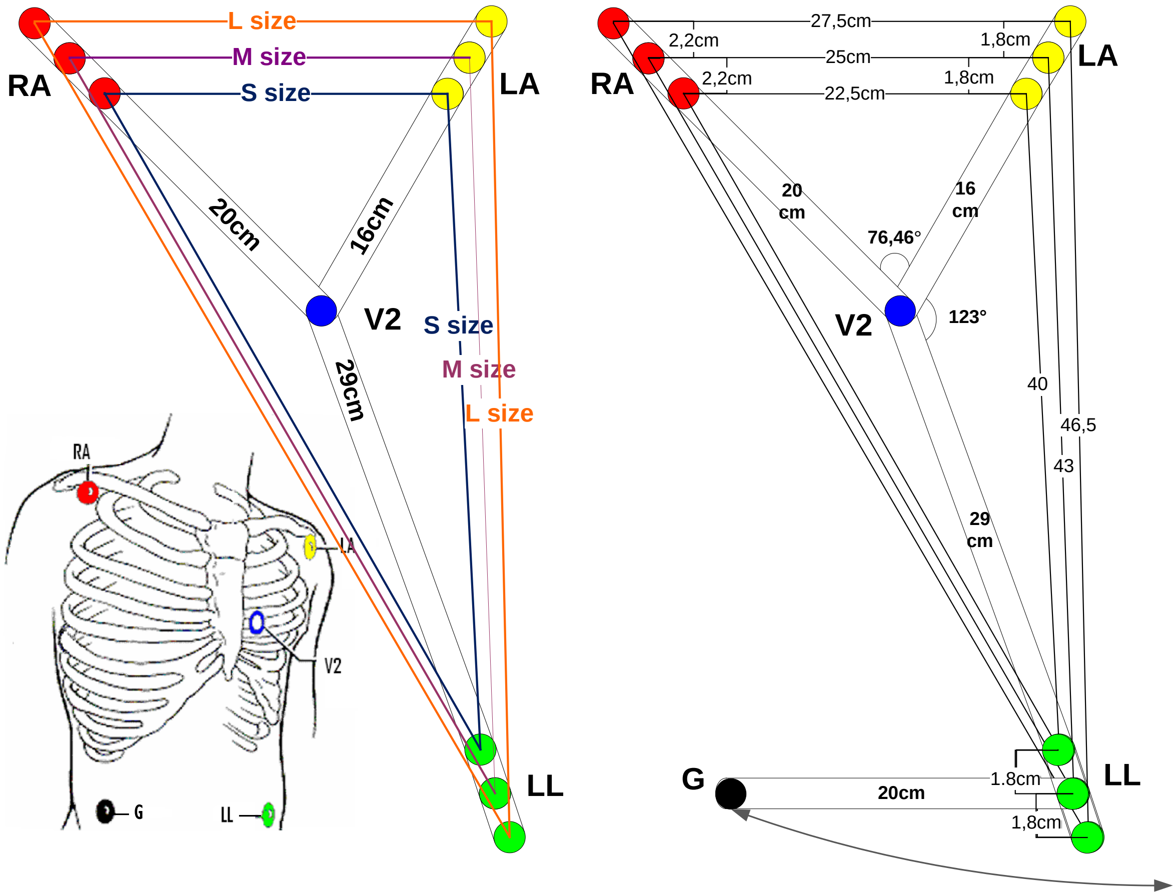


Fig.I.4_3 Design of the sensor system modalities

	V2 to RA (cm)	V2 to LA (cm)	V2 to LL (cm)	LL to G (cm)
S size	*20*	*16*	*29*	*20*
M size	*22.2* (S size +2.2)	*17.8* (S size +1.8)	*30.8* (S size +1.8)	*20*
L size	*22.4* (M size +2.2)	*19.6* (M size +1.8)	*32,6* (M size +1.8)	*20*

Table I.4.1 Four stripes measurements

To reduce interferences, any of the so called active electrodes can be taken as the ground electrode or a separate one can be placed independently from the system. However, one more strip of 20 cm can be fixed at the LL electrode and the ground electrode can be attached to it. If this strip would have a flexible angle of 180 degrees, then, the ground electrode could be placed not only on the front of the torso, but also on the back of the torso. The back torso placement (if this does not change the recorded ECG characteristics) could be useful for skin impedance measurement, which could be taken at the same time as the ECG recording, for example using 3 electrodes: S of the LL size, M of the LL size and G on the user's back. The application of our proposed sensor-system on the torso is adapted to a profane user. At first, the user needs to find the 4th intercostal space on his chest with his fingers and to attach the middle electrode (V2) at about 1-2 cm to the left from sternum in that 4th intercostal space. Then he shall slide the two upper stripes towards his/her shoulders to attach the 3 electrodes marked as RA and the 3 ones as LA, where R indicates the right and L the left side. At last, he shall slide down the stripe named V2-LL to the direction of the left pelvic bone and shall attach the 3 remaining electrodes.

An additional stripe or a chain helping to fix the sensor-system in around the neck (like a bib or an apron) can be attached to this sensor system. It would serve for the sensor-system placement and adjustment and would ensure the stability of the sensors position. If this additional stripe would be attached to the sensor-system, the user would only require to find the place where the middle electrode corresponding to V2 shall be placed and to carefully slide all stripes with the electrodes at their ends to adjust the length of the stripe/chain in around the neck. This user-specific adjustment could be kept for future recordings and so would facilitate and speed sensor-system application up.

Our proposed sensor-system supports the ability to define an appropriate sensor-system size adapted to the context of use, in our case the morphological difference of each subject. It encapsulates personalisation objective, which guides to improved health care services. Hereof, the proceeding of our research aims to establish for each sensor-system user the most appropriate sensor-system size, which produces ECG signals with a diagnostic content that is closest to that of a standard ECG. In order to evaluate a possible automation of this selection process, we performed a series of experiments.

4.4. Study population

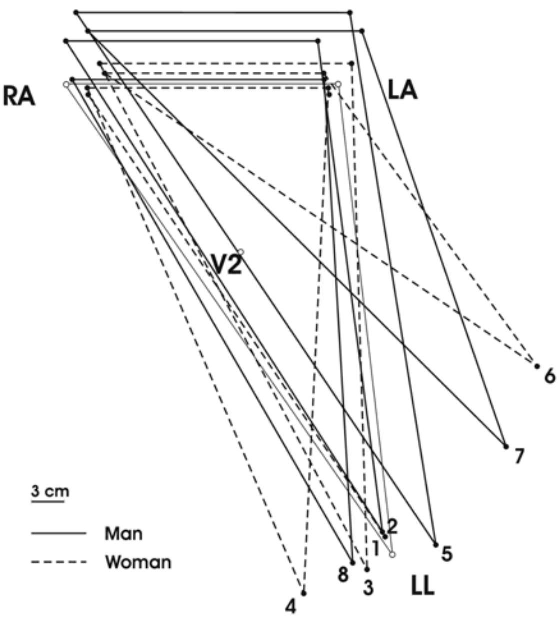
A group of 8 healthy subjects (4 male and 4 female) representing the diversity of human morphology was set-up. We registered their age, height, weight and calculated their body mass index (BMI). Additionally, we measured bust and under-bust girths for women (B) and measured a chest girth and a girth at the 5th intercostal space for men (C). Subjects were asked of what clothes' size they typically wear for the top part. The European system of clothes sizes: small (S), medium (M), and large (L) was used to record this information. We have measured in centimetres the distances among the four key electrodes RA, LA, LL and V2 when they were placed at the exact Mason-Likar positioning. Five lengths in between electrodes: RA-LA, RA-V2, LA-V2, LA-LL and V2-LL, were recorded for each subject (Table I.4.2).

Later on, already during the data acquisition, if there was a possibility, for some subjects we measured their electrical skin resistance. This supplementary information was documented together with notes about skin hairiness for male subjects.

	RA to LA (cm)	RA to V2 (cm)	LA to V2 (cm)	LA to LL (cm)	V2 to LL (cm)
1	*22*	*20*	*16.5*	*40.5*	*29*
2	*23*	*25*	*20.5*	*45*	*28.5*
3	*23*	*21.5*	*20*	*46*	*31*
4	*22*	*20.5*	*17*	*46*	*31.5*
5	*25*	*26.5*	*24*	*49*	*32*
6	*20*	*20.5*	*18*	*33*	*29*
7	*25*	*24.5*	*23*	*40*	*30*
8	*23*	*22*	*17.5*	*44*	*30*

Table I.4.2 Five distances taken among electrodes placed in their anatomically correct positions for each subject

All subjects were of Caucasian type. Their age varied in the range of 25-38 years old with the standard deviation of 4.9 (mean/median 29.1/26). The weight and height fluctuated in the range of 54-92 kg and of 160-184 cm with the standard deviation of 11.7 and of 9.6 respectively (mean/median 73.1/74.5 for the weight and 174.5/179 for the height), what have given the range of BMI in 19.8-27.5 with a standard deviation of 2.8 (mean/median 23.4/23.1). Even though, the BMI is not a very precise measurement, we noticed that if it is evaluated taking in account also the height of the subject the BMI classified pretty well the subject's morphology, from somebody quite slender to a stubby built figure. Finally, the chest/bust girth variation in the range of 82-112 cm with a standard deviation of 8.5 (mean/median 96/95.4) is rather related to men/women differences than to a variation in the figure of men or women. We noticed that the morphology of the torso can be better (though still roughly) evaluated by taking the difference between the bust and the under-bust (B) for women and the chest and the 5th intercostal's girth (C) for men. All these measurements gave a pretty good, however quite approximate, picture of the subjects' morphology. The real measurements are displayed for each subject in Fig.I.4_4.



1	*Woman, 26y, 165cm/54kg, BMI=19.8, B:82cm/70cm. High skin R. Clothes size: S*
2	*Man. 26y, 180cm/77kg, BMI=23.8, C:95cm/91cm. Not hairy chest. Clothes size: M*
3	*Woman, 26y, 180cm/74kg, BMI=22.8, B:99cm/80cm. Low skin R. Clothes size: M/L*
4	*Woman, 32y, 165cm/75kg, BMI=27.5, B:95cm/80cm. Average skin R. Clothes size: L*
5	*Man, 38y, 184cm/92kg, BMI=27.2, C:112cm/105cm. Hairy chest. Clothes size: L*
6	*Woman, 25y, 160cm/60kg, BMI=23.4, B:90cm/75cm. Clothes size: M*
7	*Man, 26y, 178cm/80kg, BMI=25.2, C:94cm/91cm. Hairy chest. Clothes size: M*
8	*Man, 34y, 184cm/73kg, BMI=21.6, C:96cm/90cm. Not hairy chest. Clothes size: S/M*

Fig.I.4_4 EPI-MEDICS project based sensor-system placement at Mason-Likar positioning for 8 subjects superposed at V2 and a brief description of each of the subject

4.5. Data acquisition

The data recording procedure involved the recording of 6 ECGs for each of our 8 subjects. The subjects were asked to remove the upper part of their clothes, to take off shoes and socks and to lie down in a supine position. A small pillow was placed under the head of the subject. None of additional skin preparation, such as shaving or cleaning the skin was performed. All ECGs were recorded using a lightweight Microtel electrocardiograph, applying 10 standard electrodes (Ag/AgCl) for each ECG recording together with press studs connectors. The ECGs sampling rate was 500 samples per second with an amplitude resolution of 5.0μV.

The data acquisition procedure included the recording of six successive 12-lead rest ECGs on each subject:

S1-ECG: a standard 12-lead ECG.

ML-ECG: a 12-lead ECG with the limb electrodes in the Mason-Likar position.

S-ECG, **M-ECG**, **L-ECG**: 3 12-lead ECGs by using our 3 sensor-system templates.

S2-ECG: an additional standard 12-lead ECG.

During all six recordings, the standard 6 precordial electrodes (V1 to V6) placed on the torso remained in the same position. The 4 limbs electrodes (RA, LA, LL and RL(G)) however were moved following the applied positioning scheme (standard, mason-likar, S, M L and standard). The ground electrode for the first and the last recording (both standard ECGs) was placed on the right leg, for the remaining 4 recordings it stayed as it was placed for Mason-Likar positioning – 2 cm above the right pelvic bone.

The same set of six ECGs was recorded again twice (roughly two months later and then three weeks later) for three of the female subjects (subjects' numbers: 1, 3 and 4) in order to assess the reproducibility of the obtained results.

4.6. Data analysis method

Our performed data analysis aimed to evaluate the diagnostic quality of S-ECG, M-ECG and L-ECG of each subject. This evaluation was performed on the basis of seven descriptors, which we defined in order to assess the similarity of S-, M-, and L-ECGs in comparison with the 1st and the 2nd standard 12-lead ECGs (S1-ECG and S2-ECG) as well as with the 12-lead ECG at Mason-Likar positioning (ML-ECG). To meet home-based self-care requirements, we decided to use only the 3-lead subset (I, II and V2) from the S-, M- and L-ECGs, as we expect that these three leads could be personally recorded by a profane user in home conditions. However, in order to be able to compare these reduced lead ECGs with 12-lead reference ECGs, we applied the Lyon program and several data reconstruction strategies to synthesise 12 leads out of the 3-lead S-, M- and L-ECGs, to get comparable data sets.

For clarity, the data analysis that we performed can be divided in 3 generic steps:

- data processing,
- data reconstruction,
- data comparison.

Data processing: all recorded ECGs were processed by means of the EPI-MEDICS software factory *[Fayn'10]*, which implements the Lyon program for data filtering, waves' typification, delineation and global as well as lead-by-lead measurements computation.

Data reconstruction: a typical data reconstruction procedure requires transformation coefficients, which allow obtaining output data out of the input data. In our case the input data was a 3-lead ECG (I, II and V2), recorded with electrodes placed approximately in the Mason-Likar positions, and the desirable output was a 12-lead ECG as much as possible similar to the standard 12-lead ECG, recorded with electrodes placed in standard positions. To obtain this output we have considered two data synthesis methods: linear regression (REG) and artificial neural nets (ANN). However, as the patient-specific transformation is usually more powerful than the generic transformation and reconstructs the standard 12-lead ECG more accurately (cf. chapter 2), we only used the computation of a patient-specific transformation to analyse our data.

The development of a data reconstruction strategy is one of the several crucial steps in the process towards a successful final result – a diagnostic quality ECG. It shall describe how transformation coefficients are obtained and how they are used to reconstruct the data out of some data subset. It shall also consider all available contextual information such as the nature of input data, the source data from which the patient-specific transformation coefficients might be calculated, and finally, the compatibility of the input and source data. The data reconstruction strategy shall identify different parts or data: a) taken from original data, b) calculated using approved and standard to all data derivations, c) synthesised data.

In our specific case, we should consider that, taking the 3-lead subset (I, II and V2) as an input and keeping it as an original part of the data will require us to synthesise only 5 leads: V1, V3-V6, while the remaining 4: III, aVR, aVL, aVF, would be calculated out of I and II leads, as it is usually performed in standard electrocardiographs. We should also pay attention that even if our goal is to reconstruct 12-lead ECGs as much as possible similar to the corresponding standard 12-lead ECG, the position of the electrodes used to obtain the 3-lead subset and the 12-lead ECGs differs. The home-based sensor-system modalities S, M and L are closer to a Mason-Likar electrodes' positioning than to the positioning of electrodes for a standard 12-lead ECG.

Two strategies have been implemented for the calculation of the PSM patient-specific matrix (Fig. I.4_5). The first strategy, called $\mathcal{M}_\alpha$, has already been used in the EPI-MEDICS project and by Hussein Atoui *[Atoui'06]*. It needs only one ECG to compute the transformation coefficients – a standard 12-lead ECG (S1-ECG in our case). Meanwhile, the second strategy, $\mathcal{M}_\beta$, was created in the scope of this thesis and needs two ECGs – a standard 12-lead ECG (S1-ECG) and a rest 12-lead ECG (ML-ECG in our case), as an input.

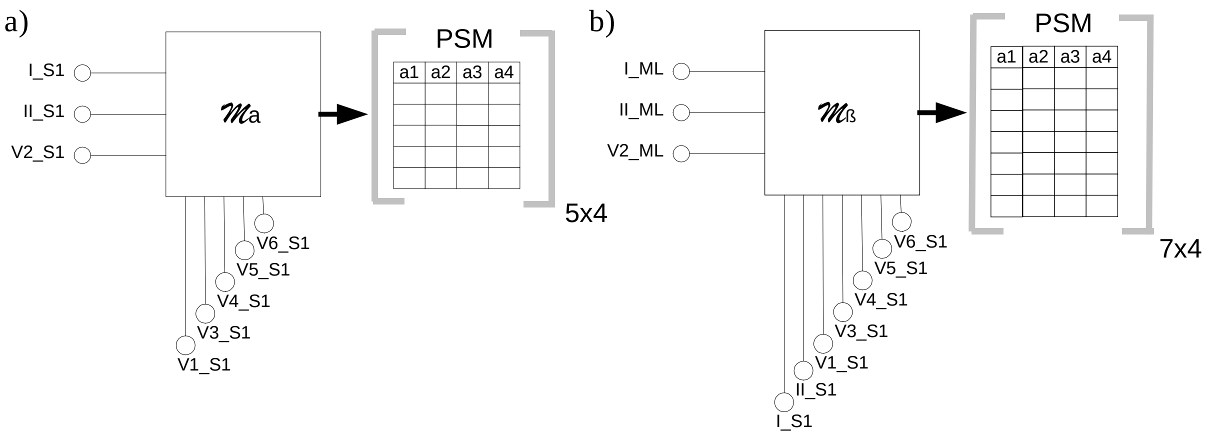


Fig.I.4_5 Two different models for the calculation of patient-specific transformations. Here S1 indicates a standard 12-lead ECG and ML a Mason-Likar 12-lead ECG

Following the $\mathcal{M}_\alpha$ strategy, the transformation coefficients are computed (Fig.I.4_5a) by means of a mapping between the five leads–V1, V3, V4, V5, and V6 of S1-ECG, and the three remaining leads I, II, and V2 of the same S1-ECG –, as these three leads are the ones that our proposed home-based sensor-system would record. A linear mapping method, such as linear regression, then would result with a patient-specific matrix (PSM), while a non-linear method, for example an ANN based method, will provide a patient-specific neural net (PSNN). The strength of the $\mathcal{M}_\alpha$ strategy lies in its capacity to encapsulate the knowledge that explains the relationships between the standard ECG leads specific to one patient in a set of linear transformation coefficients. These transformation parameters afterwards can be used and reused in self-care conditions aiming to reconstruct a 12-lead ECG out of a 3-lead ECG.

Our proposed $\mathcal{M}_\beta$ strategy differs from the $\mathcal{M}_\alpha$ strategy as it is taking not five, but seven S1-ECG leads – leads I, II, V1, V3, V4, V5, and V6 - as outputs to be reconstructed, and expresses them not via three S1-ECG leads, but via three ML-ECG leads – leads I, II, and V2, as inputs (Fig.I.4_5b). Here again, these concrete three leads I, II and V2 of the ML-ECG are chosen because these same three leads are those that are obtained with home-based ECG recording modalities. Moreover, the Mason-Likar electrodes placement is closer to the S, M and L modalities placement than the standard positioning used in the $\mathcal{M}_\alpha$ model. Therefore, the $\mathcal{M}_\beta$ model tends to include the transition pathway among different electrodes placements and so should be more accurate. The strength of the $\mathcal{M}_\beta$ strategy lies in the capacity to encapsulate not only knowledge of relationships between leads of the same ECG as it is done in the $\mathcal{M}_\alpha$ model, but also between two patient-specific ECGs, which were obtained with a different electrodes' placement. For the transformation coefficients calculation both methods: linear and non-linear can be used.

Hereof, for each subject we took their S1-ECG and using the $\mathcal{M}_\alpha$ model we computed two transformation parameters' sets: PSM_α_REG and $PSNN_\alpha$_ANN. The first of them was calculated using the REG method and the second one by means of the ANN method. Afterwards, we took S1-ECG and ML-ECG from the recorded ECGs set for each subject and using the $\mathcal{M}_\beta$ model we have calculated again two transformation parameters' sets: PSM_β_REG and $PSNN_\beta$_ANN. All these transformation parameters' sets are reusable and are typically applied when a new 3-lead ECG is recorded with a personal ECG recording modality. That is, if the personal recording modality remains the same and there is no newer standard or reference ECG, the transformation coefficients computation process is not repeated. For the transformation computations, we used parts of the Matlab programs developed by Hussein Atoui *[Atoui'06]*, which we adapted and completed to fit our specific needs.

When the sets of transformation coefficients have been computed, the signal synthesis can be performed using the corresponding linear or non-linear method. Following our experiment's procedure, for each subject we took the three leads I, II and V2 from the 12-lead S-ECG as input to the synthesis procedure and we applied all four transformation parameters' sets. The obtained sets of five reconstructed leads (according to the $\mathcal{M}_\alpha$ strategy) were complemented with the three leads used as input to finally form the set of eight principal leads. Meanwhile, the obtained sets of seven reconstructed leads (according to the $\mathcal{M}_\beta$ strategy) were complemented only with an originally recorded V2 lead. The remaining four leads in both cases were derived from these eight leads, what, finally, resulted in 4 synthesised 12-lead ECGs out of each initial 3-lead S-ECG. The same procedure was repeated with leads I, II and V2 taken from the 12-lead M-ECG and the 12-lead L-ECG signals, therefore at the end we had 12 synthesised ECGs for each subject.

All new ECGs were also processed using the same Lyon program.

In real cases the most suitable strategy should be selected according to the context of use. Then only if there is no previously recorded standard 12-lead ECG, a generic transformation should be applied. However, if there exists a previously recorded standard 12-lead ECG, a patient-specific transformation following the $\mathcal{M}_\alpha$ strategy should be computed and used. Finally, if there exist a standard 12-lead ECG and a reference 12-lead ECG corresponding to the home-based modality, the ML-ECG is our case, the $\mathcal{M}_\beta$ strategy can be used.

Data comparison: three types of ECG comparisons: C1, C2 and C3, will be performed between the typical beats of the original and synthesised ECGs using seven descriptors.

- (C1) by applying $\mathcal{M}_\alpha$ to the S, M and L ECGs, and by comparing the newly reconstructed ECGs to the ML-ECG;
- (C2) by applying $\mathcal{M}_\beta$ to the S, M and L ECGs and by comparing the 7 reconstructed leads to homologous leads of S1-ECG;
- (C3) identical to C2, but by taking S2-ECG instead of S1-ECG for the comparison.

The choice to compare the ML-ECG with synthesised 12-lead ECGs in the comparison study named C1 might look questionable as one might expect this comparison to be performed with a standard 12-lead ECG, S1-ECG or S2-ECG for instance, like it was performed in the EPI-MEDICS project. However, the $\mathcal{M}_\alpha$ strategy cannot reconstruct well the standard 12-lead ECG out of S, M and L ECGs, as it does not take in account the RA, LA and LL electrodes displacement knowledge. The $\mathcal{M}_\alpha$ model encapsulates only knowledge of relationships among leads and therefore taking in account that the reconstruction was performed out of 3-lead S, M or L ECGs, which were positioned close to the Mason-Likar positioning, the best result that this reconstruction can give is the ML-ECG, which then would be the “perfect” S, M or L ECG. Meanwhile, the $\mathcal{M}_\beta$ strategy takes in account the displacement of the RA, LA and LL electrodes and indeed reconstructs a standard 12-lead ECG, so it is reasonable to compare the synthesised ECGs with the standard 12-lead ECG: on one hand with the S1-ECG, as it was used to compute the transformation coefficients, and on the other hand with a completely independent standard 12-lead ECG, which in our case is the S2-ECG.

Paired t-tests will also be performed among the 3 types of ECG comparison studies C1, C2, and C3. Finally, we will also perform a repeatability test on the data of 3 subjects with the aim of checking the stability of the data and of data analysis over time.

Three different descriptors' types are selected for comparing original and synthesised ECGs:

- cross-signal measurements, via the correlation (r) and the RMS values(M) of the differences between the two signals to be compared,
- standard ECG measurements differences, via the differences of the QRS axes (QRSaxis[1]) in the frontal plane and the differences of the QRS peak-to-peak amplitudes (K) measured on each standard lead,
- ECG-based values of the output of a committee of 100 ANN measuring the probability of the risk of an acute Myocardial Infarction (MI) (p). The ANN configura-

[1] QRS axis is expressed in degrees. It indicates the average direction of the ventricular depolarisation wave in the frontal plane, with reference to the horizontal plane.

tion parameters were determined and assessed within the EPI-MEDICS project on a large database of 3000 patients admitted in an emergency department *[Atoui'06]*.

Seven final descriptors are then computed for comparison of the 8 principal leads of the original and synthesised ECG and their values are normalised between 0 and 1 in order to facilitate the results comparison (Table I.4.3).

Formula	Description	Notation		Scaling
		ANN	REG	
$\sum_{\{leads\}}(1-r)$	*Sum of lead by lead correlation values subtracted from 1*			/0.5
$\sum_{\{leads\}} M$	*Sum of lead by lead root mean square values*			/1200
$1-\underset{\{leads\}}{Median}(r)$	*Median of lead by lead correlation values subtracted from 1*			/0.06
$\underset{\{leads\}}{Median}(M)$	*Median of lead by lead root mean square values*			/140
$\Delta \underset{\{leads\}}{Median}(K)$	*Difference between the medians of the lead by lead QRS peak to peak amplitudes from the original and the synthesised ECGs*			/1300
$\Delta QRSaxis$	*Difference between the QRS axis values of the original and the synthesised ECGs*			/20
Δp	*Difference between the probability of MI of the original and the synthesised ECGs*			/0.4

Table I.4.3 The seven final descriptors used to compare original and synthesized ECGs

In order to obtain optimal correlation and root mean square (RMS) values computation, our proposed technique encompasses four main stages (Fig.I.4_6):

- Base line correction of each of the compared signals.
- Alignment of the compared signals, as they were not recorded simultaneously.
- Intermediate decision, which selects the best couple of correlation and RMS values produced at the alignment level.
- Final decision, which selects the best couple of correlation and RMS values produced at the level of different base-line corrections.

1st stage: The input to this algorithm is composed from two 12-lead ECGs that are intended to be compared: the synthesised ECG and the reference ECG. The reference ECG depends on the comparison study. At first, the technique aims to correct the base lines for each signal[2] separately. As there can be several different base-line corrections methods, this process will produce one new ECG per base-line correction method. In our algorithm we took in account two methods: a partial correction (p) and a complete correction (c). The partial correction aims to shift the whole signal upwards or downwards, while the complete correction modifies the whole signal progressively by subtracting the calculated base line. Hereof, in the first stage of our algorithm we construct 4 couples of corrected reference and synthesised ECGs using a simple recursion method to form all possible pairs: Ref_p and Syn_p, Ref_p and Syn_c, Ref_c and Syn_p, and, finally, Ref_c and Syn_c (p – partial, c – complete method).

[2] with signal here and afterwards in this method explanation we mean the QT segment of the typical beat

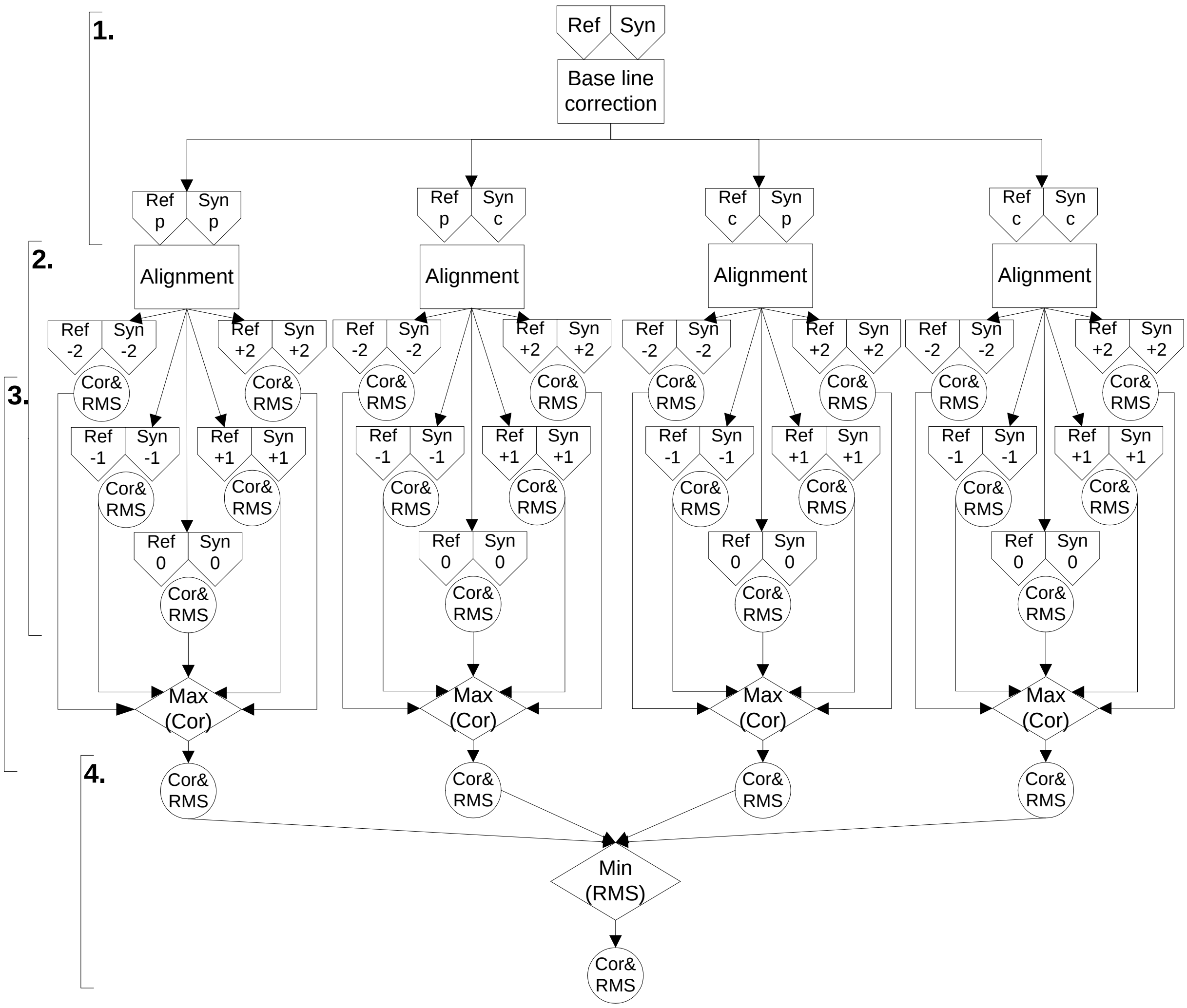


Fig.I.4_6 Optimal correlation (Cor) and root mean square (RMS) values computation method for typical beats of the synthesised and the reference ECGs comparison

2nd stage: As we have already mentioned above, the reference signal and the synthesised signal were not obtained simultaneously. Therefore, an alignment of the signals is an essential step, especially for obtaining realistic cross-signal parameters, such as correlation and RMS parameters. A cross-correlation method is used to calculate the signals' shift (one according to the other). The calculation is performed on the Qonset-Qoffset segment from the V2 leads of both signals. The V2 lead, here, was selected purposely as it is not affected by the synthesis algorithm and also because during the recording sessions, the V2 electrode was never moved while recording all six ECGs used in this experiment. However, here again we chose to let correlation and RMS values to have a decisive power. Thus, after having aligned the two signals to be compared on the basis of the signals' shift resulting from the cross-correlation computation of the Qonset-Qoffset segments of the V2 leads, we introduce a set of 5 additional shifts varying from -2 points to +2 points around the already aligned signals, by steps of 1. This modification will result into 5 slightly different though almost aligned couples of reference and synthesised ECGs segments, which for we then calculate the correlation and the RMS values for each of the 8 principal leads.

3rd stage: The first decision at stage 3 is to select between these 5 sets of results the one which corresponds to the best aligned couples of signals. The decisive parameter in this step is the correlation value. We have chosen as the best out of 5 couples the couple which produces the maximum median correlation value of 7 lead by lead (I, II, V1, V3 to V6) correlations. The V2 lead correlation is excluded here as in both ECGs (the reference and the synthesised) they are supposed to be practically the same (no electrode changes, no synthesis method applied). This decision is repeated at each of the 5 ECG couples results to the one best couple per different base-line correction.

4th stage: The final decision selecting the best couple and so the optimal correlation and RMS values for the originally given two ECGs is based on the minimum value of the median RMS value of the 7 lead by lead RMS (I, II, V1, V3 to V6) values. Again the V2 leads comparison is excluded as their natural similarity may influence the decision. The correlation and RMS values provided by the chosen couple are the foundation values on which the computation of the first four descriptors: correlation median, correlation sum, RMS median and RMS sum out of eight leads are finally calculated.

4.7. Results obtained with comparison studies: C1, C2 and C3

In the previous section we have defined three comparison studies, which we performed on series of ECGs obtained from 8 subjects. Despite the principal difference lying in the ECGs that are compared in each study, the comparison method remains the same for all three evaluations and embraces the same 3 sensor-system modalities, the same 2 ECG synthesis methods and the same 7 descriptors that characterise the obtained results. Herein, we, first of all, will briefly describe basic rules about how we will present the results, then, the questions that we have raised to observe the data and, finally, we will overview the results in details. This explicit review on the obtained evaluations will contain two parts for each comparison: observations related to a specific subject and observations related to the totality of the subjects. A global survey on the ensemble of all three comparison studies is presented in the section 4.8.

Each comparison study is based on data grids compounding the calculations for each subject (annex 4.2). It illustrates the results for 8 subjects: 4 women and 4 men. We classified the subjects by gender and we lined both groups from the left to the right according to their clothes size, from S to L respectively. The values of our seven defined descriptors are represented for each modality: S, M and L, one per subject. The closer to zero is the descriptor value the smaller is the error present in the synthesised ECG in comparison with the corresponding reference ECG.

The principal goal of data examination is to determine the sensor-system size that provides the best results (the smaller difference in comparison with the reference ECG) for a particular subject. Then, the detailed characterisation of that choice via the number and the category of descriptors (cross-signal measurement, standard ECG measurement, and interpretation measurement) would indicate the confidence we can have in that choice. Moreover, a superposition of the selected size with the clothes size, as well as the height and the weight of the subject would give insights on a correspondence between the chosen sensor-system and the subject's morphology. The tendency of the descriptor values to rise or to descend monotonically through all sensor-system sizes keeps an interest in observing the possible linearity in the error present in the synthesised ECG and the expansion of the sensor-system size. Finally, the data synthesis method shall be inspected and observations of its performance noted down.

The questioning about each study results should include sensor-system size selection issues, such as: is it possible to determine an optimal sensor-system size for each subject, do these sizes correspond more or less to the morphology of the subjects, or is there a unique size that would fit everyone. It should also describe the variations of the errors in the synthesised ECGs and the synthesis method that overall performs the best, if such a method exists. Additional observations about mutual correspondence of descriptors should be noted as well.

Comparison study C1

The transformation coefficients for the C1 study were computed out of the first standard ECG (S1-ECG) and applied to reconstruct 12-lead ECGs from the 3-lead S-, M-, and L-ECGs. This reconstruction method is based on the knowledge on relationships among the leads of one standard 12-lead ECG. Therefore, taking as inputs the 3 leads I, II and V2 of the S, M and L ECGs tends to reconstruct 12-lead S, M and L ECGs. These three reconstructed ECGs have a common reference – the ECG at Mason-Likar positioning. Hereof, the subject-by-subject results obtained in the C1 study (Fig.I.4_7) are comparisons between the typical beat of the ML-ECG and the typical beat of each of the synthesised S-, M- and L-ECGs of each subject. The closer to zero are the computed descriptors at a particular sensor-system size, the better this size suits to the subject.

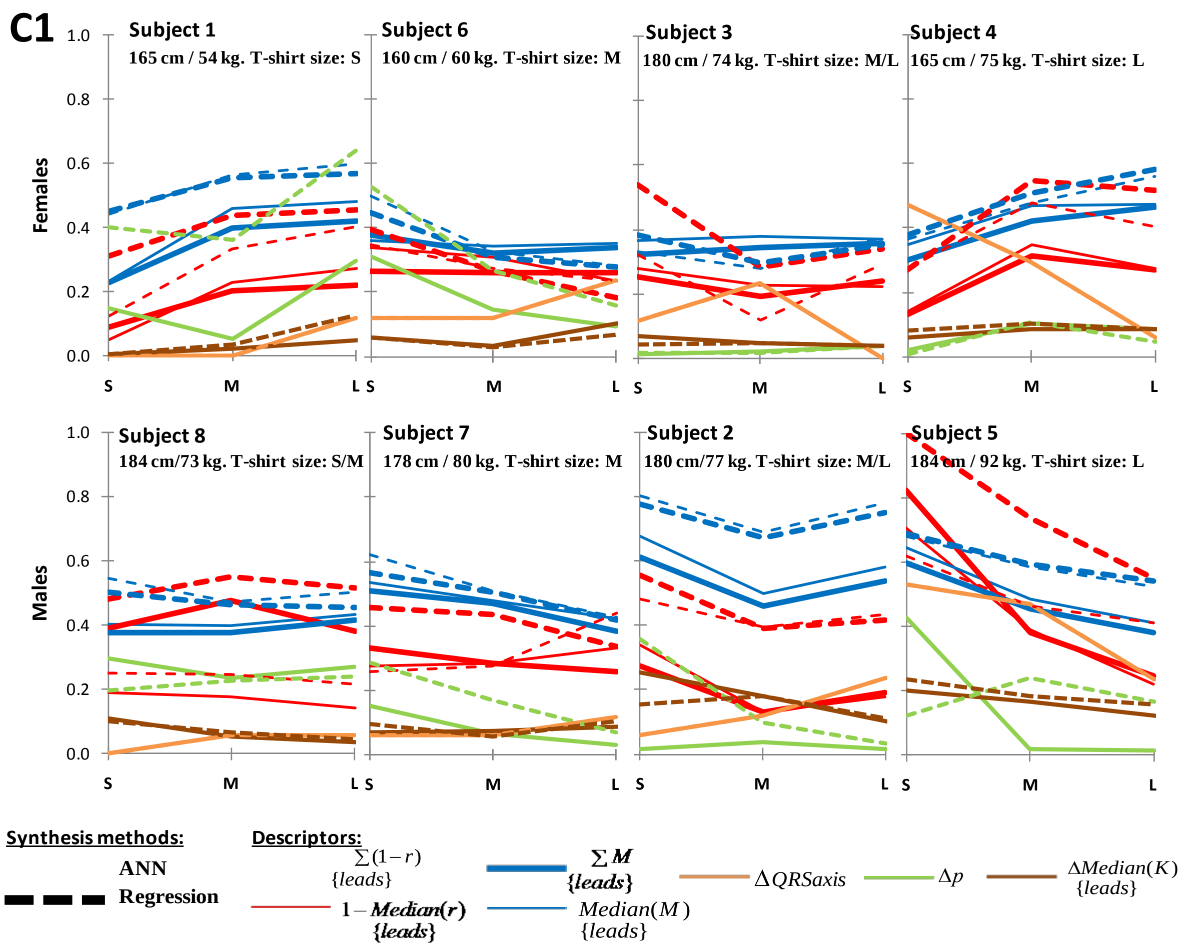


Fig.I.4_7 Normalised comparison results between a 12-lead Mason-Likar positioning ECG (ML-ECG) and 3 simplified ECGs for 8 healthy volunteers: 4 women and 4 men

1st subject: This subject's data analysis results straightforwardly demonstrate that the lowest values for almost all descriptors are obtained for the sensor-system size S, so indicating that this size is an optimal sensor-system modality for this subject. The only parameter that indicates the M size could be better is the difference of interpretation parameters. However, a superficial look at the subject's morphology and at her clothes size would rather confirm the sensor-system of S size selection. There, partially ignoring the interpretation parameter (as the misjudgements are related only with sizes S and M), we even could see a tendency of growing error when the size of the sensor-system augments. Another point clearly supported by all descriptors is the better performance of the ANN synthesis method – the smaller error is present in all reconstructed ECGs when the ANN method is used.

6th subject: This subject's data evaluation seeking to decide the optimal sensor-system size is complicated. The cross-signal measurements together with the difference of interpretation parameters clearly eliminate the S size modality, but diverge for a decision among the M size or the L size as an optimal one. Differences on both standard ECG measurements are smaller with the M size, while the differences of the interpretation parameters as well as of most of the cross-signal measurements, especially in the ECGs synthesised with the linear regression method, demonstrate smaller errors with the modality of size L. Overall, the L size modality would probably be the one to suggest, however the assessment of the subject's morphology contradicts this choice. This female subject clothes size and height would rather fit size S and M as it is indicated than the L size. In summary, we could observe the tendency of a smaller error towards the L size modality in seven parameters. However, four parameters indicate the inverse and the three remaining parameters are more or less smooth, therefore we cannot really recommend an optimal modality for this subject. Even the identification of the best signal reconstruction method out of these data is complicated. The ANN performs clearly better with the S size modality, but for the L size modality, especially for the cross-signal parameters, the ECGs reconstructed using the regression method have a smaller distortion from the ML-ECG.

3rd subject: This subject's data examination indicates the M size modality as the optimal one for this subject. The obtained descriptor values are not completely homogeneous to decide without any hesitations, however, taking into account the morphology parameters of this subject, the M size seems to be the best compromise. The differences between the standard ECG measurements as well as the difference of interpretation parameters do not indicate any preference for a particular modality, as the distortions are pretty small. Meanwhile, the cross-signal measurements are clearly smaller with the M size modality. Only the QRS axis parameter is contrary to the final choice and, actually, performs worse with the M size modality. We have not observed any clear tendency on a growing error for this subject, so at the end we would recommend the M size modality for this subject. The ANN method seems to best reconstruct the ECGs, which are more similar to the ML-ECG, however, for the M size modality, the performance of the ANN method is close to the performance of the regression method for some descriptors and for some of them the regression method overpasses the ANN method.

4th subject: The optimal modality for this subject is the S size one. Only one measurement (the difference of QRS axis) contradicts all other parameters that have smaller values when the reconstructed S-ECG and ML-ECG are compared. Though here, the contradiction that comes is related to the morphology of this female subject. Clothes size L is indicated in the questionnaire for this subject, but her morphological parameters do not fit the modality of the size S. Despite this ambiguity, we would however recommend the size S modality for this subject. The tendency of growing error is seen in a couple of descriptors, but as there are contradictions, this probably should not be taken in account. Overall, the ANN data synthesis method is undoubtedly better in reconstructing ECGs of all sensor-systems' modalities.

8th subject: The diversity in descriptors values obtained from this subject's data throughout different sizes modalities discourages an optimal sensor-system size selection. The values of each separate descriptor do not fluctuate a lot, however some contradictions even in the same group of measurements can be observed. For such type of subjects, we would say that any sensor-system size would do and that the conclusion of course might or might not fit with the clothes size the person wears and with his/her general morphological parameters. We could not really notice a tendency of a growing error, but we could firmly recommend the ECG synthesis based on the ANN method for this subject.

7th subject: The L size modality is optimal for this subject. The standard ECG parameters differences and one of the cross-signal parameters with a small difference would suggest either the S or the M size. However, the majority of the descriptors are in line with the subject's morphological parameters and stand for the L-ECG as the best. The questionnaire of this subject is filled with the clothes size M, though taking in account the BMI and the subject's height the L size would probably be suitable as well. Also, we observe that there is a tendency that subjects with smaller clothes size produce the greatest distortion errors in the comparison results with the ML-ECG. And we can firmly state that the ANN method reconstructs ECGs better than the linear regression method.

2nd subject: Most probably this subject should wear an M size modality sensor-system. The descriptors of the cross-signal parameters solidly confirm this size. It also corresponds to this subject's morphology the best. However, the standard ECG measurements differences contradict each other, as well as the difference of interpretation parameter, which demonstrates the lowest error with the L size modality. A tendency of an augmenting or descending error is not really present in the data of this subject. However, we could unambiguously see that the ANN method is the most accurate in reconstructing all ECGs.

5th subject: The examination of the data quite easily indicates the L size sensor-system modality as the optimal one for this subject. There exists a small misjudgement related to the interpretation parameter, though this could be rather related to the reconstruction method than to the sensors placement itself. A descending error tendency is present in this subject data, what correlates well with the morphology of this subject. In almost all parameters we could state that the ANN method is the one to be chosen.

	Females				Males			
Subject	1	6	3	4	8	7	2	5
Clothes size	*S*	*M*	*M/L*	*L*	*S/M*	*M*	*M*	*L*
Modality	*S*	-	*M*	-	*any*	*L*	*M*	*L*
Method	*ANN*	-	-	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*

Table I.4.4 C1: optimal sensor-system modality and ECG synthesis method suggestion for each subject

The global observations results of the C1 study presented in table I.4.4 show that for almost all subjects we could recommend one sensor-system size. However, the certitude of this choice (not measured) varies pretty much. There probably cannot be one unique size for all subjects as in some cases the descriptors' values demonstrate quite high errors. The tendency of a smaller or greater error in relation to a smaller or a bigger sensor-system size remains occasional and depends on the subject. Overall, the ANN method seems to reconstruct the ECGs better than the regression method, as it has been already reported by Hussein Atoui *[Atoui'06]*. We have not noticed a specific difference between the female and male groups and would think that descriptors values are specific to each subject and form a subject specific profile.

Comparison study C2

We used the same transformation coefficients for the C2 and C3 comparisons' studies. They were calculated out of two ECGs: the first standard 12-lead ECG (S1-ECG) and the 12-lead ECG recorded at Mason-Likar positioning (ML-ECG). The obtained coefficients were applied to reconstruct the standard 12-lead ECGs from the S, M, and L 3-lead ECGs. This reconstruction method is based on the knowledge of the leads relationship in the ML-ECG and also of the effect of the Mason-Likar electrodes displacement on the standard 12-lead ECG. Therefore, the subject-by-subject comparison we performed (Fig.I.4_8) is between the typical beat of the standard ECG and the typical beat of each of the synthesised 12-lead ECGs derived from the S, M and L ECGs. In the C2 study we took the first recorded

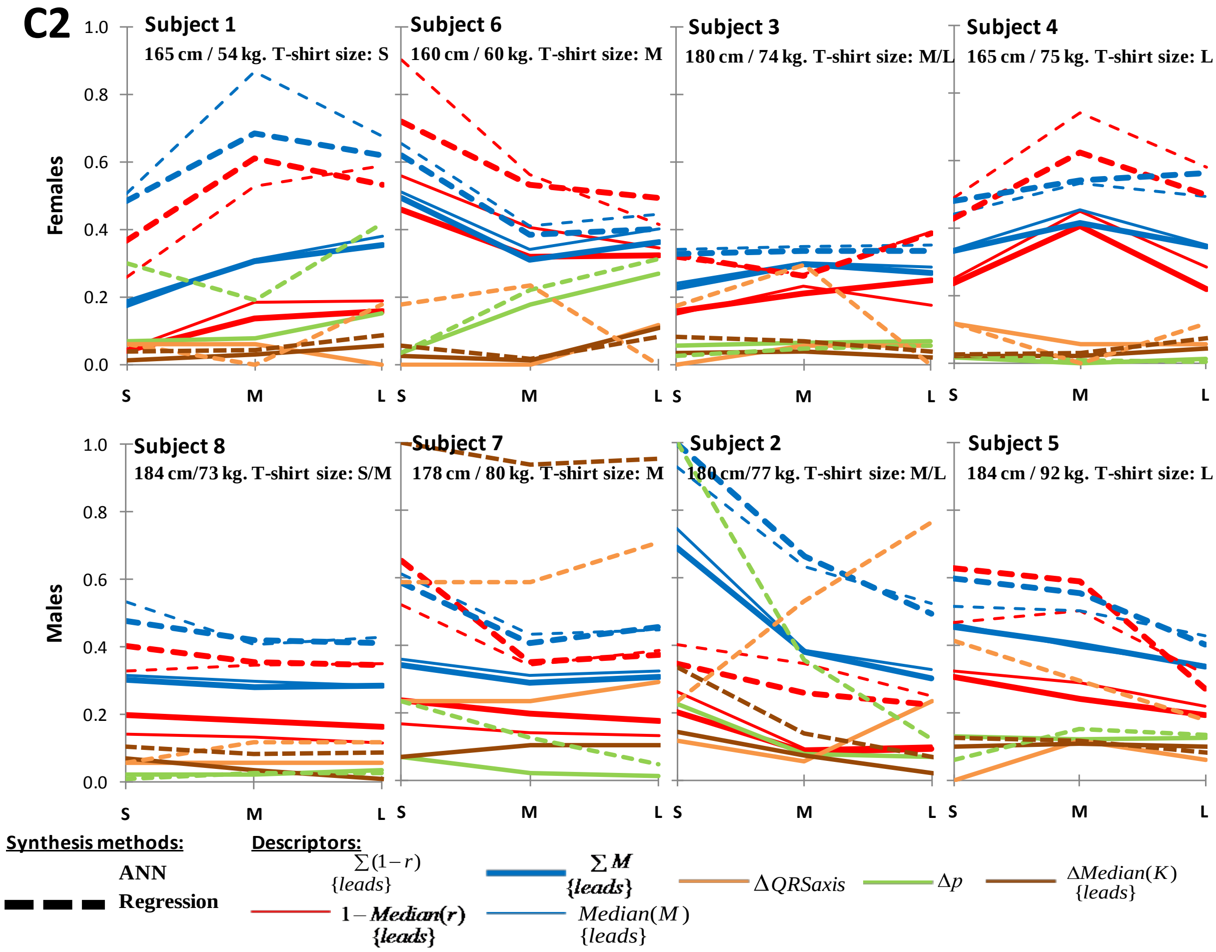


Fig.I.4_8 Normalised comparison results between a first standard 12-lead ECG (S1-ECG) and 3 simplified ECGs for 8 healthy volunteers: 4 women and 4 men standard ECG – S1-ECG - to be compared with, while in the C3 study we used the second standard ECG – S2-ECG (newer used for any transformation coefficients calculation).

1st subject: The examination of this subject's data suggests the size S as the optimal sensor-system modality for this subject. This choice is confirmed by all cross-signal measurements and in part by the differences of the interpretation parameters and the standard ECG measurements. It is also confirmed by the subjects' morphology and the clothes size the subject wears. However, the interpretation parameters differences as well as the differences of QRS axis, when obtained using the linear regression method, are indicating rather the M size modality as the optimal one, which however gives a quite great error for all cross-signal measurements. The QRS axis difference obtained using the ANN method is the only descriptor which indicates the L size modality as being the best. We could not observe any tendency of a growing or descending error with respect to expanding modalities. And, with the exception of the measurement of the QRS axis difference on the ECG reconstructed from the M-ECG with the regression method, which provides a very small error, the ANN method certainly provides the smallest error for all descriptors and for all size modalities.

6th subject: The final modality choice based on the data of this subject is quite complicated. Indeed, even though the measurements' groups agree inside the group, their indications vary depending on the group. The measurements in the cross-signal group have the smallest errors with the L size modality, the interpretation parameter – with the S size modality and the standard ECG measurements' group – with the M size modality. The cross-signal measurements and the interpretation parameter have the tendency to augment or decrease with expanding modalities. However, their variations are completely opposite to each other. We could not advice a concrete sensor-system modality for that subject. Though, we could state that the ANN reconstruction method performs better than the linear regression method.

3rd subject: The judgement on this data resembles the previous case – we cannot really distinguish one optimal sensor-system modality for this subject. Yet, the data fluctuation does not seem significant and keeps pretty smooth, what allows us to advice any of the sensor-system sizes for this subject with the prejudgement that any of them would provide results of more or less similar quality. We could not really see a tendency in the error while moving through the sensor-system sizes axis, though we would state as for the previous subject that the ANN method is more accurate than the linear regression method.

4th subject: The comparison data obtained from the recordings of the 4th subject clearly eliminates the M size modality, which can be qualified as the worst one, and keeps the S and L size modalities which are pretty equivalent one to the other. All cross-signal descriptors, as well as the differences of the interpretation parameters and of the QRS peak to peak amplitudes show a smaller error with the size S modality and only the difference of the QRS axes is smaller with the L size. A simple examination of the person's morphology, however, would rather stand for the size L modality, so augmenting the chances to be chosen as the optimal one. Therefore, for this subject we would recommend both the S and the L sizes modalities, which both should ensure similar quality ECGs' recordings. We could not observe any trend in the error which could be related to the modality size. However, the rule that the ANN reconstruction method performs better is confirmed with this subject's data as well.

8th subject: This subject's data display a very clear decreasing error pattern, which is getting smaller while modalities are changing from the size S towards the size L. Two sensor-system modalities, M and L, tend to be pretty similar in their descriptors' values, however with a small difference the L modality might be a little better. The subject's morphology would rather indicate the M size, though. Considering this subject's height, we could recommend as a first choice the L modality and then as a substitute the M modality. Here again, the ANN method performs much better than the regression method.

7th subject: The recommendation for the 7th subject is uncertain as different groups of descriptors suggest different modality. The interpretation descriptor is indicating the L size modality as the best. However, the cross-signal descriptors agree on eliminating the S size modality, but diverge in choosing the best modality while different data synthesis method was used. The cross-signal descriptors obtained with the linear regression method are better with the modality M, while the cross-signal descriptors obtained with the ANN method demonstrate a smaller error with the L modality. Finally, the standard ECG measurements, although the difference is small, perform better with the S modality. Overall, taking into account the morphology of this subject, we would rather recommend the L size modality. It does not correspond to the clothes size indicated by this person, but seems to suit well person's height and BMI. We could not observe a clear trend in the error changes, but we could repetitively confirm that overall (with several exceptions) the ANN method performs the best for this subject.

2nd subject: The analysis of this subject's data easily concludes with the L size modality as the optimal one. All parameters except the QRS axis differences have a tendency to perform better when modalities change from S to L. This subject indicates L as his clothes size and his height does not let to doubt that L should be the best sensor-system modality from all three available modalities. All parameters here as well are demonstrating smaller errors when they are obtained from ECGs on which the ANN reconstruction method was used.

5th subject: According to this subject's data, most probably the L size modality shall be used. All cross-signal descriptors support that choice as well as the analysis of the subject's morphology. The differences of the interpretation parameters and of the QRS peak to peak amplitudes are pretty stable for all modalities and only the QRS axis differences obtained from the ECGs with the ANN reconstruction method indicate S as the best size. We could observe, however a descending trend for almost all other descriptors, where the L size modality ensures better results than the S or M size modality . And the ANN reconstruction method performs better than the linear regression method on this subject's data as well.

	Females				Males			
Subject	1	6	3	4	8	7	2	5
Clothes size	*S*	*M*	*M/L*	*L*	*S/M*	*M*	*M*	*L*
Modality	*S*	-	*any*	*S/L*	*L/M*	*L*	*L*	*L*
Method	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*

Table I.4.5 C2: optimal sensor-system modality and data synthesis method suggestion for each subject

Taking into account the results from all subjects (Table I.4.5), we see that, when comparing with the C1 study, we could recommend an optimal sensor-system modality slightly easier or sometimes with a greater conviction, though different in certitudes still exit. Obviously there is no one unique size for all subjects as the descriptors' values and their tendencies are specific to each subject. This study only confirmed that the ANN method this time unambiguously reconstructs ECGs better than the regression method. Here as well, we have not noticed a specific difference between the female and the male groups.

Comparison study C3

The comparison study named C3 differs from the C2 study just in the standard 12-lead ECG used to compare with the synthesised ECGs derived from the S, M and L ECGs (Fig.I.4_9). This time it is the second standard ECG (S2-ECG) obtained in the recording series of 6 ECGs per subject which is used as reference.

1st subject: The examination of the results obtained out of this subject's data does not provide a clear answer upon which modality is the optimal one for her. The interpretation descriptor together with the differences of the QRS peak to peak amplitudes and the QRS axes (the latter obtained with the linear regression method) indicate the modality of the size M as the best, while the cross-signal parameters are rather neutral or indicate the size L as the best. The same happens with the differences of the QRS axis descriptor when the ANN method is used to calculate it. These results do not really agree with the subject's morphological parameters, which are putting priority on the S size. For this subject, this time we could not really find a tendency and so could not distinguish any particular modality of the

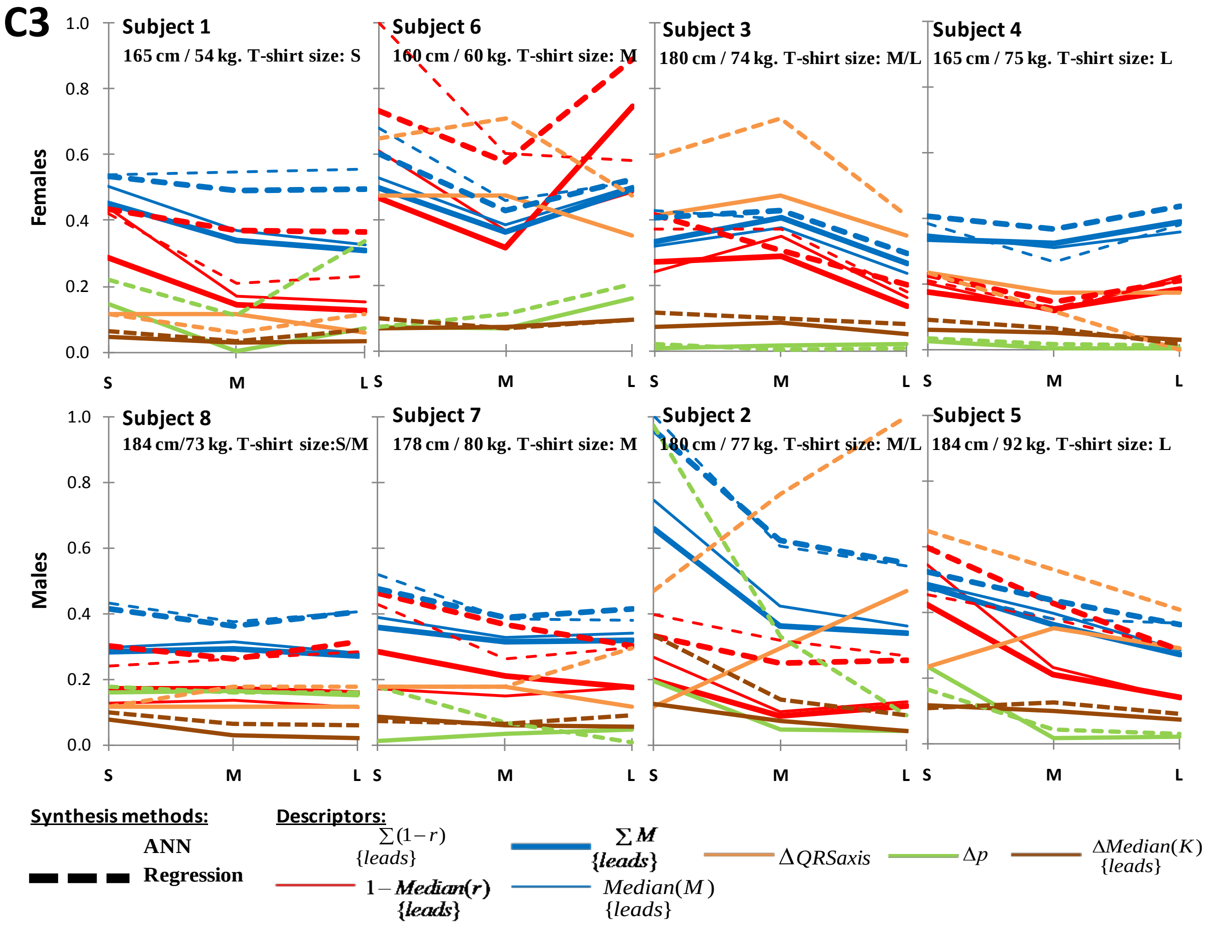


Fig.I.4_9 Normalised comparison results between the second standard 12-lead ECG (S2-ECG) and the three 12-lead ECGs derived from the 3 simplified ECGs for 8 healthy volunteers: 4 women and 4 men sensor-system as being the optimal. However, except for the small exception of the M size modality QRS axis difference data, the ANN method clearly reconstructs ECGs better than the linear regression method.

6th subject: A very similar situation happens when analysing the comparison results of the 6th subject. Even though the cross-signal parameters in unison are the best for the sensor-system corresponding to the M size modality, the QRS axis differences are smallest for the L size modality and the differences of interpretation parameters are smallest for the S size modality. We would probably, however, recommend the modality size M for that subject as the morphological parameters suit this size. Also, the differences of the QRS peak to peak amplitudes are pretty smooth throughout all modalities. We could not observe any tendency in the error that could be related to the expansion of the modalities size, though we could easily confirm that the ANN method is more accurate than the linear regression method.

3rd subject: The analysis of this subject's data is quite simple – all descriptors perform the best at the size L. Thus the L size modality would be recommended as optimal for that subject. The selection corresponds, also, to the morphology of this subject. We could not observe any error tendency that would have a linear relationship with the sensor-system modalities, however here again we could confirm that the performance of the ANN method is undoubtedly better than the one of the linear regression method.

4th subject: The data of the 4th subject are pretty smooth for all modalities. However, the parameters of the cross-signal group perform slightly better with the M size modality. The differences of interpretation parameters and of standard ECG measurements are slightly better at the modality of size L. As the subject's morphological data are slightly controversial - the clothes size is indicated as L, though the person's height would probably suggest the M size _, we would recommend the M and L sizes for this person. We could not really observe some augmenting or descending tendency of the comparison error with respect to the sensor-system modalities of different sizes, though we can confirm that the ANN method is more accurate than the linear regression one.

8th subject and 7th subject: The 8th subject together with the 7th subject share pretty similar tendencies on their data. The comparison results are pretty smooth for all sizes and even if there are some fluctuations, different parameters compensate each other. For these both subjects, despite their morphological differences, we could not identify an optimal sensor system. There is no error trend, where the error raises or gets smaller. However, we would clearly confirm that the ANN method is more powerful and performs better for the data of both subjects.

2nd subject: The 2nd subject's data are tendentious and give a smaller error when the L size modality is used. The subject's morphology corresponds to the choice L, so that would be the optimal modality we could recommend for this subject. Only the QRS axis descriptor is indicating completely contrary results, though we would base our opinion on the majority of parameters. Indeed, we could observe, also, an error tred in the data of this subject. The descriptors perform worse when the data were obtained with the S size modality and best when the modality of L size was used. These subject's data underline a clear tendency that the ANN method is more accurate for all parameters and for all modalities of all sizes.

5th subject: The optimal modality for this subject is the one of the size L. This size is also confirmed by the subject's morphology. We could clearly observe a tendency of the descriptors' values to decrease in function of the sizes and we of course can state that the ANN method is the one that should be used on this subject's data.

	Females				Males			
Subject	1	6	3	4	8	7	2	5
Clothes size	*S*	*M*	*M/L*	*L*	*S/M*	*M*	*M*	*L*
Modality	-	*M*	*L*	*M/L*	-	-	*L*	*L*
Method	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*	*ANN*

Table I.4.6 C3: optimal sensor-system modality and data synthesis method suggestion for each subject

The global overview of the results (Table I.4.6) from the C3 study pretty much confirms what we have already noticed in the comparison studies C1 and C2. We could define an optimal sensor-system modality for almost all subjects, though the confidence we can have in this choice varies and sometimes even implies the selection of two modalities for one subject. We do not think that there could be a unique size for every subject. We usually observed a linearly increasing or decreasing error in function of the sensor-system size expansion, but we could not state that for all subjects. Whereas, we could unanimously vote for the ANN method, which should be used for the data reconstruction as almost all parameters demonstrated the smallest errors for almost all sizes when this synthesis method was used.

4.8. Paired t-test results obtained comparing the C1, C2 and C3 studies

The results obtained from the C1, C2 and C3 comparison studies after a careful examination permit to project an optimal sensor-system modality for almost all subjects. However, the analysis of all data doesn't seem to result with the same optimal modality for each subject in all comparison studies. Table I.4.7 displays the results of these differences expressed by the p value of the paired t-tests between all descriptors of the comparison studies C1, C2 and C3. Table I.4.8 demonstrates these differences expressed separately via 3 types of parameters. Two descriptors: the sum of the correlation coefficients and the median of the correlation coefficients are presented under the name of correlation. The same grouping is made for the RMS values and the third parameter represents the differences of interpretation for Myocardial Infarction.

	C1/C2	C1/C3	C2/C3
1st subject (f)	***0.87***	***0.71***	***0.65***
6th subject (f)	*0.02*	*0.00*	*0.00*
3rd subject (f)	***0.23***	*0.03*	*0.01*
4th subject (f)	***0.53***	*0.00*	*0.00*
8th subject (m)	*0.00*	*0.00*	***0.60***
7th subject (m)	***0.08***	*0.00*	*0.00*
2nd subject (m)	***0.88***	***0.37***	*0.03*
5th subject (m)	*0.00*	*0.00*	***0.55***

Table I.4.7 Paired t-tests results taking into account all descriptors, f – female, m – male

	Correlation			RMS			Δp		
	C1/C2	C1/C3	C2/C3	C1/C2	C1/C3	C2/C3	C1/C2	C1/C3	C2/C3
1st subject (f.)	***0.19***	***0.74***	***0.71***	***0.75***	***0.91***	***0.89***	*0.02*	*0.02*	***0.10***
6th subject (f.)	*0.00*	*0.00*	*0.03*	*0.00*	***0.56***	*0.00*	***0.51***	***0.15***	***0.11***
3rd subject (f.)	***0.65***	***0.95***	***0.73***	*0.00*	***0.06***	*0.00*	*0.00*	***0.06***	*0.00*
4th subject (f.)	*0.00*	*0.01*	*0.00*	***0.85***	*0.01*	*0.01*	***0.06***	***0.11***	***0.14***
8th subject (m.)	***0.08***	*0.01*	*0.00*	*0.00*	*0.00*	*0.01*	*0.00*	*0.00*	*0.00*
7th subject (m.)	***0.56***	*0.04*	***0.13***	*0.01*	*0.00*	***0.14***	*0.01*	*0.03*	***0.14***
2nd subject (m.)	*0.00*	*0.00*	***0.44***	***0.21***	***0.23***	***0.48***	*0.01*	***0.10***	*0.00*
5th subject (m.)	*0.00*	*0.00*	***0.60***	*0.00*	*0.00*	*0.01*	***0.51***	***0.13***	***0.47***

Table I.4.8 Paired t-test results for 3 descriptors, f – female, m – male

The results show a data heterogeneity for most of the subjects. Summarizing the whole, we could indicate that the results highly depend on the subject himself. For example, the studies related to the 1st subject are significantly different, while the comparison results for the 8th subject are more or less similar.

4.9. Repeatability study results

The results heterogeneity is also confirmed by the repeatability test. The experiment on the data repeatability included 3 female subjects and 3 sets of data recorded on each of them at about 2 and 3 months intervals. The same data acquisition procedure of 6 ECGs per recording session per subject was used. The first data set from all three subjects was obtained in October, 2008. Then, after two months, we repeated the recording session and got the data for the second set. Finally, the third data set was obtained at the beginning of January 2009, around 20 days after the second recording session. We processed and analysed the acquired data on the base of 3 comparison studies: C1, C2 and C3 (annex 4.3). As an example, we present hereafter the data analysis results of the C3 study in detail.

1st subject: The examination of the 1st subject's data throughout all data sets (Fig.I.4_10) suggests that recording an ECG after some time, even if the recording procedure remains the same, does affect the electrical ECG parameters. The obtained descriptors' values are hardly comparable among the sets. The sketchy comparison seeking to find out an optimal sensor-system modality is still possible, but strongly lacks of certitude. For example, the 1st data set would more-or-less suggest the modality of L size if we take in account the descriptors obtained from ECGs synthesised with the ANN method and would suggest the modality of M size, if we would check values obtained using the linear regression method. The 2nd and 3rd data sets are balancing between the S size and the L size modality. Descriptors obtained with the ANN method would stand more for the sensor-system of the size S, while values obtained with the linear regression method performs better with the L size modality. The value of the interpretation descriptor changes in each set and performs better with size M (1st data set) and size L (2nd and 3rd data sets) modality. Consequently, we cannot fix an optimal sensor-system size for this subject that could be used over the time. However, with only a couple of exceptions the ANN data reconstruction method seems to perform much better than the linear regression method throughout all data sets.

1st subject: **2008.10.20 (1st set)** **2008.12.19 (2nd set)** **2009.01.09 (3rd set)**

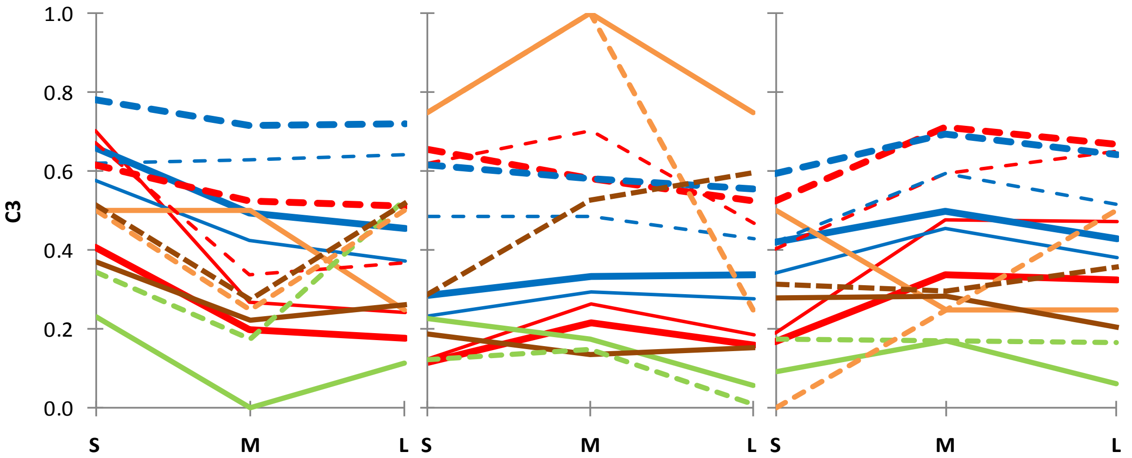


Fig.I.4_10 Normalised comparison results between a second standard 12-lead ECG (S2-ECG) and 3 sets of synthesised 12-lead ECGs reconstructed from 3 simplified ECGs recorded at a few weeks interval for the 1st subject

3rd subject: **2008.10.20 (1st set)** **2008.12.19 (2nd set)** **2009.01.09 (3rd set)**

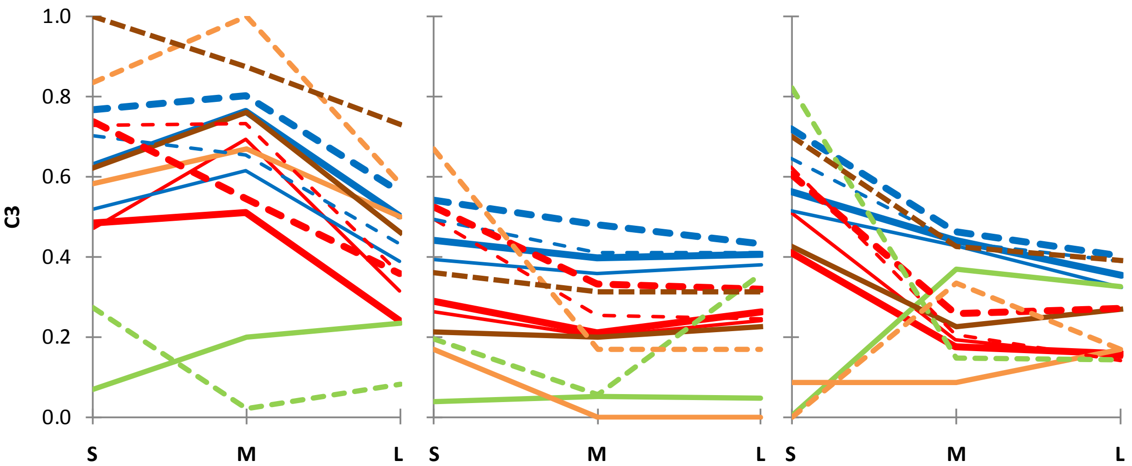


Fig.I.4_11 Normalised comparison results between a second standard 12-lead ECG (S2-ECG) and 3 sets of synthesised 12-lead ECGs reconstructed from 3 simplified ECGs recorded at a few weeks interval for the 3rd subject

4th subject: **2008.10.20 (1st set)** **2008.12.19 (2nd set)** **2009.01.09 (3rd set)**

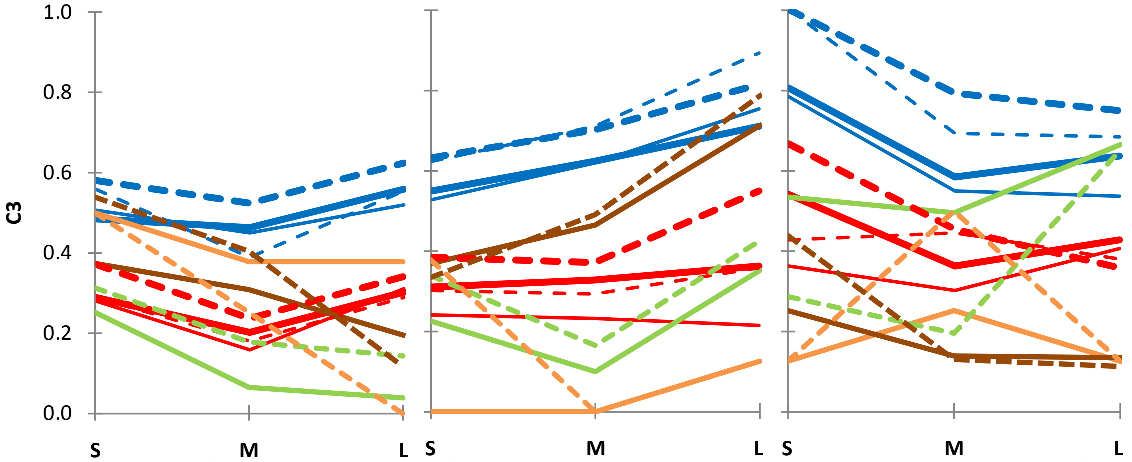


Fig.I.4_12 Normalised comparison results between a second standard 12-lead ECG (S2-ECG) and 3 sets of synthesised 12-lead ECGs reconstructed from 3 simplified ECGs recorded at a few weeks interval for the 4th subject

3rd subject: The examination of the results of the 3rd subject's data (Fig.I.4_11) results with completely different conclusions than the one performed on the 1st subject's data. The data sets of the 3rd subject do not mandatory resemble each other much, but we could clearly observe a tendency of a descending error towards the modality of size L. Only the interpretation descriptor fluctuates pretty differently in each data set and suggests a S, M or L modality size depending on the data reconstruction algorithm and the data set itself. It might be too early to judge that the L size is the optimal sensor-system for that subject and that it can be used over the time, but the primary results suggest exactly this conclusion. Overall, we could conclude again that the non-linear data reconstruction method – the ANN method in our case, performs better than the linear regression method.

4th subject: The results of the 4th subject data (Fig.I.4_12) are miscellaneous in almost the same manner as the 1st subject's data are. The majority of the descriptors of the first two data sets perform better either with the S or the M size modality, while the descriptors of the 3rd set seem to show smaller errors with the M and the L size modalities. Differently from the 1st subject's data, here we could observe some tendencies of a smaller error, for example towards the size S in the first two data sets. However, as the data sets are not homogenous, it's impossible to detect the optimal size which this person could use over the time.

To check our observations including all three studies (C1, C2 and C3) we performed a paired t-test (Table I.4.9). The results of this test on the descriptors' values of the reconstructed ECGs from simplified ECGs recorded at about 2 and 3 months interval show significant differences ($p<0.05$) within each ECG comparisons' type Ci, i=1,2,3, for at least one and at most two subjects, but not specifically to one subject or to one comparisons' type.

	1 set vs. 2 set			1 set vs. 3 set			2 set vs. 3 set		
	C1	C2	C3	C1	C2	C3	C1	C2	C3
1st subject	*0.00*	***0.13***	***0.37***	***0.34***	*0.00*	***0.06***	*0.00*	***0.24***	***0.63***
3rd subject	*0.00*	*0.00*	*0.00*	*0.00*	*0.02*	*0.00*	*0.00*	***0.31***	***0.09***
4th subject	*0.06*	***0.84***	*0.00*	***0.33***	***0.07***	*0.00*	***0.23***	*0.02*	***0.29***

Table I.4.9 Paired t-test results for the 3 sets of data per subject

4.10. Discussion

A total of 280 ECG comparisons were performed on the basis of 20 comparisons per subject (6 « involving 2 synthesis methods for the 3 S, M, L modalities » for each of the 3 types C1, C2 and C3, plus the S1 vs. S2 and S1 vs. ML ECG comparisons) and for a total of 14 "subjects" (8 subjects during the 1st experiment, and 3 subjects recorded twice for the repeatability study). The small number of subjects limited our study, although it already clearly demonstrates the complexity and the variety of obtained results. The observation of these results shows that there is no clear tendency of a greater inter-signals difference when the used modality is shifted from a smaller towards a larger size (further from the heart). Also, there is no linear values progression according to the modalities shifts per sizes. Yet, we notice that the descriptors' values are specific to each subject and form a subject specific descriptors' profile. Globally, we see that the ANN method seems to reconstruct ECGs better than the regression method, as it has been already reported in the literature *[Atoui'10]*. By looking at the descriptors' values for some subjects we could define the most suitable sensor-system modality. Our study shows that the complexity of the human's body surface with respect to a diagnostic quality ECG recording cannot be only explained by the gender and/or by the clothes size.

This study implies the possibility that two types of subjects could exist and so should be considered: (I) those for whom personal settings are needed in order to take in account their morphology for proposing them an optimal sensor-system modality, the most advantageous available data reconstruction, analysis, and decision-making algorithms, and (II) those for whom personal settings are not required. The latter could, actually, be an issue of two situations: (IIa) any settings adjustment would provide a good quality ECG, and (IIb) none of the possible adjustments can increase the quality of the obtained ECG. Some generic settings shall thus be made available to the subjects of the second group. These suppositions came out of the fact that the recordings obtained over a pretty short period of time for some subjects differ significantly, even if the same electrode placement and same processing and analysis algorithms were used. While, for other subjects, moderate changes of the sensor-system or of the analysis algorithms have not caused any significant fluctuations on the obtained data even over the time. Our small study suggests that 2/3 of the subjects are of the type (I) and remaining 1/3 of the subjects are of the type (II).

Herein, before "playing" with home-based sensor-systems, we would see an interest of firstly collecting observations on the influence of the human's morphology on a series of standard 12-lead ECGs obtained on one subject over the time and analysed with various methods. The results of this new study could stand as a fundamental background for analysis and evaluation of self-care systems. Herewith, we should also underline a lack of any studies of this type, what in consequence results in a lack of established data analysis and data assessing methods for self-care systems. This also calls for new databases of carefully selected ECGs that would go with appropriate meta-data describing well the morphology of the subjects and the subtlety of the recording, processing and analysis methods. Finally, new standards might be previewed to promote the quality of self-care services.

Indeed, medical self-care devices and services should not be seen in isolation but as a part of a systematic approach to care delivery. High level attention should be paid to the understanding of the citizens' heterogeneity in terms of needs, environment, and circumstances. Our study *[Krupaviciute'10b]* demonstrates that different settings related to an ECG recording, processing and analysis make influence on the quality of an ECG and so may alter the final diagnosis. These variations need to be formalised and their assessment and control methods need to be established. We firmly believe that due the lack of such type of studies and so the formal knowledge these studies could contribute with, the specialist-independent home-based systems will have difficulties in reaching the market and thus to fulfil the promises of a better health-care. We expect our study to be a starting point for future research and thus to be criticised, compared with, expanded and extended.

4.11. Conclusion

The most important outcome of our study is that the sensor-system size does not correspond to the person's clothes size. Also, with respect to our results, we can see that one size for everyone can introduce important signal distortions and that a subject specific sensor-system personal to each citizen shall be used to provide high diagnostic quality signals.

All three comparison studies: C1, C2 and C3 almost unambiguously demonstrate that, in general, the ANN synthesis method reconstructs ECGs that are closer to their reference ECGs in comparison with those reconstructed using the linear regression method. However, some specific cases perform better with the linear regression method and, in the context of this thesis' aim, these cases should be compulsory taken into account. Indeed, our goal goes further than searching general tendencies, which might be new and interesting as well, till finding the best match of sensor-system modality and of corresponding algorithms to provide a diagnostic quality ECG for a specific person. We firmly believe that the personal adjustment can be and should be performed automatically on the basis of the ECG signal information and not of the chest/clothes size of the individuals.

We highlight the importance of using different signals recordings, synthesis methods, transformations' strategies and descriptors typologies for the personal modality selection. As a matter of fact, this approach produces a lot of data and thus makes more complex their processing. Though here, we suggest appealing to the Central Limit Theorem, which supports the plan of having more data in order to approach the truth. Then, the more methods on the more signals are used and the more descriptors are evaluated, the more precise selection can be performed. The integration of various services which perform the needed computations could be one of the possible solutions for mastering the complexity. This integration could be based on a dynamic and services oriented architecture, which adapts to the context of each specific user. The results of this research will be presented in Part II of our thesis, which aims to contribute to the domain of context-adaptive business processes and so could be adaptable to solve personalisation issues in various domains, including the self-care in cardiology.

In conclusion, we have proposed a three sizes sensor-system that can be used in self-care and that was evaluated on 8 healthy volunteers. The results obtained so far confirm that there is a need to carefully select the "best" sensor-system modality for each subject in order to ensure the diagnostic quality of his vital signs recording. The results show that the evaluation process shall be automated by involving a large typology of ECG signal descriptors. A decision-making strategy, however, must be tested on more clinical data.

Part II

Ambient intelligent assistance

Chapter 5 – State of the art

5.1. Introduction 93
5.2. Business process modelling and management 94
5.2.1. Business process modelling and implementation via services 95
5.2.2. Automated services composition 97
5.2.3. Business process management 100
5.3. Architectural foundation for services composition 101
5.3.1. Service-Oriented Architecture 101
5.3.2. Model-Driven Architecture 104
5.3.3. Event-Driven Architecture 104
5.3.4. Discussion 105
5.4. Knowledge engineering & Knowledge management 106
5.4.1. Knowledge engineering 106
5.4.2. Knowledge management 107
5.4.3. Process-centred knowledge 109
5.5. Context-awareness versus Situation-awareness 110
5.6. Decision-making and support 113
5.6.1. Decision support 113
5.6.2. Decision-making 114
5.6.3. Intelligent assistant for decision-making and decision support 115
5.7. Summary 116

5.1. Introduction

Adam Cheyer, cofounder of a Silicon Valley start-up named “Siri” – an application capable of working as virtual personal assistant software for iPhone OS, in 2009 presented an "intelligent software assistant" as one of 10 emerging technologies *[Naone'09]*. The initial version of Siri, released in the third quarter of 2009, is aimed at users of mobile phones and is capable, for the moment, to perform only specific types of functions, such as helping with reservations, checking flight status, or planning weekend activities. The application uses natural language processing to answer questions, make recommendations, and perform actions by delegating requests to an expanding set of web services. The importance of Siri’s release stands in the fact that it is one of the first applications shifting the interaction in the internet from the search paradigm to an assistance paradigm, that is, to a new type of software, which is focused to help users to complete tasks rather than just to collect the information or to retrieve it from the net. Indeed, when the task is to find information and the problem can be solved by reaching a web page, the search engine paradigm is optimal. Though, when the task is to solve a problem involving personal context, different preferences, choice-making, and applying multiple information sources to a task, the assistant paradigm is better suited. Siri’s team forecasts that in a five to ten year period, having our own virtual assistant, which with time will get more personal and more intelligent, will be as natural as it is to have a shadow.

By definition, an intelligent software assistant is a software agent, which is capable to perform a task with minimum specific directions from a user and may also be known as “cognitive assistant” or “cognitive agent” *[Rasmussen'94]*. It combines an amount of different technologies, such as a dialog, natural language understanding, signal processing, learning, planning, reasoning, service delegation and integrates them all into a single assistant, a human kind assistant that can help a user to get things done according to his concrete situation and needs. The purpose of developing such an intelligent software assistant is to formalise specific domain knowledge and to accumulate experience of human-machine interaction that will help people to be more efficient in their lives, in solving problems, performing tasks and making decisions. *“If applications were able to detect a user’s expertise, then software could automatically adapt to better match expertise. Detecting expertise is difficult because a user’s skill changes as the user interacts with an application and differs across applications. This means that expertise must be sensed dynamically, continuously, and unobtrusively so as not to burden the user.” [Hurst'07]*.

This all becomes more realistic with the growing power of pervasive and ubiquitous computing during the last decade, which significantly increased the information processing integration into everyday objects that are located anywhere. These enhanced objects are now easily accessed anytime by the user or can record contextual user information autonomously. Contextual real time data together with boosted networks let to anticipate a possibility to design a new type of powerful software that provides a personal aid. This new intelligent layer is aimed to complement the sensor data with domain knowledge of a specific assistance scenario (e-reservation, e-planning, e-learning, e-health) and to allow delivering a wider range of services for accomplishing our tasks, mainly enhancing the autonomy of users and improving the ability of information systems to successfully assist the user.

For example, the need of an intelligent software assistant is very well pronounced in the healthcare domain. The great expectation in this domain is to improve a care model, which encourages citizens to develop new self-care habits by actively following their health status on their own in order to prevent diseases. The primary healthcare provision, then, is under-taken by the individual rather than in centralised healthcare units and *"far from being a "passive subject", the patient becomes one of the main actors and the main responsible for his/her own daily management and possible behaviour changes (self-determination). This leads to a sort of "assisted self-management" way of care delivery that could be both efficient and effective if the organization succeeds." [Leonardi'07]*. The functionality of such assistance shall allow the *"dynamic content provision, adaptation to each patient's profile, and mechanisms for data validity"*, as well as it implies that *"a system has to be designed in a flexible and extendable manner, enabling "plugging" of different interfaces serving a common purpose" [Maglaveras'05]*. Computerised devices providing this type of assistance will be a part of a pervasive environment and will ask for *"a distributed and modular architecture"* capable to handle *"devices heterogeneity and their dynamicity, concerning their availability as well as the nature of data and services they can provide" [Gripay'08]*. The pervasive environment is mostly characterised for its resources heterogeneity. At the moment, the management of these resources is based on their typification, that is on their determination in advance. Though, surely with the asked functionality the system shall take in account all resources, no matter have we identified them in advance or not, which demands for a continuous evolution of information models in order to be capable to adapt to changing clinical requirements, as *"the list of medical concepts can never be complete" [Martinez-Costa'09]*.

We shall emphasise that all fields of applications, like the above mentioned e-reservation, e-planning, e-learning, e-health, share similar problems, such as expert activities modelling, automated composition/orchestration of distributed services, dynamical system reconfiguration, personalisation, which are currently dealt separately in the distinct scientific communities. For the hereafter review purpose, we will group the current research works into five classes related to: business process modelling, service-oriented architecture, knowledge engineering, context-awareness and decision-making. In each of them we target to underline the current research results that could influence the development of intelligent software assistant. Each of them separately lack of theoretical and methodological power to propose a final solution of a flexible architecture capable to deploy a dynamic business process, which instantiation is tailored to the specific user context. Though, the ability to grasp them in the context of each other should permit to build a base for a methodology of an ambient intelligent software development, which we are targeting at in this thesis.

5.2. Business process modelling and management

In order to achieve a successful communication between the intelligent assistant and the user as well as automatic and correct communication among the different modules in the intelligent assistant itself during all communication periods, it is necessary to model a working scenario of such an application. The most related modelling practice is a wide spread routine of business activities modelling as business processes. Business activities started to be modelled seeking to automate business and to expand its accessibility by overcoming time and distance limitations as well as reducing business costs. All this is absolutely coherent in the intelligent assistance case, therefore the intelligent assistant can be seen as a business activity and so its working scenario as a business process. Thinking in processes rather than in procedures or functions was widely accepted by system designers/architects in the early 1990's and so in the field of software engineering there is now common to use the term of business process modelling (BPM). *"Business process modeling is a set of practices* *or tasks that companies can perform to visually depict or describe all the aspects of a business process, including its flow, control and decision points, triggers and conditions for activity execution, the context in which an activity runs, and associated resources." [Bloomberg'06]*.

5.2.1. Business process modelling and implementation via services

The business activities' modelling is generally performed for activities based on well known, already defined procedures. These procedures are documented with one of the modelling languages, such as flow chart, Gantt chart, PERT diagram, the family of IDEF (Integration DEFinition) languages or using BPMN (Business Process Modelling Notation). The latter was introduced by Business Process Management Initiative and from 2005 maintained by OMG (Object Management Group). The latest version of BPMN is version 2.0 *[OMG'10]*. The aim of such a documented business process is to better analyse the business activity, to improve process efficiency and quality, and, finally, to automate as more steps in the process as possible. All these languages usually suggest *"the choice of shapes and icons used for the graphical elements"* that *"defines many semantic concepts used in defining processes, and associates them with graphical elements, markers, and connections" [OMG'10]*. Michael Havey in the book "Essential business process modeling" *[Havey'05]* summarises the most important business process modelling terms that are as follows: process definition, process instance, activity, manual activity and automated activity. I'll quickly define them here again, as they will be often used in following chapters. A process definition is *"a basic algorithm or behaviour of the process"*, that is a chain of steps transforming given input data to the required output. A process instance is *"an occurrence of the process for specific input"*. Then an activity or a task is *"a step in the process"* representing some part of work that is logically separated from other parts, for example a task performed by one specific role or one operation. An activity in the process can be manual or automated. A manual activity is *"a step in the process that is meant to be performed by a human process participant"*, while an automated activity is *"a step in the process that is performed directly by the execution engine"*. Additionally, a process execution direction – an order of activities, shall be determined using a so called sequence flow. The execution can be set step by step, in parallel, or with a condition that some defined clauses are fulfilled. Thus, a business process is a set of activities determined in an ordered sequence to reach a fixed purpose in some domain.

The literature on business process modelling is large and diverse, however all approaches can be separated into two categories: *"the one created before web-services – workflows and the other based on web-services" [Cauvet'08]*. The term workflow has first been introduced to describe a controlled business sequence where people are participants. It facilitates simple routing of tasks and activities from person to person. The systems that offer workflow functionality always revolve around users/groups/queues, their associated forms and various user assignment schemes. Automated activities (steps) are usually provided as well, but only as a secondary offering. A workflow approach describes *"how a certain result can be reached" [Josuttis'07]*, it looks further into the details of all steps or activities and is often criticised *"for its rigid process specification" [Cauvet'08]*. At present, it can be simply understood as a concrete implementation of some business process, a process instance, because currently the business process modelling domain is undergoing the dramatic change from static workflows to dynamic business processes. This change was initiated by enterprises, which searched for the solution to customise processes according to changing customers' requirements, and led to rethink business process modelling. Enterprises started to ask for business processes that could better respond to the needs of the customer by adapting themselves "on demand". Therefore now a "*flexibility is becoming a new guiding principle in the design of business processes" [Cauvet'08]*. The terms of mobility and flexibility refer to the way in which processes evolve as they execute through the exchange of information among participants whose relationships evolve as a result. This newly spreading BMP approach is matching well the demands of the intelligent assistant, which, indeed, is not based on some defined practice known in advance. The intelligent assistant acts more likely as an expert which communicates with the user, obtains required data and according to them provides a personalised output. The direction, upon which many researchers agree by now, is that this type of flexibility could be reached by *"reusing business services that encapsulate business process fragments" [Cauvet'08]*, where *"the lowest activities of a decomposed process are services" [Josuttis'07]*, hoping that dynamic composition of services will generate tailored processes.

By definition, a service is a piece of self-contained business functionality. *"Whether a service is implemented as a fine-grained component performing a discrete operation or as an application suite exposing an entire business function, each can be considered a self-contained, self-describing, modular unit that participates in a larger ecosystem." [Chappell'02]*. Services can be atomic or non-atomic (compound/composed). Atomic services commonly implement basic functionalities that are not further broken down to a finer lever and usually correspond to tasks (atomic activities) in a business process model. A non-atomic service accepts and encapsulates other services: *"Eventually these small ecosystems can all be combined into a larger, more complicated, orchestrated business macrocosm" [Chappell'02]*. Composite services are designed as coarse-grained services structures that include more business functionalities and implement sub-processes (compound activities) of business processes. From the implementation point of view, a sub-process is equivalent to a business process itself – a set of atomic services guided to be executed as a defined process and as a whole are implemented as a composite service. Each service is made from two parts: service (the implementation) and service description (the interface). Nowadays we often add "web" to the service definition in order to underline that this service is *"located somewhere on the Internet and is accessible through standard-based Internet protocols such as HTTP or SMTP" [Chappell'02]*. Yet, the web-service has the same functionality as a service, simply, a web-service has a sort of envelope, which makes web-service searchable and invokable in the net. A web service is additionally specified as having specific behavioural characteristics: XML-based, loosely coupled, course-grained, has an ability to be synchronous or asynchronous, supports RPC (Remote Procedure Calls) and document exchange. The three major web-services technologies are: (i) SOAP (Simple Object Access Protocol), which provides a standard packaging structure for transporting XML documents, (ii) WSDL (Web Service Description Language) – an XML technology that describes the interface of a web service in a standardised way, and (iii) UDDI (Universal Description, Discovery and Integration), which provides a worldwide registry of web services for advertisements, discovery and integration purposes. Web-services *"represent a new model in the utilization of the network in which self-contained, modular applications can be described, published, located and dynamically invoked, in a programming language independent way" [Berardi'04]*.

Frequently, several different and often remote web-services can implement the same activity. Then a composition mechanism responsible for finding the "best" matching service and composing it into the predefined business process has to be set. Also, *“when a functionality that cannot be realized by the existing services is required, the existing services can be combined together to fulfill the request” [Sirin'03]*. What we have just called a predefined business process in the literature is often named a process workflow document/plan or simply a workflow. In general, services composition consists of two parts: *“the static part that involves the definition of services, including the services operations and their interfaces, and the dynamic part that defines the associated process workflow” [Feuerlicht'09]*. Naturally, it is expected that services composition should be performed automatically *[Fahkri'11]*, that is after the request of composing such a workflow is set, *“a composer agent will act independently, without any end-user intervention” [Laukkanen'04]*. This automated composition – a base of intelligent assistant – has been actively researched in both semantic web and workflow communities.

5.2.2. Automated services composition

Semantic web research for the automated services composition focuses on the part of services description. It stresses the need at first to semantically annotate services and then to compose them into workflows. The aim of this annotation is to provide a common framework that allows data to be shared and reused across applications, enterprises or over the Internet. However, the task of annotation is complex *“due to the heterogeneity of the web service structure, non standardization of registry format and complexity in capturing the semantic part of the web service functionalities” [Mustapha'06]*. The design of message structures determines the compatibility of service interfaces, and consequently the composability of services into higher level business functions. There is a number of different semantic web projects focussing on various aspects of web-service description and composability. Commonly this now is based on ontology languages as they are widely accepted as being the most appropriate paradigm in the semantic web domain. *“In an ontology the knowledge of a certain domain is represented formally as a set of objects and the relationship among them.” [Laukkanen'04]*. Ontologies consist of hierarchically modelled domain concepts, which are clearly determined through properties, like that supporting knowledge representation and reasoning and so consequently they can precisely define shared understanding of services annotations. The semantic web research society uses ontologies to provide a specific domain related vocabulary, which allows describing web-services properties and capabilities. In 2001, a partnership of semantic web researchers, under the program of DARPA Agent Markup Language (DAML), agreed to develop an ontology for web-services, using the semantic web ontology language DAML+OIL (OIL stands for Ontology Inference Layer). This has resulted in the creation of an Ontology Web Language for Services (OWL-S, formerly DAML-S), which can be used to describe services with an unambiguous description in order to discover them and to compose them into a workflow. OWL is an artificial intelligence and description logic based language for describing content on the web. Most importantly OWL, and thus OWL-S, has a well-defined semantics, which contrasts it with other efforts. As OWL-S was established, various projects started to build extensions to improve it or to adapt for different cases. For instance, the need for a computer-interpretable process model prompted researches to extend OWL-S with a First-order Logic Ontology for Web Services (FLOWS). FLOWS provides a well-understood theoretical model semantics and enables characterization of reasoning tasks for semantic web services in terms of classical notions of deduction and consistency *[Gruninger'05]*. Then, web service discovery can be characterised as deductive queries, and web service composition as satisfiability. The similar result was targeted by a suggestion to *"define QoS properties of web services: execution cost, execution time, availability, successful execution rate, reputation and frequency" [Ko'08]*. Another important problem was to enable the deployment and interoperability of semantic web services between different ontologies. *"Web services as passive participants within a composition often require mediation." [Barrett'06]*. This strength is a characteristic of the Web Services Modelling Ontology (WSMO). The WSMO framework ensures the connectors (mediators) between components with mediation facilities. There is also a demand for a framework which could provide richer semantic specifications of web services and to span the full range of service-related concepts. This work was supported by the W3C (World Wide WEB Consortium) organisation and developed towards the Semantic Web Services Framework (SWSF), which includes the Semantic Web Services Language (SWSL) and Semantic Web Services Ontology (SWSO). This framework targets to enable fuller, more flexible automation of service provision and use, to support the construction of more powerful tools and methodologies, and to promote the use of semantically well-founded reasoning about services. It aims to permit the specification of many different aspects of services and to provide a foundation for a broad range of activities across the web service lifecycle. All these examples illustrate how greatly service annotation problem solving benefits from various research projects, though there still remains unsolved issues. The part which burdens the whole job and which can be time-consuming is, actually, the construction of the domain ontology. This part was tried to be automated too, using for example techniques based on Artificial Intelligence (AI), though *"researches have shown that automated ontology building has not produced promising results" [Mustapha'06]*. In summary, important studies were accomplished by the semantic web community seeking to support/standardise web-services annotation in order to decrease efforts, time and costs in their development, their integration and the maintenance, thus the problem of their use in practice remains and waits for new solutions.

The workflow (plan) community researches are concentrating on different planning techniques that could be applied to the problem of automated composition of web services. What simply means that if we want to find a service and to compose it into the workflow to make some precise part of a whole work, we should know the totality of the tasks needed to finish the work, the sequence of these tasks and their compatibility with the functionality of the available services capable of implementing the task. In realistic situations the plan is already set in advance (prior to execution) by system designers or programmers, because *"the planning problem is far from trivial, and can be hardly addressed by "classical planning" techniques" [Traverso'04]*. Usually, several traditional workflow tasks (human, non-transactional, or transactional tasks) and web-service tasks are placed and interconnected on the canvas. Even though in the design level any business process model describing languages (mentioned above) can be used, in the implementation level another technology integrating and managing web-services needs to be designed in order to ensure *"the control and data flow from one service to another" [Barrett'06]*. This role is taken by Business Process Execution Language (BPEL), which mainly focuses on the executability and control of web services. The designed process at this stage is described, this time, in a low level (if-then) language for manual definition on how involved web-services interact to fulfil the designed business process. The BPEL is based on XML grammar and comes from a consortium consisting of BEA Systems, IBM, and Microsoft. It combines and replaces IBM's Web-Services Flow Language (WSFL) and Microsoft's XLANG specification. Today an international non-profit Organization for the Advancement of Structured Information Standards (OASIS) is in charge of the standardization of BPEL, which is now usually called Web-Services Business Process Execution Language (WS-BPEL) or Business Process Execution Language for Web-Services (BPEL4WS). The BPEL (WS-BPEL and BPEL4WS) is a process modelling and execution language based on conditions and iterations. The BPEL deals explicitly with the functional part of business processes: control flow (branch, loop, and parallel), faults and compensations, embedded subprocesses, asynchronous conversations. It directly addresses message exchange between parties, data manipulation between partner interactions, supports long running business transactions, parallel activities and provides a consistent exception handling. The BPEL, actually, enables to describe information exchanges internally or externally of the application, that is to define the rules for a business process flow in an XML document which can be afterwards given to a business process engine to "orchestrate" the process. The concept of web-services' orchestration describes an automatic coordination and management of complex web-services based systems. With orchestration, the process is always controlled from the perspective of one of the business parties, often called a central controller, which coordinates all web-services participating in the process execution according to steps and decisions set in advance. The BPEL can then be called a web-services orchestration standard, which defines the way of services aggregation to business processes.

Researchers separate the area of workflow composition in three cases: *"predefined workflow with pre-defined tasks, pre-defined workflow with on demand selection of tasks, and composition of the whole workflow on demand" [Laukkanen'04]*. The composition is, usually, performed by a so called workflow composer agent, which, actually, acts differently in each of these three cases. The first case describes the static situation where all tasks and their order in the workflow are set and all web-services implementing these tasks are known in advance, that is the whole composition is designed before the runtime of workflow. This is usually done using BPEL. In the second case we speak of the situation where tasks and their order have already been defined for the workflow, though the concrete web-service implementing the task is not known and shall be retrieved and composed into the workflow at runtime. This usually happens when e-workflow compositions is done: *"the composition of workflow processes that model e-service applications differs from the design of traditional workflows, in terms of the number of tasks (Web services) available to the composition process, in their heterogeneity, and in their autonomy" [Cardoso'03]*. For this case, techniques based on results of semantic web research are used. There web-services are described using ontologies (OWL, OWL-S), which enables the reasoning about them like that permitting the automation of web-services discovery and composition. Finally, the third case adds a second composition at runtime where not only the web-services should be selected and composed in a process at runtime, but also tasks themselves should be selected and composed in the final goal-oriented workflow at runtime. This composition type matches the best intelligent assistant's requirements for the personalisation and for the continuous context-awareness. Thus, the research on this composition type is fundamental in order to support the development of such intelligent assistant. Here, workflow modelling is viewed as a goal-driven problem. At first, Artificial Intelligence (AI) based algorithms were analysed as a most suitable candidate. *"A number of researches have advocated using AI planning techniques to address the task of Web service composition including using regression-based planners, planners based on model checking, highly optimized hierarchical task network planners and their combinations." [Sohrabi'09]*. Though, simply finding a plan that achieves a goal was insufficient as it is not the case that web-services invocations can be modelled as deterministic. Workflow composition needs to take into account not only the cost and time needed to accomplish a task, but also the factors included in service level agreements. A composition based on the exploitation of flexible templates as generic procedures was also proposed, though, currently AI based workflow planning was judged as not considering long term optimality while constructing the plans. Researchers recognise that execution monitoring shall enable the recovering *"from unexpected behavior of services due to service failures, and the dynamic nature of real-world environments" [Feuerlicht'09]. "The planner needs to deal with the nondeterministic behavior of web services, the partial observability of their internal status, and with complex goals expressing temporal conditions and preference requirements." [Traverso'04]*. Efforts to model non-determinism of web-services resulted in research on dynamic services orchestration *[Sakkopoulos'09], [Benatallah'03], [Fahy'08]*. This research attempts to answer the double composition problem described by the third workflow composition type: the selection of web-services that are not known a priori and the selection of a services' linking path, which is not known in advance as it depends on the usability context (i.e. on personal demands, intermediate data, services availability). One of the current approaches to this problem speaks of building a common language from a single information model, which can be used by systems *"to express their needs in a machine-programmable manner and translate those needs into a form that network can understand" [Fahy'08], [Robinson'08]*. The second direction works on technique which generates and checks (using mathematical algorithms) an automatic plan according to the limitations and conditions, so overall rules based. This technique focuses on retrieving and modelling the so called *domain control knowledge [Baier'08]*. One more area receiving a great attention is *web-services knowledge management* with an expressed need of *"an end-to-end management infrastructure - for applications, systems, and networks – that gives them flexible control over business processes involving Web services, including control of specific steps in the processes" [Peltz'03], [Adamopoulou'07]. "A robust management infrastructure must both monitor the environment and provide capabilities to adapt and optimize it in real time. The need for this type of management will become increasingly important as Web services are combined, aggregated and orchestrated to form more meaningful business processes." [Peltz'03]*. The creation of such an infrastructure strongly considers the experience gained into the business process management area. Here, a bit of confusion in modelling and managing business processes can be sensed, as for both the acronym is BPM. Though, we should see the business process modelling as a part of a business process management and so it's more common in the literature to use BPM when talking about business process management.

5.2.3. Business process management

The business process management can be perceived as the next step after the business process is modelled and put into action. *"Business process management usually refers to a technology-enabled means for companies to gain visibility and control over long-lived, multistep processes that span a wide range of systems and people in one or more organizations." [Bloomberg'06]*. The BPM integrates, automates, and optimises processes that were originally not connected *[Ciuksys'07], [Lee'08]*. It enables to gather information from various distributed systems and establish various key business performance indicators to more effective management and monitoring of business resources. For example, a specific measurement framework for software development processes was proposed by European and American research communities and resulted into CMM (Capability Maturity Model)

[Herbsleb'97] and ISO/IEC 15504 (Software process assessment) *[ISO'03]* standards. Though currently, general tools automating the BPM are not available. Usually, companies order software developers to code them a company-specific – compatible to their needs – BPM utility, which often requires to be complemented by new functions or tracking possibilities and so is continuously maintained. The business process management together with web-services knowledge management aims to include an additional notion of an intelligence dimension to the whole modelling, as the value of creating business processes is in the intellectual assets that those processes represent. The concept of enhanced business process, which is capable to dynamically adapt itself to the context in order to provide coherent results as well as services' capability to be dynamically reconfigured and/or built according to the incoming contextual data in order to implement that dynamic business process, remains a critical business problem. Ongoing research calls for new solutions to support this type of information handling. One of the most promising candidates could be based on an architectural framework oriented to services, that is, on the well known service-oriented architecture (SOA). Researchers in this domain raise the objective: *"How does service-oriented architecture has to be designed in order to allow for flexible interactions between learning and business applications realizing integrated scenarios." [Leyking'07]*, which we shall overview in the next section.

5.3. Architectural foundation for services composition

5.3.1. Service-Oriented Architecture

The Service Oriented Architecture (SOA) is already a practical standard for those oriented to a truly services-based business. The SOA expresses a concept of software architecture that defines the use of services to support the requirements of software users. Though, *"SOA is not a concrete architecture: it is something that leads to a concrete architecture" [Josuttis'07]*. Indeed, it is rather a set of ingredients that allows building a particular, usually enterprise architecture. The SOA may be intended for a use across the public Internet, or built strictly for a private use within a single business or among a finite set of established business partners. *"Service-oriented architecture (SOA) is a paradigm for the realization and maintenance of business processes that span large distributed systems. It is based on three major technical concepts: services, interoperability through an enterprise service bus, and loose coupling." [Josuttis'07]*. Here, the Enterprise Service Bus (ESB) is a middleware infrastructure based on messaging protocols aiming to support easy information exchange between remote parties. The abstraction of the "bus" is used to describe a messaging engine, where applications connect in order to interchange data so avoiding a number of point-to-point connections. The "bus" acts as a message broker ensuring the validation, the transformation and the routing of messages. It allows the connection of diverse applications so supporting the heterogeneity of the system. The capacity of such collaboration, usually by diverse applications with different interfaces, is described as a property of interoperability. The "bus" also reduces a system dependence on a concrete component as a replacing application can be easily integrated into the system. This improves system's flexibility and allows to decrease knowledge that one component has of others, which supports a loosely coupled system design. The interoperability and loose coupling are key factors of increased SOA popularity in business. Business always needs to adapt to customer demands, trends, the economy, and various regulations in order to succeed in a rough competition. The orientation of already existing applications towards services in order to facilitate these applications integration and reconfiguration promised this desired flexibility. The main cause why web-services based technologies were ready for success was their answer to data heterogeneity and the capability to reuse applications and methods. Though, this flexibility and so the benefits that go with it are not reached via one iteration, usually this happens progressively with the growing maturity of enterprises. That means that a higher level of SOA adaptation is reached gradually. Service-oriented integration (SOI) maturity models define this incremental transformation into several stages. Enterprises can use them as roadmaps for getting in position to deliver the business agility. One of these models is proposed by IBM (the earliest version goes back to 2003) and is called Service Integration Maturity Model (SIMM) *[Arsanjani'05]*. It is mapped to the Capability Maturity Model Integration (CMMI) *[Forrester'11]*, which is a natural way to understand maturity. Another model, called SOA Maturity Model (SOAMM) is based on CMM (Capability Maturity Model) and is used by Microsoft to assess the state of SOA *[Beack'06]*, *[Arsanjani'05]*. There exists also a Combined SOA Maturity Model (CSOAMM) composed comparing and mapping SIMM and SOAMM as they both are recently proposed as standards *[Jardim-Goncalves'07]*. If we take SIMM as an example, then by listing SIMM proposed seven steps of incremental enterprise architecture transformation towards more mature levels of service integration and by describing them briefly, we can see the capacities of SOA as follows:

First level: Applications architectures and topologies are monolithic and lack of integration with other systems across the enterprise. This level is called *Silo* and mostly puts attention on data integration.

Second level: Application architectures and topologies are monolithic with minimal separation of concerns between architectural layers of application tiers. This level is named *Integrated* and demonstrates examples of how applications are connected one with another.

Third level: SOA development practices are applied inconsistently across the organisation. Most application architecture topologies have both a physical and logical separation of concerns in presentation, business logic, and data tiers. The use of SOA enabling technologies such as an ESB is inconsistent across the enterprise. The level is called *Componised* and demonstrates the benefit of functional integration, as applications are seen or are divided into components responsible for some type of tasks.

Fourth level: Service components of application architectures employ SOA patterns such as separation of concerns between logical and physical layer of the presentation and business logic. Service integration is achieved using the ESB in some but not all business units. It is called a *Simple services'* level as it already takes into account basic process integration. This is an early version of SOA.

Fifth level: Application architectures are designed with a separation of concerns at the logical and physical layers. ESB integration patterns are used to support application and process integration to achieve sharing of services. This level is called *Composite services* level. The focus is set to support services as a part of chains. Connecting business processes and services is essential and services here can already form on-demand interactions. This level is the one where we can finally see principles of the SOA working.

Sixth level: Application architecture is decoupled from infrastructure components. There exists an extensive use of ESB architecture patterns to support the BPM. The level is named *Virtualised services* level and is characterised by decoupling services from applications and by the supply chains integration. Management, monitoring and events handlings are performed by separate services. The architecture is a grid-enabled SOA. New business models are deployed and the infrastructure can be set.

Seventh level: Adaptive enterprise. The application architecture supports dynamically reconfigurable business and infrastructure services and provides the SOA solution for internal or external partner consumption. This highest level is named *Dynamically reconfigurable services*. The final aim – business on demand – is reached here and an automatically reconfigurable architecture is introduced. The system can compose services and applications at runtime. The infrastructure is based on dynamic sense and response.

From this short summary we perceive that service maturity relates to integration capabilities at various levels of business: processes, governance, applications architecture, and infrastructure. Elements affecting maturity include business goals, business models, service identification and coupling to support business needs, their management, metrics that need to be supported by services, their standard usages and deployment capabilities as well as tools and technologies, and therefore, confirm that the migration towards SOA based enterprise is time consuming. Since a high level SOA is not widely implemented it's difficult to discuss SOA maturity in practice. Though, the seventh maturity level defined by the SIMM has a strong appeal to match the requirements of the intelligent assistant and therefore SOA is then well intended to be used for its design and implementation.

Scientists are already discussing of how indeed to reach this level in reality. One of very recent related works proposes a service-oriented architectural framework *"based on ontologies and agent technology" [Sasa'08]* for mental process tasks' semi-automation. It defines mental process task as a human task, which result depends on every individual task owner. *"The procedure used to accomplish a task is not predefined, but is the result of their own information, perception of the world, their experience etc. These are mental process tasks, as we name them, and they cannot be automated easily." [Sasa'08]*. This paper defines four levels of human tasks:

(i) Human tasks: the human task owner performs the task and enters results into the process.

(ii) Human tasks with user support: the system is capable to propose alternatives, agendas for the human task owner, who performs the task according to the system's propositions and enters results into the process.

(iii) Semi-automated human tasks: the system performs the task itself, though it demands a confirmation from the task owner.

(iv) Fully automated human tasks: the system has a high level of automation - tasks are performed without interrupting the task owner.

Starting from the third level, the system shall be capable of working on behalf of task owners. Though, if in the fourth level, the system is independent and the overall business process execution is automated. In the third level the system still needs a human verification. Researchers, here, agree that the fourth level is *"a long-term future vision, which may not even be always desirable" [Sasa'08]*. Sasa's et al. research work also demonstrates that the current SOA can support only first level human tasks and proposes to improve/extend SOA in order to reach a higher level of self-standing system automation. This important outcome of SOA extension is also valid for the design of an intelligent assistant. It means that in order to use the SOA for the development of cognitive agents, the SOA needs to be extended to incorporate methods in order to control domain knowledge and perform process configuration management. In the next sections, we will review works that could help to answer of how the development of this specific SOA extension could be done.

5.3.2. Model-Driven Architecture

Claus Pahl demonstrates *[Pahl'07]* that the extension of SOA or the integration with other architectures is possible and even needed. He shows the importance of domain modelling as a mean of abstraction and the need of Model-Driven Architecture (MDA) for service-based software systems: *"model-driven architecture as the development approach, applied here to service-oriented architecture as the platform and ontology as the semantic modeling technique"*. MDA is widely accepted to be a foundation for reusing domain knowledge and is treated as a domain engineering technique *[MDA'11]*. Domain engineering is an activity of collecting, organizing, and storing past experience in a particular domain in the form of reusable assets as well as providing adequate means for reusing these assets *[Czarnecki'00]*. The outcome of domain engineering techniques is a domain model, which defines the vocabulary used in the domain, classifies and if possible decomposes concepts, ideas and observations (heuristics), and includes varying consistent features' specifications. MDA uses the domain model and additionally provides a method for reusing knowledge concerning a specific application or system. It is intended to provide a comprehensive solution for application's interoperability and portability in the future as it separates system design from concrete system architecture. MDA aims to create a platform independent model (PIM), which represents a conceptual design realising the systems' functional requirements, hoping that it will survive changes that often appear in realisation technologies and software architectures. It is an open approach to interoperability using the widely applied Object Management Group's (OMG's) modelling specifications Unified Modelling Language (UML), Meta-Object Facility (MOF), and Common Warehouse Meta-model (CWM). C. Pahl in his study claims that the "*development of service-based software architectures requires the integration of domain modelling and architectural configuration aspects in order to implement services as reusable and composable entities in a process oriented environment*." He takes three different MDA layers: "*The computation-independent layer focuses in domain captures. The platform-independent layer focuses on architecture configuration and services process compositions. The platform-specific layer focuses on interoperability and discovery support." [Pahl'07]*, and demonstrates how they can be supported ontologically for service-based software development. This model-driven approach facilitates SOA implementation and naturally complements it with models of *"services, workflows, policies, SOA protocols, monitoring mechanisms" [Tsai'08]*. In particular, MDA is the notion of models transferability and is intended to support the raised claim of flexibility for SOA.

5.3.3. Event-Driven Architecture

Besides the domain control knowledge management, another important dimension is pointed to the process orientation – the need to achieve a final goal taking into account related contextual information during the process execution, that is to implement what we call a dynamic business process or a dynamic business process management to be precise. In simple terms, this dynamic BPM shall support rapid "on the fly" process adjustments. The major contributor and probably the candidate for the solution in the sense and respond areas is a Complex Event Processing (CEP) technique as it helps with situation awareness in real time. *"Events offer a mechanism for coordinating change in more ad hoc business processes. Without events, processes are limited to structured orchestrations that do not address unexpected process changes or nonsequential tasks." [Gartner'10]*. Indeed, the main characteristic for processes to be dynamic is the ability to respond to changes, which can be called external events. CEP is a part of Event-Driven Architecture (EDA) – an architecture, which generally takes care of events production, events detection, their utilisation and essentially of the reaction to them. EDA can be described as working in a style of SOA as the occurrence of an event can trigger the invocation of one or more services. Or opposite, a service may generate an event. *"The event-driven action may include the invocation of a service, the triggering of a business process, and/or further information publication/syndication." [Michelson'06]*. Goal-oriented and semantic processes may also benefit from EDA as goal states can be reached by the dynamic selection of tasks based on current events and situation. A very important characteristic of EDA is an extreme louse coupling. Event owners are not obliged to know event's subsequent processing or interested parties and so the system becomes highly distributed.

5.3.4. Discussion

Another candidate, though non-architectural, for a dynamic BPM solution is Adaptive Case Management (ACM), which is currently a subject of the OMG standardisation. The vision of case management is presented as a revolutionary way to change how "knowledge workers" – *"anyone who has a complex job to do that is hard to define and that requires "experience" to do it" [Swenson'10]* – get things done. Case management is responsible for the management of collaborative processes that require knowledge coordination, correspondence to content and resources in order to achieve the defined goal and where the path of execution cannot be pre-defined. ACM consists in goal orientation, strong communication, records keeping and document management, what as a whole represents knowledge management. Issues concerning knowledge management and process management potential integration we will discuss in the next section.

Taken as a whole, a general problem of modern distributed computing systems is that their complexity, and in particular the complexity of their management, is becoming a significant limiting issue in their further development. The SOA contribution in controlling this complexity is essential, though not sufficient to support automatic, adaptive and situation/context-aware systems like for instance the intelligent assistance development case is requiring. A key solution, supported by various different research studies, might be an extension of SOA by merging it with some others, even completely new technologies. The appropriateness of SOA in this context is widely assured by the research community. *"Several researches show that SOA is important toward the integration of distributed healthcare service." [Pan'08]*. Interoperability is mentioned as *"particularly important in biomedical research systems, which need to coordinate a variety of disparate types of data" [Komatsoulis'08]*. *"The healthcare industry provides many more scenarios in which web services can be put to use effectively." [Chappell'02]*. The SOA properties of interoperability and loose coupling become fundamental in developing a broad range of architectural products. Therefore, SOA is seen as a good candidate to stand for business process governance and management support. Even though, the SOA paradigm is never used as a method description utility allowing formalising these new tasks implementations, *"we assumed that a modern, web-based, multi-tier application and development environment would continue to provide a viable base for the future systems" [Black'04]*.

5.4. Knowledge engineering & Knowledge management

In the previous sections, we showed how ontologies are used by the semantic web community. However, the more generic ontologies purpose is to accumulate knowledge in a specific reusable mode. *"An ontology is a formal, explicit specification of a shared, machine-readable vocabulary and meanings, in the form of various entities and relationships between them, to describe knowledge about the contents of one or more related subjects domains throughout the life cycle of its existence." [Strassner'07]*. Historically, the first step to knowledge engineering is classification. People working in different domains (botany, biology, astronomy) spent hundreds of years in taxonomic attention to plants and animal, celestial bodies and others. Then, the next step after classification was the finding of some normality or regularity. In the last 150 years, nearly every science has made this transformation from the taxonomic – entity oriented perception, to the investigation of rules and strategies – process oriented perception, which resulted in acquiring knowledge in a certain domain. To be sure that we clearly understand the concept of knowledge, we should clearly see it in the context of other terms: data, information, knowledge, competence, intelligence *[Debenham'89]*.

D*ata* can be obtained using some instruments, are discrete or continuous, have units and have a reputation of being objective.

Information implies the subject developed by a sender and receiver and has an intention of a message. It aggregates data and is subjective to the person that created the message.

Knowledge is the sum of what has been perceived, discovered, or learned in a specific domain. It is a result of information acquisition and is inseparable from the context. It's at the same time a static representation of what is in the memory and a dynamic representation of continuous knowledge construction process.

Competence is an effective knowledge application to the given situation. It refers to an evaluation and partly to an evaluator.

Intelligence is an ability to combine knowledge, decision and action seeking for a good result. It is capable to elaborate new solutions.

Knowledge according to Nonaka and Takeuchi *[Nonaka'95]* can be *tacit* versus *explicit* and *individual* versus *corporate*. The knowledge formalisation makes tacit knowledge explicit in a way of forming written practices and rules. The new knowledge creation through the knowledge combination, using induction, deduction methods or hybridisation is performed making knowledge from individual to corporate. Knowledge accumulation allows saving time (reusability, information access), avoids of losing knowledge when experts leave, reduces possible errors in application, and enriches the global professional expertise.

5.4.1. Knowledge engineering

An engineering discipline called Knowledge Engineering (KE) and responsible for integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise emerged in the 1960s and 1970s and gained commercial interest starting in the early 1980s *[Liebowitz'01]*. *"Knowledge engineering grew out of the field of artificial intelligence and focused on building computer programs that simulated the behaviour of experts in well-defined domaina of knowledge." [Liebowitz'01]*. The process of KE involves the capture, representation, encoding, and testing/evaluation of expert knowledge and outcomes to a knowledge base, containing the set of facts, relationships among them and related heuristics (rules). Knowledge engineering strategy can be roughly divided into four basic steps that are repeated in loop: knowledge identification, knowledge development, knowledge enhancement and knowledge preservation.

Knowledge identification includes knowledge localisation, characterisation, organisation according to the priorities and evaluations.

Knowledge development consists in knowledge updates, enrichment, acquisition, comparison, and erasure.

Knowledge enhancement takes part through the way knowledge is spread, accessible, used, combined and created.

Knowledge preservation works with knowledge extraction, formalisation, modelling and retention.

Various technologies may contribute to knowledge creation (for example data mining techniques), modelling (via databases, ontologies), transfer (networks, forums) and application (expert systems, workflows). *"Among many enterprise assets, knowledge is treated as a critical driving force for attaining enterprise performance goals. This is because knowledge facilitates the better business decision makings in a timely fashion." [Han'09]*. In order to improve the competitiveness of an organisation a Knowledge Management (KM) discipline was introduced in mid-1980 *[Tisseyre'99]*.

5.4.2. Knowledge management

"Knowledge management is the process of creating value from an organization's intangible assets. Intangible assets, also referred to as intellectual capital, include human capital [people knowledge], *structural capital* [patents, trademarks, databases], *and customer or relationship capital." [Liebowitz'01]*. KM is strongly rooted in the knowledge engineering discipline. It deals with knowledge as a key resource of a modern organisation. The benefit of KM lies in the formulation of know-how, formalisation of experience, and collective education as well as knowledge diffusion. KM helps to reduce the time in repeatable applications of knowledge and to improve organisations' efficiency, flexibility, and reactivity for innovations and decisions. The major challenge in knowledge management is to pass from the stack of knowledge to expertise accumulation and competence. In order to establish KM in organisations, it is needed to save the expertise acquired through projects guidance, to express the technical expertise in a readable and shareable manner, to diffuse knowledge, to make a link in between the knowledge capitalisation and competences management, to create networks which allow sharing competences through distance, to take experience from the past in order to forecast future situations, to manage confidentiality levels, and, finally, to secure knowledge. All this leads into a huge variety of appropriate tools, technologies, strategies and processes (obviously based on numerous research studies), which turn data and information into valuable knowledge resources. Ontologies there play a principal role.

"Ontologies turn out to support a number of modeling tasks – from domain modeling to architectural configuration and also service and process interoperability." [Pahl'07]. In general, ontologies provide a mechanism to understand the problem description, harmonise statements of entities that are described from multiple perspectives (which is not a trivial task as different terminologies might be used) and unambiguously interpret limiting constrains what all together helps to accomplish the raised goals. As already stated ontologies are applied into the web-service domain as a conceptual framework for the semantical description of web-services in order to improve their composability into workflows. Ontologies also play a critical role in simulation applications – simulation integration, simulation

composability *[Benjamin'09]*. The increased use of distributed intelligence approaches to simulation modelling (distributed simulation, federated simulation, agent based simulation, etc.) accustom the employment of ontologies. For example, an ontology-based service-oriented simulation framework was proposed by Tsai et al. for storing *"relevant information useful to compose simulation applications, and items stored also cross reference to each other to facilitate reusability and rapid application composition" [Tsai'08]*. Ontologies are also used as data model, for resources description as well as for solving semantic interoperability problems during the service usage. *"A unified view of services is introduced by means of a service ontology, service classification and service layer architecture." [Weigand'09]*. The strong point of ontologies is that they help to understand systems easier, and serve as a communication system / translator. *"The network and system management community has recently become interested in applying ontology-based semantics to the development and operation of management systems." [Strassner'07]*. Macris et al. propose an approach to capture and to represent the knowledge concerning healthcare processes training material *"as an ontology-based knowledge network (a training scenario serving a specific training need)" [Macris'09]*. In autonomous systems, that is the ones that can manage themselves, ontologies are seen to be particularly important, as they can enable network components to be reflective and reason about themselves. This is a necessary first step in enabling networks and systems to become *"self-configuring, and from there self-healing and self-optimising" [Strassner'07]*. The standardisation process performed by the World Wide Web Consortium's Semantic Web group on the Resource Description Framework (RDF) and the Web Ontology Language (OWL) has provided a representational stability and has motivated the increasing adoption of ontology languages in management systems design, though still some weak points of ontologies that we overview in the following need to be sustained.

The first issue that engineering using ontologies faces is the construction of the ontology. The fundamental problem is that *"ontologies represent knowledge in a fundamentally different way than established information and data models" [Strassner'07]*. The solution to this might be to provide a *"precise definition of each concept and relation, and define the common unified schema for the mapping of the local database schemas" [Temal'08]*.

Another research issue related to the construction of ontologies is to determine the appropriate scope and the granularity of the ontology *[Benjamin'09]*. In some domains the information is so vast and heterogeneous, and the relations among it are so complex that in order to avoid errors it becomes essential to have an IT technologies based foundation. It is crucial for example in the medical domain: "*the Healthcare domain not only provides challenging opportunities from managing knowledge but also is one of the domains where it* [the knowledge] *is most poorly understood and deployed" [Gibbons'10]*. The knowledge granularity becomes important when the alignment of several ontologies representing the same family of business processes shall be made. A mediating model needs to be created to aggregate two or more ontologies. *"Recently, the documenting ontology mapping patterns has been proposed, which supports domain-specific mapping decision making." [Strassner'07]*. However, the knowledge about families of similar business processes and this knowledge reusability in enterprise engineering projects still remains an open problem *[Ciuksys'07]*.

One more challenge is the inter-domain ontologies mapping that shall facilitate the shared use of data within and across disciplines. This latter can be answered by the use of a meta-ontology. *"At the highest levels is a top-level ontology that includes abstract concepts and relationships valid across domains." [Temal'08]*. Usually, ontologies are more or less local, that is goal-oriented, role-dependent. Sowa defines it as: *"The subject of ontology is the study of the categories of things that exist or may exist in some domain. The product of such a study, called an ontology, is a catalog of the types of things that are assumed to exist in a domain of interest D from the perspective of a person who uses a language L for the purpose of talking about D. The types in the ontology represent the predicates, word senses, or concept and relation types of the language L when used to discuss topics in the domain D." [Sowa'1999]*. While, a meta-ontology is based on the extrapolation of common properties, arbitrary chosen axiomatic concepts and rules of local ontologies and is expressed, usually, in the languages of mathematical theories (set theory, graph theory, functional analysis, etc.). It can be seen as a conceptualisation framework enabling the comprehension of actions in several related domains. The meta-ontologies domain is recently strongly investigated in software engineering. This way of conceptualisation can successfully address for example the granularity management problem by describing modularisation of down-level ontologies.

One more meta-ontologies' application, straight forward related to this thesis, is cognitive intelligence. Meta-ontology can be seen as a goal-oriented knowledge ordering (conceptual modelling) tool that aims to provide the designer of complex engineering system, i.e. an intelligent-agent-based conceptualisation, with a structured set of methods and rules, which allow him to control top-down a goal-oriented processes or activities conceptual modelling *[Gadomski'89]*, *[Gadomski'09]*. There the demand is not only to construct the knowledge base, but to build *"medical knowledge in such a way that a computer program used by a non-expert can use that knowledge to perform like a medical expert in making diagnostic decisions" [Gibbons'10]*, which, indeed, speaks of the way to merge KE with Expert System development. This particular focus on real-life eBusiness applications demands for a set of intelligent advisors, which guide the users through various stages of the process.

5.4.3. Process-centred knowledge

Hence, automating sophisticated, personalised expert activities, which are based on the application of professional knowledge to process an individual case, remains a great challenge. An automated case management simulates anyone who has a complex job to do, which is hard to define and that requires experience to do it. The opposite of this knowledge based work is a routine work, which can be defined and predicted. The knowledge-based work depends upon the situation and is highly affected by how the situation develops. From such perspective two types of knowledge are distinguished: descriptive knowledge: relation-based rules, physical laws, theories, models, and operational knowledge: algorithms, methods, instructions, procedures, and specifications causing IA's actions. This additional notion of knowledge in a sense of content and correspondences coordination should ensure the correct execution of tasks. This now starts to be known under the term of process-centred knowledge models, which need to be identified and shall include not only domain, or inter-domain knowledge, but also to provide a context-rich knowledge. *"Most of these representational aspects have not been formally considered as people developed existing models. Thus, there is a semantic gap between using ontologies and using current management information." [Strassner'07]*. Macris et al. confirms that *"the development of a process-oriented healthcare information system essentially involves redesigning existing healthcare processes with the objective of shifting the focus from traditional transactional processing to cooperation and collaboration" [Macris'09]*. Analysis of processes provides a scope of what is needed from domain knowledge and helps to discover a feature model, which describes possible variabilities in process execution. *"Since dynamic process management becomes a core perspective enabling expert activities automation, the main concentration is to distinguish knowledge that helps to achieve tasks construction on-the-fly, a correct reaction to situations/events and overall a better process performance. The purpose of the process-centred knowledge model is to identify and categorize the type of knowledge to be created and accumulated in a process-centred way." [Han'09]*. There, solutions describing and managing context should be strongly considered, *"if knowledge is separated from the business process context, it does not lead to the ability to take the right action for target performance" [Han'09]*. Many KM researchers recognised that context is a *"crucial component to improve the understanding of knowledge." [Ahn'05]*. However, the subjectivity of the observer can affect the whole work, so automatic tools capable to process the contextual information need to be integrated into the knowledge accumulation process. Automatic contextual knowledge handling tools could help also during the execution of the process as process actions related knowledge can be created at runtime. Umar et al. already demonstrate the ability to *"automatically update the database by extracting information from the internet" [Umar'09]*. This is one of the intelligent agents' characteristics.

5.5. Context-awareness versus Situation-awareness

In previous sections we have looked at the modelling possibilities of a set of actions that the intelligent assistant shall provide and at the infrastructure and its potential management in order to maintain such execution. In this section, we want to consider the role of contextual information, which, in many respects, is a leading force of the cognitive assistance. We expect the contextual information to be considered, merged with a previous experience and dynamically related in order to achieve a meaningful whole when we employ an intelligent assistant. Though, *"the issue of proper organizing contextual data for their effective and meaningful exploitation by autonomic pervasive services is, and will increasingly become, a challenging one" [Castelli'08]*.

In the computing environment, context stands for *"any information that can be used to characterize the situation of entities (i.e. whether a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves" [Abowd'99]*. In general, context can be considered as a piece of knowledge related to a concrete unrolling situation. It can be understood as some set of facts or circumstances that surrounds or that can influence the situation or event. Context is characterised as being *"dynamically constructed" [Soylu'09]*, because even though some context dimensions are static (like date of birth), most of the context dimensions change more or less frequently. This implies that context can be temporal and so observations on how the contextual data evolve should be considered. Contextual data are also inter-related, that is they can only be seen in the light of other sources of contextual information. *"Context is a very fluid notion, and several researchers claim that it is very hard to formally define context and to abstract it in terms of variables and data models." [Castelli'08]*. Typically, contextual information can be generated by a variety of heterogeneous sources and at different levels of granularity. So techniques to acquire the contextual information (by a direct input of the user or retrieving it with the help of various sensors), to store it into manageable models, to interpret, to aggregate and, finally, to apply it are needed *[Chen'03]*, *[Abbas'08]*. These techniques were investigated from early 1990 under the name of "context-awareness", which, speaking of, is now migrating towards "situation-awareness".

The application capable to adapt to changing circumstances and to respond according to the context of use (that is to manage the context) is called a context-aware application *[Baldauf'07]*. Two groups of context-aware applications are distinguished: applications providing active awareness and the ones providing passive awareness. Active context-awareness happens when an application automatically adapts to discovered context by changing its behaviour. While an application providing passive context-awareness only presents the new or updated context to a user or a system. Definitely, the intelligent assistant shall demonstrate the active context-awareness and to be identity-aware though self-managed. However, the passive-awareness can be applied too, for example, while handling various exceptions. Context-awareness and adaptivity make a great sense in a pervasive computing environment. Research studies on the context-management technologies – technologies that develop a reusable and scalable middleware, which makes the design, development and deployment of applications that sense and react to various types of contextual information easier – contribute a lot to enhanced pervasiveness *[Black'04]*.

Early works in context-awareness concentrate on acquiring contextual data from data sensors and lead to data models creation to store this acquired data. This approach partially addresses the data model challenge, where the need is to conceptualise given raw contextual data. *"However, they generally miss in identifying a uniform model and a common semantics to describe the data." [Castelli'08]*. These constraints oblige developers to build new query languages and to constantly develop new components that correspond to the information at a hand. Missing common semantics, capable to manage multiple context information, limits current web-services to the search of contextual representation, which only matches their supported input source. In general, it implies the lack of any support for data manipulation and organisation. Consequently, enlarged and richer contextual information data models were proposed. Though, they have not brought advantages in facilitating querying by services as simplicity and generality were lost. Besides, the process of *"analyzing retrieved contextual data atoms and turning them into useful knowledge may be non-trivial" [Castelli'08]*. In other words, getting access to information content does not automatically imply the capability of comprehensive understanding about the situation – to be situation aware.

The recent proposals *[Temal'08]*, *[Chen'03]*, *[Ralyte'08]* have considered the issue of adopting specific ontologies to model context information in order to simplify querying and enable efficient context-reasoning. This has a specific interest to pervasive computing as services can perceive different information, depending on their current state. This feature is very important as it allows tailoring information to users' needs and viewpoints *[Simonet'11]*. The involvement of ontologies allows separating contextual data related to a domain (domain specific ontologies) from contextual data related to context management (policy ontologies): *"both the context model and the policy model are rooted in information models, so that they can govern managed entities" [Strassner'09]*. It is paramount that applications have general information about the context in which they operate at their disposal in order to reach the interoperability, the adaptability, to propose generic interfaces, the automatic code generation and the mobility of code *[Preuveneers'04]*. However, general knowledge of contextual information does not answer to the key point of the intelligent assistant of how context is managed across a number of contextual data while focusing to serve a specific situation at a given time. The lack of such recommendations about the functional needs and the context management creates a gap between fundamental research on context representation and actual context aware prototypes. The problem of a specific situation tailoring, in real, it is the aspect of a case management in the business process management domain, which then raises the proposal that perhaps we should be looking at suitable BPM technology to better manage the context.

Other interesting proposals focus on contextual data models that capture only specific aspects of context, which are of interest, and identify them in a uniform well-defined structure *[Pedrycz'07]*, *[Kocaballi'07]*, *[Lin'02]*. They actually granulate information and organise it together according to its similarity, functional properties, and spatial/temporal adjacency. This research area is called granular computing and considers the general issue of processing "information granules". The very important benefit of such an approach is the possibility to identify regularities in data, an eventually provide higher knowledge-level views. Though, *"to the best of our knowledge, there are no studies directly related to applying granular computing ideas to support context-awareness by pervasive services." [Castelli'08]*.

Hence, despite the presented approaches that address specific engineering challenges related to the context management, none of them propose full-fledged solutions for a layer that encapsulates mechanisms and tools to analyse and (self-)organise contextual information into sorts of structured collections of related data items, which could reflect comprehensive contextual knowledge related to a situation of interest and which can be exploited when needed. The main problem becomes not the lack of information, but finding what is needed at the time when it is needed. The expectation is to find bits of useful information, process them together with other bits to arrive at the relevant information which is required for decision-making. That is, models besides the contextual information shall provide also a way that it can be useable cognitively. This design objective has been named situation-awareness. *"Future pervasive and autonomic services must evolve from a simple model of "context-awareness", in which they access isolated pieces of heterogeneous contextual data and are directly in charge of digesting and understanding them, towards a model of "situation-awareness", in which they access properly structured and organized information." [Castelli'08]*. Mica R. Endsley defines that the situation-awareness is composed from three parts *[Endsley'00]*:

(i) perception of elements in a current situation,
(ii) comprehension of current situation, and
(iii) projection of future status.

The first level – the level of perception – is fundamental. Erroneous or incomplete data form an incorrect picture and lead to a total process failure. The second level is responsible for meaning-making. It should support cognitive constructions by integrating, interpreting and determining the relevance of data bits. *"The meaning must be considered both in a sense of subjective interpretation (awareness) and in a sense of objective significance or importance (situation)." [Flach'96]*. The highest level should embrace the ability to forecast future situations, to project events and their dynamics in order to anticipate the future and to allow timely decision-making.

The research on situation-awareness *"can be classified by the subject that performs this process – human or computer" [Kokar'09]*. Human-based situation-awareness process, which Endsley and many cognitive science scientists are focusing on, needs to be measured, evaluated and possibly supported. The aim here is to develop the operator's skills, if possible with computer systems help, to reach the level of an expert. While, the computer-based situation-awareness process deals with situations itself and should simulate a domain expert work on its own. The computer situation awareness process still lacks a more systematic treatment. It needs to be defined and implemented and calls for unambiguous specifications development. One of the trends in this direction is that of using ontology-based computing as a paradigm with some parts supported by mathematics and rules *[Kokar'09]*. Another research direction uses so called *"concept maps"* as the graphical representation of knowledge that could depict the concepts and relationship between them *[Teo'06]*. Sowa in his book *[Sowa'1999]* provides an overview of the AI treatment of context and proposes an approach to represent contexts (situations) in the formalism of conceptual graphs. AI techniques are usually based on a predicate logic that explicitly states the fact that the proposition p is true in the context c. All these formalisations let to envisage that the access to contextual information will no longer be direct, but rather via a layer of knowledge models, which pre-digest contextual information and provide it into a compact and expressive form to services. This layer will ensure a higher-degree of adaptability and autonomy of pervasive systems and will lead to proper and accurate computerised decisions.

5.6. Decision-making and support

Ubiquitous and pervasive computing environments demand systems to adapt to situations and challenge them in asking to *"be able to make implementation decisions at run-time, rather than at design time or compile time" [Paluska'08]*. Decision services are recommended to be implemented in order to incorporate business rules/policies and to support an automatic services' orchestration during a dynamic process execution. *"A decision service may be seen as the application of a set of business rules (that make up a policy) to generate a statement." [Weigand'09]*. These services are incorporated into the execution pathways at the decision points. Decision points are defined when the goal of a process is set.

The goal of a process is described via two elements: decision points and technical specifications, which characterise how to satisfy these decision points. Then, the decision-making evolves via a series of steps, where each of them has at least one decision point and ultimately culminates in a problem resolution. It is usually a systematic procedure based on a process of human thinking. This rational process comprises three major phases: definition of the scope of the problem, design and choice of the final solution.

(i) The aim of the first phase is to identify the problem and clearly define it gathering required information and organising issues to be worked on.

(ii) The design phase results in a finite and limited number of alternatives that can be finally evaluated. The evaluation is performed on identified and prioritised criteria.

(iii) Finally, the choice phase performs a more detailed analysis of selected solutions by verifying conditions and checking them against the real world constraints. It establishes the performance of alternatives related to a particular criterion. Following the analytical or experimental evaluation results obtained in this phase, the best alternative is chosen and the decision what to do next is taken.

Speaking of an intelligent agent we should not mix decision support systems with decision-making systems:

5.6.1. Decision support

A Decision Support System (DSS) aims at providing assistance to human operators in processing huge amounts of data, determining alternatives and if possible evaluating them.

It acquires data from a mass of routine transactions, analyses it using advanced statistical techniques, extracts meaningful information and if some decision rules are set, narrows the information to the range of choices, otherwise provides the evaluation on the selected data cut. It usually presents results in the form of various charts, pivot tables, diagrams or graphs. DSS management tools provide an interface to the operator – a real decision-maker – to limit or to enlarge the data set, which needs to be processed, to define rules for alternatives construction and to “play” with various section views. Thus, DSS facilitates the analysis, but does not replace the human, who makes decisions.

DSS is a tool that supports and improves human decision-making. It helps to analyse complex problems and to process vast amounts of analytical data and to arrive at decision options. It is usually adaptable to information format and uses analytical methods and reasoning to model data. Researchers and technologists have built and investigated DSS for approximately 50 years. They organise them in five broad DSS categories including: communications-driven, data-driven, document-driven, knowledge-driven and model-driven decision support systems. The earlier versions of DSS were stand-alone systems, usually developed under a concrete organisation’s order and based on various models. Newer DSSs are based on data warehouses that can include relevant external data and use Online Analytical Processing (OLAP) and data mining techniques for data analysis and Business Intelligence (BI) techniques for we could simpler say data monitoring in order to spot, dig-out and analyse the “hard” data that for example help to detect opportunities for innovations. Recent works suggest to additionally contextualise the decision support systems *[Beuscart'09]*. This initiative is essential in medical applications. Clinical decision support systems need indeed to take in account user-specific information in order to protect the patient’s safety *[Koutkias'09]*. DSSs are capable to study the effect of changes in one or more parts of the model and its influence on other parts of a model (sensitivity analysis), to check the impact of changing input data (what-if analysis) and to find the input values necessary to achieve a desired level of output (goal-seeking analysis). Evaluation of the overall effectiveness and applicability of DSSs has been a research topic from 1980 *[Phillips-Wren'09]*.

5.6.2. Decision-making

A decision-making system is in charge to replace the human-expert and to make decisions on its own. When we define a decision-making we refer to a mental process of a human to select a logical choice from the available options. Trying to make a decision, a person weights the positives and negatives of each option and considers all alternatives. For an effective decision a person must be able to forecast the outcome of each option and to revise previously made decisions (experience).

Automated decision-making systems are usually based in a broad sense on Artificial Intelligence. A typical category of such systems are expert systems - computer systems that behave like humans. An expert system, typically, consists of three parts:

(i) a knowledge base, which contains the expertise, usually, acquired by interviewing experts,

(ii) an inference engine, which interprets the submitted problem through the rules and information logic stored in the knowledge base, and

(iii) a user interface that allows the user to express the problem in a human language.

The expert system works converting the knowledge of one expert, called a Subject Matter Expert (SME) into a software code, merging this code with other codes (based on the knowledge of other experts of the subject) and using queries to describe a problem and to retrieve the answer. Such expert systems may or may not have learning components and are usually using knowledge engineering techniques. An interesting characteristic supported by an expert system is real-time adaptation. Real-time expert systems are designed to be used over time and are capable to respond to constantly changing input data *[Moore'88]*. However, the features that the representation of such data include in order to define possible changes are criticised as being not enough truth worthy (we cannot envisage all in some domains) and as taking too much time to check each data bit and each conclusion.

AI technologies contribute for several types of decision-making systems based on: Case-Based Reasoning (CBR), Artificial Neural Nets (ANN), intelligent agents, genetic algorithms and fuzzy logic.

Case-based reasoning is a technique which uses a human thinking manner. Once a person has accumulated a large number of similar cases, he starts making generalisations in order to acquire a general rule that could work for a set of data and facilitate new data processing. An induction method (from specific to general) is the one helping to enhance understanding, extract knowledge and respond to new cases/problems. However, the task of making decisions from the set of collected data is not an easy one. The CBR methodology summarises this through four steps: history/past data storing, method for analogy/similarity determination, similar case retrieval for a new case and, finally, similar cases classification in order to extract common rules, new knowledge. Starting from then, a new case firstly passes accumulated rules and only if a solution is not yet indicated in these rules, a search of a similar case is performed.

The ANN are often applied for financial portfolio planning and face recognition. Recently, some works using the ANN approach were published with application to the medical diagnosis for the risk stratification related to the possibility to get a myocardial infarction *[Atoui'06a], [Atoui'10]*. However, the employment of this technique remains experimental.

Intelligent agents (bots) are most commonly found retrieving personalised information from the web, managing online auctions.

Genetic algorithms are, basically, computer instructions on creating a population of thousands of potential solutions and evolve the population toward better solutions. They are applied mostly for scheduling and logistics problems.

And, finally, fuzzy logic is a type of reasoning, which deals with incomplete or uncertain data *[Manzi'10]*. It recognises that logical statements are not only true or false, but can also vary in a range from almost certain to very unlikely. Typical application domains are real estate appraisal and environmental control systems.

5.6.3. Intelligent assistant for decision-making and decision support

The personal intelligent assistant is intended to be capable to use techniques from both the decision support and the decision-making domains. It takes a role of a domain expert and makes professional decisions at points where the system user has no competence to decide, and at the same time, it supports and even encourages the decision-making by a system user at other points. It’s all more evident with a small example from the medical domain. The user using an intelligent personal cardio assistant wants the system to guide him during the record of a personal ECG, to evaluate the quality of the obtained ECG, to choose an algorithm to process it (which better suits in his case) and a method to assess it (again which better suits in his case), as in all these steps the user has no competence to decide. But at the same time the user is free to choose to record a better quality ECG or not, to give access to his personal medical data for an enhanced assessment or not and, finally, to choose what to do with the provided results, for example in the case of a minor alarm (the status of a possible risk of myocardial infarct), to send this alarm to a specialised unit to take care of him or to take an appointment with his doctor, or even do nothing (and repeat the test). Then, we could summarise that the personalised process that is provided by the intelligent assistant acts as a decision support system, though the services' orchestration implementing this process performs as a decision-making system.

The principal goal of a decision-making and support system, such as the intelligent assistant shall be, is to maximise the service that it is capable to provide while minimising the required interactions with a user. Any mix of techniques might be used and is welcomed if it answers to this demand with a sufficient precision.

5.7. Summary

Scientific projects on knowledge management (KM) and business process management (BPM) integration in order to construct an intelligent layer capable to provide an automated expertise to the user upon the data obtained using ubiquitous computing techniques are relatively new and the complexity of this challenge often misleads in achieving applicable results. Usually researchers' attention is captured by some physical or practical constraints, whereas a task execution procedure itself is already known. Then the automation (which is often new and valuable) terminates upon solving these technical limitations, but it does not replace a professional care.

Another common target is to model fundamental knowledge, though without taking into consideration the context this knowledge is applied in. This fundamental knowledge modelling enables new services design, but it is often insufficient to provide these services to end-users. Indeed, without application knowledge (in the sense of how to apply, to employ the knowledge), professional care services cannot be implemented and deployed correctly. Here we would remark that knowledge should not be seen as some undivided unit and that a clear separation of the fundamental domain knowledge from the context (process-centred or service-centred) knowledge should be very helpful in order to use it properly.

Services personalisation is another related and contributing research area. Here again personalisation very often is depicted as giving the user the possibility to be responsible for some business process or to configure the system by himself, what should be rather called customisation instead of personalisation. The customisation approach shall definitely be encouraged as it contributes to create user-friendly environments, but again it cannot be sufficient for the intelligent assistant system as here the user is not an expert and expects itself to be guided and supported.

Then the most important lesson we learned is related to the overall integration issues. The final solution asks to find a flexible architecture as a base and to model processes which instantiations are tailored to the specific user context. These processes shall be based on determined domain knowledge and should assist the user in decision-making, with a quality level that should be comparable to the one provided by a human professional.

The following thesis chapters will define a method for the development of personalised intelligent assistance systems. This method combines pervasive sensor-systems, the processing of user context, domain knowledge and process management and is expressed in terms of service-oriented architecture.

Chapter 6 – Process knowledge modelling

6.1. Introduction 119

6.2. Expert activities automation 120

6.2.1. Motivation for expert activities automation 120

6.2.2. Challenges of expert activities automation 120

6.2.3. Procedure based activities automation versus expert activities automation 121

6.3. Business intelligence view 124

6.3.1. Of-the-shelf assistance implementation 124

6.3.2. Personalised assistance implementation 125

6.4. Compositional model 126

6.5. Dynamic business process 128

6.6. Core-ontology for decision support process knowledge modelling 130

6.7. Discussion 133

6.1. Introduction

The very nature of pervasive computing implies that contextual information must be continuously monitored and taken into account for improving system usability. This fact requires ubiquitous systems to analyse a flow of input data, select meaningful information and provide services adapted to the environment and user needs. However, the complexity of ambiguously related information and huge amounts of heterogeneous data, which can now be easily captured with smart technologies based on various sensors, tend to discourage designs involving intelligence for reaching user-specific objectives. Current systems narrow themselves to data collection, speedy data transfer and data pre-processing, leaving the task of providing a final solution to an expert of a specific domain. A delivering of autonomous professional care at a time and a place when it is needed seems to be hindered by its complexity. A lack of methodology on constructing intelligent assistance systems and, therefore, a lack of enhanced tools that would enable services to modify their activities dynamically according to the context of use, block the evolution from this passive information storage to a proactive services facility for its users.

The essence of the problem lies in the difficulty to mimic expertise. Therefore, we start this chapter by analysing differences between typical procedure-based business processes automation and expert activity automation. As a result, in order to automate expert activities, we propose to integrate context related knowledge into business process modelling via the development of an intelligent process management agent. This concept is partly supported by current knowledge modelling techniques, which provide ontology-based tools to construct knowledge of a specific domain or about a particular context. Yet, the formalised knowledge that these models contain does not cover how this information could be applied to conduct an automated goal-oriented expertise process, nor how such processes could be modelled. Further on, in this chapter, we answer this part of the challenge. We propose grounded means that help understand and design the intertwined aspects of domain specific knowledge and contextual information in order to provide an intelligent and personalised control on information processing. This solution expands the typical logic layer of an application with a view of business intelligence that supports the creation of advanced, adaptive business processes, which could provide person-specific intelligent assistance when the person's context is gathered at system runtime.

We concentrate ourselves on web-based technologies as their potential allows developing a certain degree of autonomy, which significantly increases the accessibility of intelligent assistant services by using the Internet. We propose to extend the typical business process development view, mainly focusing on services composition, with a service cluster, which supports the design of complex, automatically reconfigurable services. The latter, then, are orchestrated by content and meaning instead of only by syntax, while using process knowledge accumulated into a compositional model, which performs a mapping between a domain knowledge representation and a business process context. As a proof of concept we have developed a core-ontology supporting the creation and the orchestration of decision-making services. We demonstrate the relevance of this approach in further chapters for the design of automatic composition of services helping any citizen select an optimal and personalised sensor-system and improve decision-making.

All these proposed mechanisms are intended to enhance web-based systems to become more sophisticated, complex, and capable of fulfilling their great promise of distributing expert knowledge into non-experts' hands.

6.2. Expert activities automation

6.2.1. Motivation for expert activities automation

Professional care has become a common and widely used service. Domain experts, for instance therapists, teachers, bankers or consultants, are contacted in order to help to reach some state, to solve problems or to answer a question specifically related to the demander. However, the mechanism of how this service is provided has not changed for years. It is essentially based on a human-to-human relationship: a person or a company (a client in general) comes with the problem and then the expert employs one or a set of techniques such as: interrogations, tests, simulations and follows typical or personal scenarios by making exceptions when needed to finally respond to the request of his client. As an example we could take a patient, which after getting a chest pain contacts his general physician (an expert) expecting that with his guidance (analysis, prescribed medications, etc.) he could get well soon. Techniques applied by an expert require knowledge and practice in the use of domain specific utilities, an experience in working with such type of problems or clients, an effective expert-client communication, and a delegation as well as a supervision of tasks, when some intermediate tasks are performed by or with the help of other actors.

This already well-settled procedure of human-to-human professional care, however, has its disadvantages from both: the client and the expert side. The client must find an available expert, must adapt to his working hours and schedule. An expert should develop good communication and marketing skills, while a client should learn some technical vocabulary helping him express better his needs and protecting him from misinterpretation. The process itself requires creating a mutual confidence between expert and client, which, if needed, may be based on credibility checks. In general, the professional care may be rather expensive, if not in terms of money, then in terms of time, travelling and equipment costs and, therefore, not necessarily accessible to everyone. All these issues, together with the spread of advanced technologies, promote the idea of expert activities automation into a digital service (even if basic, at first) that would be accessible and available to all citizens, anywhere and at any time. This automation is anticipated in various domains, such as preventive medicine, where early diagnosis is expected to be performed automatically, at anytime and anyplace. Also, a thorough automatic follow-up of knowledge assimilation processes could improve and boost e-learning. Similarly, the sales and consulting sectors may benefit from services offering the user an automatic help that is more adequate to his needs.

This all could be reached if the delivery of a professional care would be changed from the human-to-human interface towards a human-service-human, that is expert-service-client interface. However, the automation of such complex, nonstandard and expert-requiring business activities, which we call expert activities, is a very challenging task.

6.2.2. Challenges of expert activities automation

Currently, new technologies propose various automations related to expert activities, such as: electronic application forms, electronic data transfer and management systems, data processing and analysis algorithms, various reminders, alerts and others. Professionals often make use of them gratefully as they speed the whole process up. Yet, these technologies, even sometimes announced so, do not replace an expert, who is taking a role of a principal coordinator or advisor, and who manipulates the whole process in order to drive it towards the most adequate personalised solution.

A great majority of currently available automations are based on well-known and steady procedures, and are implemented as reusable atomic or composite business services, generally accessible via the Internet. Examples of such procedure-based activities include popular services such as booking a hotel room or reserving flight tickets. However, if a demand of a user is more complex, specific or intrinsic to this user, such service cannot fulfil the request and so the task is passed on to a human expert. Such exclusive steps typically require thorough knowledge in a specific domain, which, as a rule, is held by professionals of the area. Then, for us, the automation of an expert activity is not the one which provides some improvements, but the one aiming at assisting a user in the application of specific domain knowledge, which depends on the circumstances of use and which adequately answers a concrete demand of the user. This type of activity tends to cover human intellectual processes, such as teaching, diagnosing or decision-making. In common language, we used to name the future service that an automated expert activity would provide as “a professional in the pocket”, though, among the domain specialists, this functionality is seen as an intelligent assistant.

Indeed, the label of an “automated expert activity” is often glued to any automation that provides some optimisations related to the complexity of professional care execution or allows overcoming some technical limitations. However, none of the current automations are really capable of eliminating steps requiring process coordination and adaptation to changing and user-specific situations that are now performed by a human-expert. This top-level (from the perspective of a process) management coordinates when to stop the whole process and when it is the time to finally formulate the conclusion. The outcome of such a business process has the quality of being specific to the managing expert, yet may be confirmed (or not) by other experts or by some widely accepted - specifically defined for this purpose - assessment procedures.

In the following, we will discuss the subject of expert activity automation by comparing it with the automation of a procedure-based business activity to identify issues transforming automation of expert activity into a complex task.

6.2.3. Procedure based activities automation versus expert activities automation

Procedure-based activities are being successfully automated in various domains, while expert activity automation is often presented as a long-term future vision, although both activity types have, at a first glance, many similarities. They both tend to replace a human, to give a personal assistance, and they both require specific domain knowledge. First of all, let us give two simplified examples of each activity type. We name PA a procedure-based activity helping a system-user with plane tickets reservation and EA an automated expert activity, which is capable of locating the correct electrode position on a human torso, enabling the user to record the best quality electrocardiogram (ECG) by himself. Here, EA aims at replacing a skilful nurse, in a specialised hospital unit, recording an ECG that provides a patient-specific signal for further medical diagnoses. Our goal is then to analyse the modelling of these two examples according to their input data, working practice and output, seeking to highlight differences of the PA and EA.

Procedure-based activities and expert activities can both be structured into five broad stages: there is *a situation* (1), from which *a need* (2) is rising, then *a formal request* (3) is made and *a process* (4) is launched in order to obtain *an answer* (5) to the user needs.

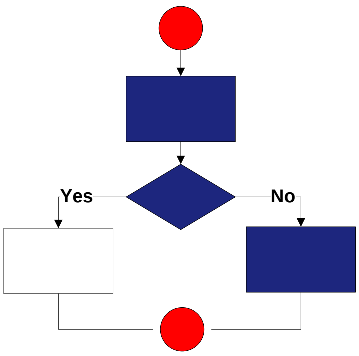
Obviously, input data have to be described before the process is launched, so stages (1), (2) and (3) include the design of input parameters. In the PA case, the system user is an active client, capable of managing the online booking service on his own. He analyses, by himself, his situation. He clearly understands his need and is capable of formulating his demands in a formal request form. For example, a travel type – one way, a destination – Berlin, an outbound date – 2000.01.01. In the EA case, a service user is, on the contrary, a passive client, who has no specific medical or anatomy knowledge to evaluate his own situation and is not sure how to express his needs. So he is not really able to formulate his demand. Here, a service performing an optimal electrode-system selection should guide the user in order to obtain the possible user-related contextual information that would help this service determine the user situation and his needs as a formal request. Hence, the request form, in the EA case, can differ and basically depend on contextual information and on available methods used to acquire and to process this information.

Then, when the user specific formal request (3) is set, the input data analysis (4) seeking to obtain a final result (5) is launched. For the procedure-based activities, a chain of tasks is usually designed, where each task might demand additional information (or decision) from the client and a final task in the chain provides an answer (5) to the initial user request. Intuitively, we can even define a chain of tasks for the PA case: a flights search, a flight selection, a client identification, payment and confirmation. The latter answers the initial user demand with a confirmation of booked tickets and details upon his flight. At present, we could repeat this online service to book the flight to Berlin and we would obtain the same final result (a confirmation and flight details), which shows that procedure based activities tend to produce repetitious results and are easy to test.

Meanwhile, expert activity is often seen as a black box, which produces hardly tailored results. Here, we suggest perceiving a mental process not as a black box and not as a complex chain of tasks, but rather as a labyrinth, which can have several entries and several exits, and for which some passing rules (obtained from previous experience or from evaluation of the current steps) can be set in advance and/or at a runtime. Indeed, expert activity design strongly depends on a specific request, that is an entry point to the labyrinth, and on the expert knowledge base (practice / experience / habits), which holds labyrinth navigation rules. So the model of expert activity (the labyrinth) instantiated for different clients may contain different passing routes or workflows of tasks to be performed, as a human expert would have proposed to each client a personal and individualised service. In our simplified EA example, if a client is a man and the signal quality is not sufficient (or qualified as insufficient), he might be asked to shave his chest in order to increase skin conductivity. Also, several iterations of signal recording or an identification of various possible recording postures, electrode placements or other changes might be done in order to improve the signal quality depending on the client context, on available data processing methods and on intermediate results. It is hard to repeat the EA activity as each time it tends to take a slightly different path adapting to the context that is never the same and the situation that might change.

The labyrinth can be roughly seen as a state machine, where a pointer changes its state by choosing one out of four possible directions: the next state is the exit (x), the next state is a new task (+), the next state is an error, the pointer comes back to one of the previous states and sets a rule based on the error analysis (-), or finally the next state is the repetition of the current task using another method in order to gather additional information, which will improve confidence about the next direction to take (°). This concept of state machine

Procedure-based activities	Expert activities
(1) The situation analysis is made or supervised by the user.	*(1) The situation analysis is completely performed by the system.*
(2) The need is defined by the user.	*(2) The system guides the user to collect contextual information.*
(3) The request is set up by the user according to a static input data form.	*(3) The request is designed by the system, which holds passing rules related to the input data and available processing methods.*
(4) The business process is static. The paths are defined in advance.	*(4) The business process is dynamic. The paths are set at runtime.*
(5) The result is repeatable and easy to test. Validation is performed by the user himself.	*(5) The result is poorly traceable. Validation must be performed by the system.*
The task management of procedure based activities is usually supported by a classic workflow modelling.	Dynamic modelling support is requested for expert activities of labyrinth type - with several entries, exits and navigation rules that are dynamically set at runtime.



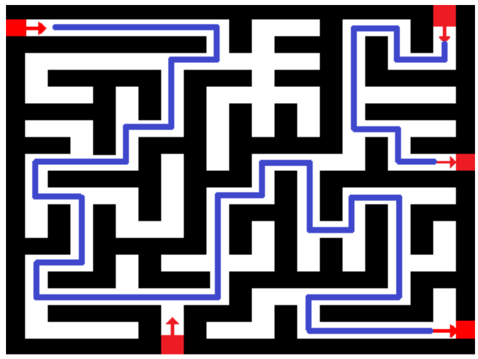


Table II.6.1 Summary of procedure-based activities and expert activities overview

explains well another characteristic of expert activity: the dynamicity. The decision is made according to instantaneous parameters and so the enhanced adaptation to the situation and its context shall be demonstrated.

This brief comparison on procedure-based and expert activities (summarised in table II.6.1) shows that expert activity is an advanced process as it is self-standing, adaptive and dynamic. It encapsulates not only specific domain knowledge, but is also guided by some intangible intelligence towards the optimal solution. This, yet to be formalised intelligence, implies that the path and the result of such a process may change even with the same input. The lack of explicitness of how to manage this process, poor traceability and unclear validation procedures remain the main blocking element for automating sophisticated and personalised expert activities, which are based on the application of professional knowledge to an individual case. Besides that, the automated use of knowledge in practice is extremely difficult, as knowledge is usually heterogeneous, voluminous, poorly structured and steadily evolving. Current methodologies and process modelling techniques used to automate procedure-based activities do not take this complex functionality into account and, therefore, are insufficient to handle the automation of expert activity.

As an answer to all these highlighted shortages we propose a solution based on the integration of business process management (BPM) and knowledge management (KM) techniques. We suggest extending the typical business process view, mainly focusing on services composition, with a business intelligence view which aims at supporting the design of automatically reconfigurable services. To fulfil this objective we design an approach that consists in setting up an intelligent environment based on a context-aware concepts' ontology and on related context handling rules for services orchestration. It is a generic method that allows the capture of the knowledge and of its granularity into a set of models that support dynamic business processes capable to implement any expert activity

6.3. Business intelligence view

In this section, we describe a new schematic method *[Krupaviciute'10a]* to model a logic layer (a middle tier) of a service-oriented application in order to enable the design and implementation of complex business processes and, thus, the development of intelligent assistance. Our approach (Fig.II.6_1) consists in three views: a business process view, which describes the business process in some application domain, an operational view responsible for low level tasks and activities implementation, though, still material independent, and a business intelligence view, which comprises components enabling an automated business process management.

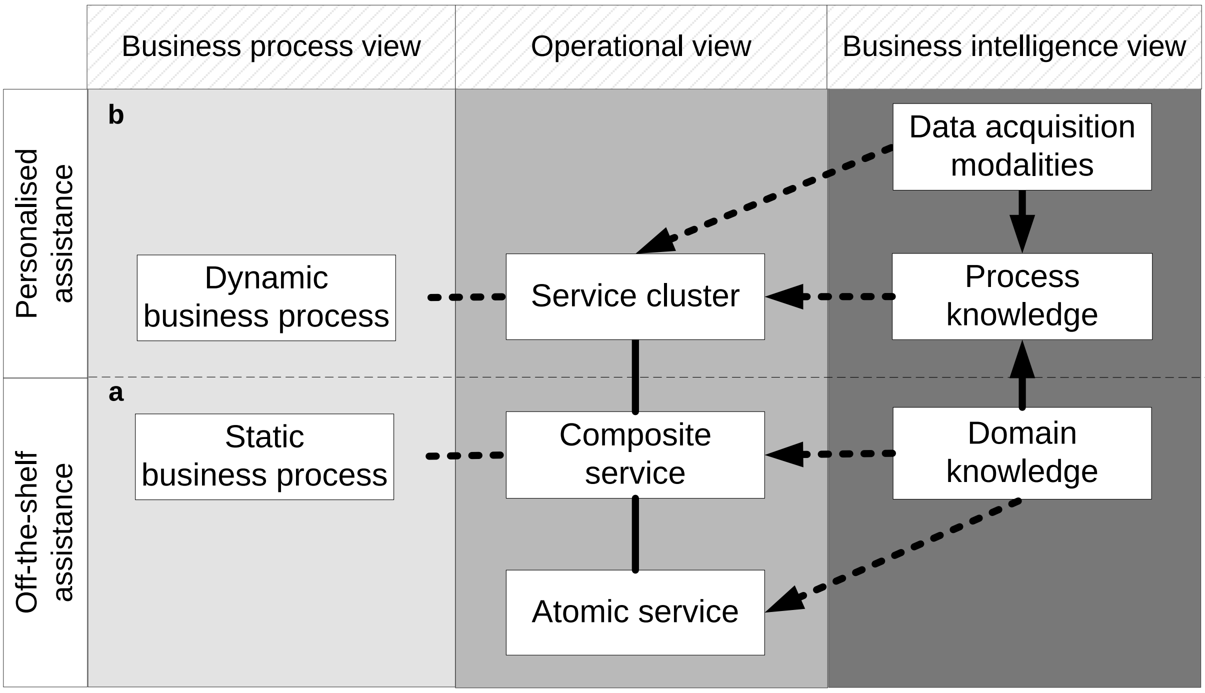


Fig.II.6_1 Approach of data and model driven services orchestration

Our approach basically extends the currently used approach – the layer a) in Fig.II.6_1 – by putting on the top a new layer – layer b) – that takes in account the modelling of process knowledge in order to automatically manage the process. Process knowledge includes contextual information related to the process and domain-specific knowledge and provides the required dynamicity.

6.3.1. Of-the-shelf assistance implementation (Fig.II.6_1 part a)

By definition, a business process is a set of activities determined in an ordered sequence to reach a fixed purpose in some domain. In service-oriented applications, an implementation of a business process involves atomic and composite services, which are based on domain knowledge representations. Here, atomic services commonly implement basic functionalities that are not further broken down to a finer lever and usually correspond to tasks (atomic activities) in a business process model. Meanwhile, composite services are designed as coarse-grained service structures that include more business functionalities and implement sub-processes (compound activities) of business processes. From the implementation point of view, a sub-process is equivalent to a business process itself - a set of atomic services that are orchestrated to execute a defined process and, as a whole, are implemented as a composite service. Frequently, several different and often remote services can implement the same activity. Then, a composition mechanism responsible for finding the “best” matching service (i.e. the fastest one, or an available one) and composing it into the predefined business process has to be set. This workflow composition automation is currently an active research domain, where various atomic services selection methods are proposed. Generally, they are based either on the representation of domain knowledge through ontologies or on services descriptors and semantic annotations.

As a rule, a typical business process has a default configuration and defined usability regulations for process execution. Therefore, in the context of this thesis it is named a static process. The business service that it provides should be named an off-the-shelf assistant rather than a personal assistant.

6.3.2. Personalised assistance implementation (Fig.II.6_1 part b)

A personalised assistance demands that the business process adjusts itself to concrete circumstances at workflow runtime in order to provide an optimal result. Indeed, situations demanding a business process to differ may appear, for example when an input data set is incomplete. For instance, in order to get a full set of data “abcd”, a predefined workflow will start with variations like “abc”, “acd” or “bcd”. The semantics of the context will also be particular if an input data retrieval method is different, which, in that case, could change the input data content. Then, instead of the “abcd” set, the business process would get a chain of “1234”. This latter change would demand that content-sensitive business process activities should be implemented via services compatible with the input data content syntax. The mentioned incomplete or different input data content causes variations in the outcome of several services, which then influences selection of the following services in the chain. Besides that, it is also possible that a business process allows several services to implement the same activity in parallel via grid computing and then selects the best result or expands intermediate, transition data sets. In this case the services’ availability for these special activities can be a source of different business process instantiations at runtime. The requirement of process adaptation to its context at runtime asks for a reflection on services modelling, demanding to create an environment where composite and atomic services could be dynamically reconfigured and/or built according to the incoming contextual data in order to implement business processes, which would simulate a human mental process. We propose to model this problem using three conceptual elements: a *dynamic business process*, *service cluster* and *process knowledge*.

A **dynamic business process** is a process that can provide different workflows, where each is adapted to the context of a specific user. This context as a whole or in part may be known at runtime. A dynamic business process is defined with two main elements: (i) a set of tasks and (ii) an intelligent management agent, which encloses process knowledge.

(i) The set of tasks corresponds to various business process functionalities, which are typically implemented via atomic or composite services and are present in a service cluster.

(ii) The intelligent management agent is implemented via a service orchestration adapted to the context of each client. A performed on demand and well supported orchestration ensures dynamic process reconfiguration. A separation of concerns is present, here, since the implementation of tasks via one or a number of services is separated from the process knowledge containing guiding rules. Therefore, they can be appended, corrected, replaced or deleted independently. Thus, rules that define a concrete path in a workflow, which is specific to a user, and a set of chosen atomic or composed services implementing tasks in this workflow, are identified at runtime according to the context and the available and suitable business process implementing services and compositional model that is used.

A **service cluster** is aimed at implementing a dynamic business process – a process that corresponds to the design of an intelligent process, where a workflow result depends on a process executor and execution circumstances. The cluster of services contains a set of business process related atomic and composite services able to fulfil goal-oriented process activities. An evaluation of the services' appropriateness and their orchestration to reach a business process objective should be ensured by the process knowledge, represented in a compositional model.

Process knowledge (compositional model) is a mapping between a domain knowledge representation and a data-driven business process context. The design of such a compositional model (or models) supports the management of business processes via process rules, and allows tuning an improvement procedure which would better personalise the activity. The compositional model takes the role of a subject matter expert (SME), a professional in the given domain with an expertise of fundamental knowledge application.

6.4. Compositional model

The compositional model is a core element for enabling the automation of expert activities. It encapsulates knowledge on a specific problem. This problem-specific knowledge is modelled in two principal stages. First, the analysis of a domain as well as the analysis of its context are performed separately by breaking them into granules - concepts. Then, in the compositional stage, complex and often new structures linking concepts to other concepts are formed. These concepts are taken either from the set of domain concepts or from the set of the domain related contextual user information concepts. After this stage, the resulting model is named a compositional model. In this manner, the model (or even a set of models) is constructed to represent a comprehension of concrete problems and so encapsulates the intelligence of an expert. Here, domain specific concepts have meaning only in relation to contextual concepts. Formally, this problem-specific process knowledge modelling procedure (Fig.II.6_2) has 3 main components: *domain-specific* and *context-awareness models, compositional model* itself and outcomes of *improvement, management, measurement* (or other) *procedures [Krupaviciute'09]*.

Domain-specific and context-awareness models represent a two dimensional approach to the definition of a problem the way an expert could operate it. Each model corresponds to a different dimension and thus describes its area data elements and determines the rules between these elements autonomously. The *domain-specific model* is a formal description of

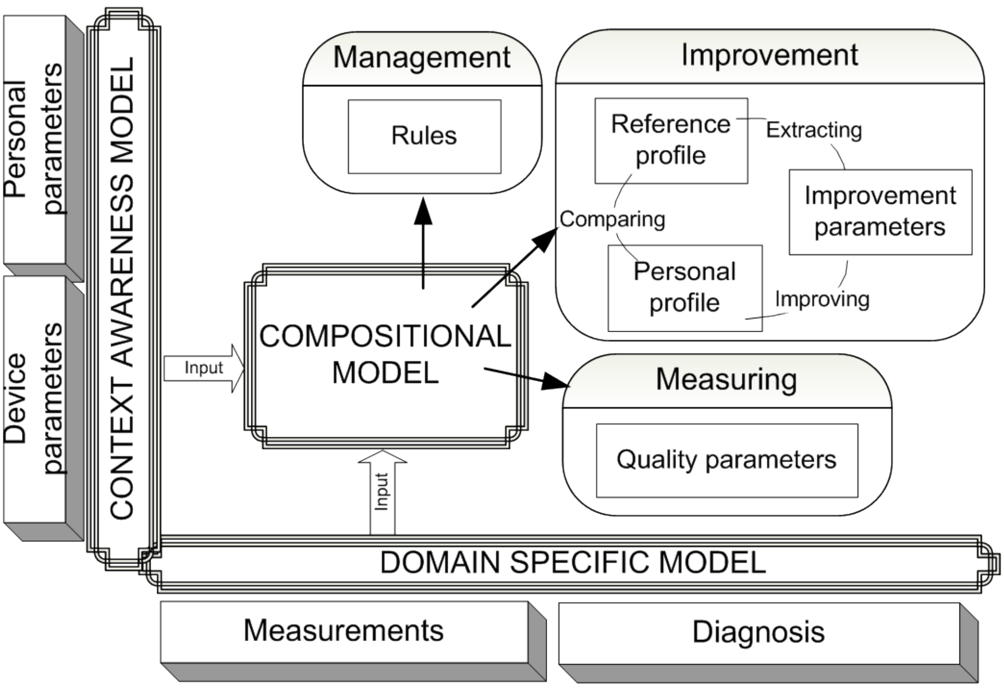


Fig.II.6_2 Process knowledge modelling schema

any particular domain and serves for a better understanding of this domain. Specific domain related knowledge obtained using various observation methods is represented in this dimension. Furthermore, it performs an acquisition of additional knowledge about a defined specific domain by iteratively processing already defined data structures and rules. The domain-specific model can be roughly described through domain related measurements and given diagnoses based on these measurements. Remembering our defined EA example (in section 6.2.3), the domain-specific model would contain an ECG signal analysis and a signal interpretation model. The measurements part would include parameters computed during the analysis of the ECG signal (heart rate, wave duration, ranges of amplitude, ratio between peaks, etc) and signal processing algorithms (signal filtering, ECG feature extraction, etc). The diagnosis part would describe rules on how measurements can be organised and merged to form meaningful data sets and how they could be assessed, seeking to provide a medical diagnosis based on ECG signals. The *context-awareness model* is a representation of information related to the domain environment. The contextual data provided by users and by devices are gathered and placed into predefined data structures. Special, but not exhaustive context organisation rules are applied to verify contextual information. On demand, this model aims at providing organised data for a detailed situation description. In the EA case, this model describes the information of the citizen whose cardiovascular system is being monitored: demographic data, body's morphology, skin/tissue characteristics, previous ECG recordings, clinical history, etc. It would also collect data from active devices and from their characteristics (number of recorded leads, etc). Different rules should control the integrity and the reliability of the information gathered from a citizen and from his medical devices.

A **Compositional model** is a communication component formally describing the relation between *domain-specific* and *context-awareness* models. Its main purpose is to manage the complexity of ambiguously related representations of both domain-specific and context-awareness models, serving as a formal language that supports a particular process. The compositional model focuses on data and rules specific to the process that should be obtained, clarified and, if needed, be processed to create new structures and rules or even new terminology. This model can be understood as a collection of data, rules, and terminology leading to a transparent understanding of a particular process and so enabling development of required solutions.

That is, in the EA example, where the process aims at adjusting a sensor-system used by an individual in order to improve the reliability of the recorded ECG and of the digital diagnosis performed by the system, the compositional model would consist in newly defined data structures, which identify the data relationships among the subject's morphology, his ECG signal and the sensor-system size.

Process management, measurement and improvement procedures provide a solid base for implementing process related services and composing them into business processes, which are dynamically orchestrated according to client specificities and which help to maintain the model itself. An instance of compositional model would present rules that select the path in a process in order to obtain an optimal result in a certain situation. This finite set of formalised rules would enable process tailoring while re-simulating it. The analysis of such already tailored processes opens possibilities for process measurement and improvement. The collected set of measurements and repeated improvements can lead to the construction of a new rule or additional data and could append (extend) the compositional model.

Also, the terminology of the compositional model allows describing a user profile, which includes domain-specific measurements obtained in a formally defined user situation. An aggregation of several profiles would result in a generic profile, which could be set as a reference profile to some conditions and used to improve the process, while comparing it with the user profile. This comparison would allow setting a target profile for the current situation. Then the defined target profile would first help extract these domain specific parameters, which need to be improved in the current profile to match the target profile. Afterwards, changes that must be performed in the current user situation to reach the desired state would be described using context and domain relationship rules.

6.5. Dynamic business process

The SME (subject matter expert) first creates a formal task record, representing the expected business process functionality, based on incoming information. Then the SME makes a thorough input data analysis and reviews the data accuracy in order to define the scope of the required knowledge. Therefore, the knowledge corresponding to the business process should be already grained in advanced. Finally, the principal SME's task is to assist the business process configuration, using knowledge application rules, and to validate the results correctness and completeness. These application rules must have recent activity awareness, otherwise they could be outdated.

In comparison, the activity of dynamic process (Fig.II.6_3) could be schematically explained in 5 stages: (a) definition of a business process step, (b) definition of possible tasks implementing this step, (c) tasks filtering according to the input data, contextual information and process rules, (d) tasks association to composed or atomic services that implement these tasks and, finally, (e) orchestration of these services (Fig.II.6_4), which is, again, based on process rules.

(a) In this stage we set generic steps that guide the process to the final stage. This modelling part resembles a classical business process modelling, where a process is de-

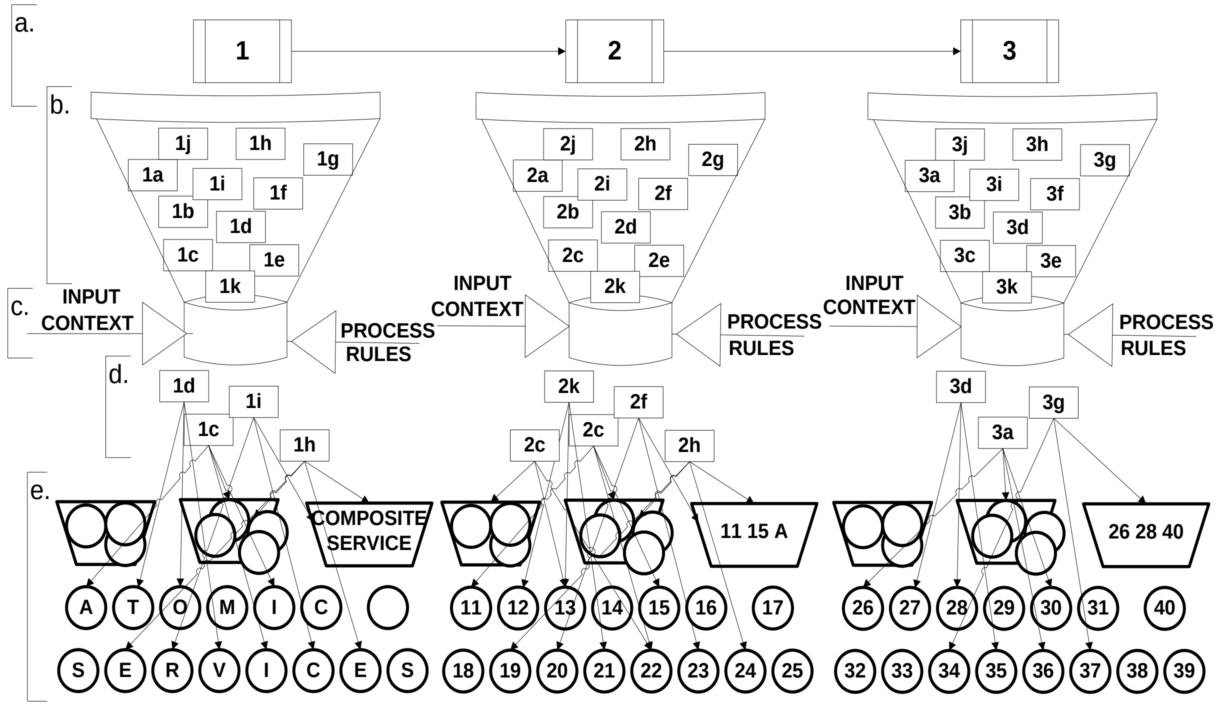


Fig.II.6_3 Schematic dynamic business process activity

fined task after task to fulfil some activity. This similarity is intentionally repeated here, just the granularity level is higher: only generic steps (therefore, it's called a step and not a task) should be defined in this stage. Specific tasks are not important at this stage as they may differ depending on context and on user specifics. For example, these steps could look as follows: data acquisition, data analysis, selection, adaptation.

(b) The second stage is the point where all tasks that are related to one step are defined. They should not be ordered, numbered or associated with each other. They shall only be expressed in terms of compositional model and be a part of a step.

(c) In this stage a user-specific workflow is defined. At each step tasks that need to be performed are selected according to process rules and the user-specific contextual information (tasks selected in the previous step, available services or some outside limitations). These tasks can already be set if possible in a sequence using process knowledge, but this sequence is still not definitive and may be changed in the final service-orchestration.

(d) At this stage the business process implementation is planned. Composite or atomic services implementing each task are selected and composed in the process.

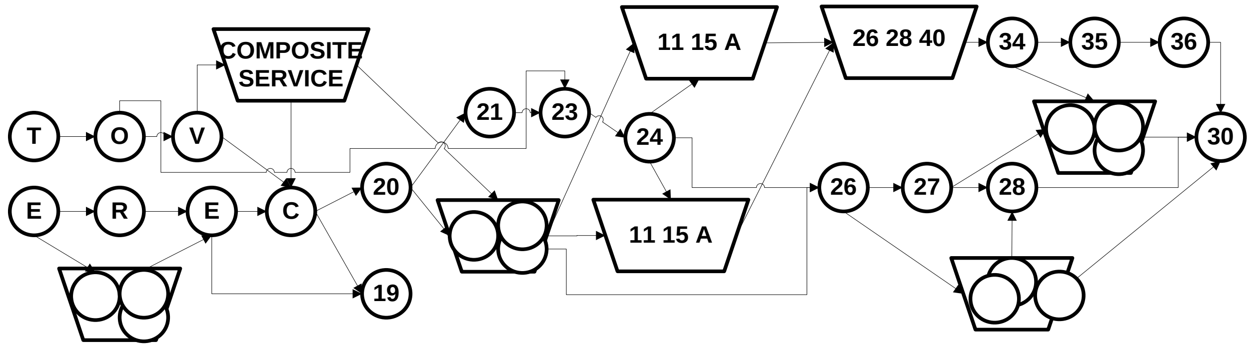


Fig.II.6_4 Symbolic representation of orchestrated atomic and composite services

(e) A final services' orchestration is performed at this stage (Fig.II.6_4). The orchestration results in a sequence of selected atomic and composed services, while step by step partially ordered tasks, process knowledge and input/output data are taken. Some exceptions: a corrupted service, inadequate or missing data could trigger the repetition of the configuration and could concern one step or even the whole process. Reconfiguration may affect the stages starting from stage c and beyond.

Following this schema, each business process intended to provide personalised assistance should have its own compositional model, which includes process knowledge. The easiest way to use this knowledge would be to model it as a concepts base and process handling rules, that is using standard knowledge modelling techniques based on ontologies. Therefore, a dynamic business process, which models expert activities, may be implemented by means of *ontology driven services' orchestration*.

6.6. Core-ontology for decision support process knowledge modelling

As a proof of concept we aim at demonstrating the relevance of this approach – dynamic business process modelling via ontology driven services' orchestration – in the cardiology domain with decision support services development and orchestration. Decision-making can be regarded as an expert activity as it is a mental process (cognitive process) resulting in an outcome leading to the selection of one among several alternatives.

However to be able to do that, we needed to develop a compositional model that enables us to model the contextual and domain specific information related to decision-making problems. Usually, even well defined and well structured expertise domains are complex and large. The related contextual information adds supplemental hesitations on its priority and final scope. This uncertainty introduces, again, a probable dependency on a human-expert and his selections, which may result in plenty of unique process knowledge models. In order to avoid this situation we propose to define a core-ontology for a group of expert activities, in our case decision-making activities, unrelated to a specific domain or a specific context. The aim is to specify a required knowledge granularity level. Such a core-ontology would embrace a set of intelligent assistants that share the same problem and could be re-used to facilitate the creation of new intelligent assistants, which thus could trigger new measurement and improvement procedures.

The core-ontology we have developed has been defined to identify and characterise the minimal required set of leading categories of concepts, which underlie the construction of a complete decision-oriented ontology dedicated to an envisaged goal in an application domain. Usually, a further development of the process related ontology requires human-expert level knowledge in the specific domain and specific information on the service receiver situation, which is typically identified as a context. To edit this ontology and the related rules, one can use standard tools, like Protégé™ *[Protégé'11]*.

Every decision-making process produces a final choice, with respect to a set of possible modalities (modes). Decision-making will, thus, be performed in function of available input data modalities, according to the data domain in general and to the specific data to be processed that may be considered as personal modalities. We thus propose two principal ontology branches (Fig.II.6_5): *data acquisition modality* and *decision support knowledge model*, which in turn contain a set of branches for decision-making expert activities automation.

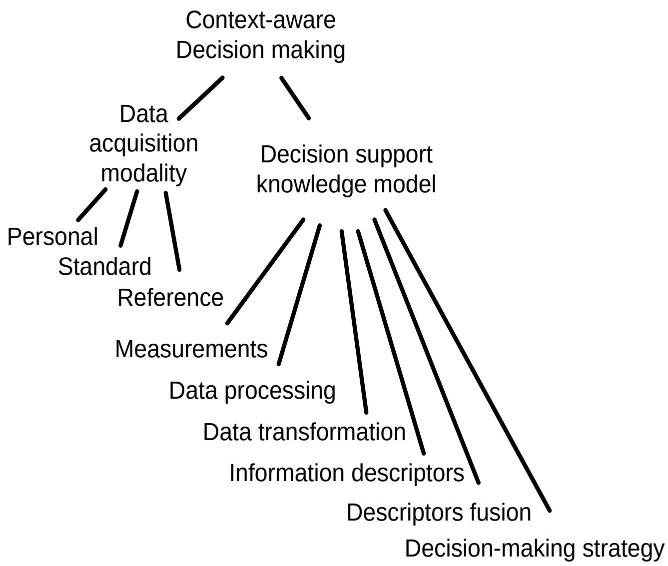


Fig.II.6_5 Core-ontology empowering decision making services orchestration

Data acquisition modalities aim at representing user related contextual information, which is extracted from a business domain using three key modalities: *personal*, *standard* and *reference*.

- *Personal* data acquisition modality is triggered by the end-user and is designed to produce user-specific data entities, such as signals, images and any others. This modality is often simplified in order to entice users to exploit it. Also, it might be used in several different contexts and can even be integrated into a common user environment. However, the content of the acquired data may be incomplete and, most likely, inaccurate.
- The *standard* data acquisition modality is commonly used by professionals. Their obtained data entities are widely accepted by domain experts as classical ones. Acquired data are normally based on well established, fundamental knowledge and there exist common practices in using various techniques for data processing and interpretation. The *standard* data acquisition modality can be set individually to a person by a professional in order to obtain a standard personal data entity.
- Finally, a *reference* modality shall be a more extended (more complete) version of the *personal* modality. Basically, it includes an accurate and rich data content, which is partly obtained using a *personal* modality. The *reference* modality represents the most adequate and accurate expected result of the *personal* modality.

Gathered data with the use of any of the three modalities may be directly input to the system by the user himself, or initialised automatically from other systems, or via sensors integrated into the user's environment.

Decision support knowledge models: The objective of a *decision support knowledge model* is to manage contextual information related to decision making activities and to supply service-developers with domain specific knowledge. The following taxonomy of six *decision support knowledge models*: *measurements*, *data processing*, *data transformation*, *information descriptors*, *descriptors fusion*, and d*ecision-making strategy* is present in this core-ontology.

Initially, all raw data sets (obtained by means of different data acquisition modalities) are loaded in the decision-making system and the activity of information retrieval is launched for each data entity. This activity is based on two knowledge models: *measurements* and *data processing*, which support the design of services, such as data filtering, pattern recognition, data correction and measurements extraction.

- Here, the *measurements knowledge model* aims at identifying each entity's characteristics, which should be computed in order to describe the entity through a functional set of measurements.
- In the meantime, the *data processing knowledge model* encloses knowledge on operational fundamentals based on data analysis algorithms. In addition, this model ensures algorithm classification into specialised groups and includes unified descriptions of the analysis methods to be used.

Afterwards, activities preparing the entities' comparison are launched in order to unify entity metrics, to identify significant comparison sets and to determine data assessment features. Here, first of all, syntactically and semantically comparable entity descriptions are prepared following the *data transformation knowledge model*.

- The *data transformation knowledge model* is designed to support the modification of data entity description in order to meet the required measurements set and, if necessary, to synthesise additional information about the entity. The model contains source and target objects definitions, transformation parameters and transformation methods of a source object into a target object. Following this, two main types of services are developed: a transformation characterisation service for repetitive transformation coefficients determination and computation, and an entity transformation service.

Then, the data assessment parameters computation is based on the *information descriptors knowledge model*.

- The *information descriptors knowledge model* aims at identifying meaningful comparison assessment parameters as descriptors and to model these descriptors' computation methods.

Decision-related comparisons are identified according to the *descriptors fusion knowledge model*.

- The *descriptors fusion knowledge model* defines decision-related data (entities, data processing or transformation methods, etc.) compositions into specially constructed comparison structures, like grids, in order to assess them according to already predefined information descriptors. The descriptors' fusion will result in aggregate compounds, like macro-data *[Fayn'99]*, from the initial information descriptors by using merging techniques. At this step, services providing assessment, descriptors computation and composition are developed.

Finally, the decision-making activity can be launched in order to evaluate comparisons and select the best or the most relevant final choice.

- The decision-making activity development is based on the *decision-making strategies knowledge model*, which aims at defining decision methods for data evaluation, and decision strategies, such as scoring or trend-analysis for the final choice selection. Then it might also model recommendations for the end-user according to the performed selection.

The main strength of this core-ontology lies in its ability to encapsulate the business process related knowledge in a global compositional model. This thoroughly designed knowledge then enables the construction of specialised context handling rules coherent with a business process. Only a complete service related knowledge model can ensure the rules correctness. The latter are the principal driving forces for the personalised configuration of dynamic decision-making business processes. These rules will be executed by an engine which implements two principal mechanisms: dynamic services' orchestration and dynamic services' composition. Hence, according to the business process execution circumstances, a selection of goal-oriented business process activities is first performed. Then, the services that may implement the required business process activity with the given input information or with the expected intermediate data are selected. This allows a complete services' readjustment directed by the contextual information and, thus, a dynamic business process development.

6.7. Discussion

Procedures of current professional services, characterised by human-to-human contacts, are likely to change with the expected development of new services automating expert-level tasks. Even so the role of a domain expert should not disappear, as some could have feared. On the contrary, it is supposed to be permeated with more credibility. These services will bring the possibility to describe expert knowledge in an expressive way, to measure and even to evaluate its preciseness. Automated personal care services would take a double role, on one hand, allow professionals to automate complex and user related tasks in their expertise domain and, on the other hand, enable non-expert users an autonomous use of personalised professional services in this domain. Current domain-experts' role to diagnose, to advise or to teach the client will be slightly shifted to a knowledge engineering role, that is, they will become principal modellers of these services. They will design dynamic business processes and will be responsible for their maintenance. Knowledge engineers will perform regular system usability analysis in order to improve system performance. Also, frequent knowledge exploitation updates are likely to happen. This all together will enable improvement procedures for a knowledge application, as measuring is always a starting point for any enhancement.

To contribute to this vision and to trigger the whole process, we designed a core-ontology for decision-making to capture and to represent knowledge into specially defined knowledge models, which then facilitate composite services creation and their dynamic orchestration. The goal of our proposed core-ontology is to permit professionals to automate decision support tasks in their expertise domain. We firmly believe that the service cluster implementation and management based on a context-aware business process knowledge model facilitates the automation of dynamic business processes and expands service applications to new domains where complex solutions are required, or enhances current service adaptability to a user while avoiding loading an extra knowledge charge on these users. Our approach and the obtained results confirm the widely accepted importance of context-aware modelling. Though, instead of modelling separately the domain knowledge and its context, we propose to model process knowledge on the basis of context-aware domain knowledge. In this approach, the process knowledge is formally expressed in an ontology and a rules engine that provides a solid base for implementing process related services and to compose them into a business process which is dynamically orchestrated according to the client specificities. We conceptually demonstrated that the intelligent assistance could be grounded with a knowledge model upon a specific business process. Subsequently, core-ontologies of such business processes might be reusable in other domains.

To our knowledge, this is a new approach and it goes beyond the previous approaches as it integrates different aspects taken from several domains by using what fits best to solve the present, though, often bypassed problem. However, tools implementing these concepts and supporting real expert activity automation are still awaited. The first of them would be a platform to model dynamic business processes and an engine which can execute rules from the compositional knowledge models. The creation of such a platform might awake another problem, related to new terminology. Currently, the meaning of many terms, for example expert activity, dynamic process, personalised system, adaptive system is somewhat confusing. A clear distinction between words used for marketing purposes and for real problems need to be put forward at least for professionals, which would make these problems less mystic and would encourage their solution. Then, developing some of these assistants even related to one problem would be a big step forward as their comparison would allow creating evidence-based perspectives and would contribute in improving our approach. However, the next big issue that we see mostly missing is the design of an architecture, which on intelligent assistant based systems shall be created. This architecture should take into account the complex relationships between context representation models and the domain specific models and should support problem specific models. Consequently, it should satisfy adequate user requirements for personalisation. This problem we investigated in depth and the results that we have obtained are presented in the following chapter.

Chapter 7 – Data and Model driven Service Oriented Architecture (DM-SOA)

7.1. Introduction 137
7.2. Engineering decisions 137
7.3. General framework 138
7.3.1. Global architecture 139
7.3.2. Business logic layer 140
7.4. Generic DM-SOA for decision support systems 141
7.5. Domain-specific DM-SOA for decision support system 147

7.1. Introduction

Like any other complex structure, software should be built on a solid foundation in order to keep system's performance, security and manageability quality attributes. This fundamental basis, known as system architecture, includes a scope of the system, its behaviour describing system components and their relationships with each other, main working principles that meet functional and non-functional system requirements, finally, system constraints if such exist.

In this chapter we propose a new architecture, which aims at supporting the development of an intelligent assistant system. Such a system shall deliver personalised professional care services to any citizen and, thus, replaces a human expert, so enhancing the accessibility of a professional care at anytime and anyplace. The architecture is based on principals from model-driven architecture (MDA) and service-oriented architecture (SOA) and employs our proposed dynamic business process management approach that we discussed in chapter 6. The presentation of this architecture consists in three levels of particularity:

1) a general framework, suitable to develop any intelligent assistant;
2) a generic architecture for decision support services development;
3) a domain-specific architecture for optimal personal ECG recording system selection.

Our proposed architecture is intended to be used as a tool to design the intelligent assistant system or to estimate and if possible to reduce its development cost.

In the scope of this thesis we haven't included a technical view of the architecture, which compounds the deployment of the system as well as technical/hardware components. We project that the intelligent assistant system might be deployed in a variety of environments and distributed or non-distributed system patterns may be used while implementing this type of systems. The issues of a physical components separation across different servers, networking protocols, firewalls, router configurations and others are specific to a concrete implementation and should be defined in the technical view of a final architecture of a specific intelligent assistant.

7.2. Engineering decisions

Typically, a software architecture embraces a series of decisions, where each of them has a considerable impact on the quality and overall success of the system. In our architecture we have considered four principles as architecture cornerstones:

(1) service-oriented: loosely coupled, reusable, composable components
(2) model-driven: platform independent, models-based domain engineering
(3) automatically manageable: process knowledge based agents
(4) reconfigurable: dynamic business processes deployment, and data-driven adaptive services orchestration.

We have chosen to use service oriented architecture (SOA) as an architectural background for our intelligent assistant system. This architecture addresses the complexity, inflexibility, and weaknesses of existing approaches to business process design and workflow management. It promotes that software processes need to be broken down into services which are then made available and discoverable on a network. A service there stands for a self-contained piece of business functionality and from the perspective of a system it is seen as a black box, which communicates only via a predefined interface. To ensure this communication, a specific infrastructure based on messages is created and it ensures services high interoperability. The SOA approach accepts heterogeneity and leads to decentralisation so introducing an ability to decouple the application architecture from the infrastructure components. This SOA taken course of a loosely coupled, therefore flexible, scalable and with a greater fault tolerance system is a promising method aiming at business processes' automation support.

Model-driven architecture principles, in particular, address the problem of the increased data and methods heterogeneity, which is a key problem of complex systems, such as life science systems. A model-driven architecture implicitly uses the principle of divide-and-rule, which helps to give a structure to any given problem. In our case this results in knowledge granulating in separate models, in which things can be done relatively easier. Functional system requirements are covered with a conceptual design of these models so ensuring the mapping of appropriate domain-specific concepts with the user context. This conceptualisation gives platform independent models and makes sure that the design could survive modifications caused by changing technologies during its realisation.

Then, in chapter 6 we have demonstrated an approach of an intelligent personal assistant implementation via an automatically manageable dynamic business process. We will apply this key approach in our architecture by creating the management component, which overwhelms process-related knowledge. Process knowledge is a collection of contextual and domain-specific data and rules that are specific to the process and that enable process management agents to automatically guide the process towards an optimal solution to the concerned user.

Finally, we concluded that the personalisation of an intelligent assistant shall be grounded with a dynamic domain-specific knowledge application to continuously supplied heterogeneous user data. In chapter 6 we demonstrated that this can be implemented using an ontology-based services orchestration, that is a settlement of various services capable to perform the needed computations and that is automatically reconfigured by means of the created ontology and orchestrated in order to implement a user specific workflow provided by the instance of the dynamic business process. The achieved business service is a distinct business facility optimised to a specific system user.

In the following sections we will show how these key engineering decisions ensure the construction of an acceptable system, which at the same time is enough flexible to be less likely broken by changes while it is evolving.

7.3. General framework

This section provides a high level overview of the architecture we propose for the intelligent assistant system development (Fig.II.7_1), which is data and model driven SOA. The architecture at this level outlines the strategy of layered system responsibilities, where each layer is associated with a particular system development part. This approach helps to isolate system development tasks one from another, so facilitating the creation of the system as the focus can be set on several smaller parts rather than on a whole complex system and even speeding the development process as tasks can be implemented in parallel. The presented framework allows establishing criteria for system design, modelling of all the related specifications that define the domain-specific problem, and helps developing a system and its content in terms of messages and services.

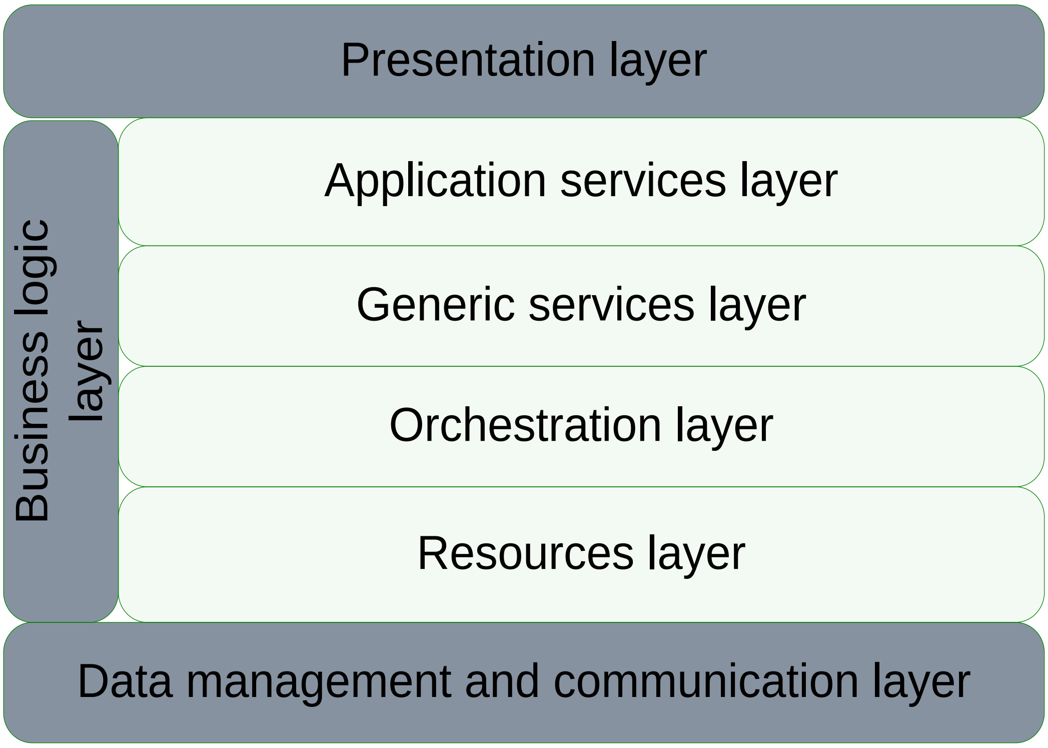


Fig.II.7_1 Data and model driven Service Oriented Architecture

7.3.1. Global architecture

The general framework is based on the usual three-layer architecture, which is the most widely spread use of N-tier client-server architecture. Hereof, our architecture is an aggregation of the three typical layers:

1) presentation layer,
2) middle (business) layer,
3) data management and communication layer.

The **presentation layer** deals with the presentation logic and contains all the components needed to allow system interactions with an end-user.

The **middle layer** incorporates the level of **business logic**. It contains the implemented functionalities of the application. Any programming language can be employed to realise this level's functionalities by coding the behaviour of the whole application or only some services. In our architecture, this layer consists in four sub-layers as follows: the application services layer, the generic services layer, the orchestration layer, and the resources layer. The communication among these layers is coordinated in the orchestration layer. The orchestration layer is controlling the interchange via business rules taken from the resources layer in order to instantiate a dynamic business process via composite and atomic services to provide a user-specific workflow.

The **data management and communication layer** encloses data storage and the logic necessary to access and to manage data. It ensures data consistency and validity. Interfaces containing specific operations without exposing details concerning the database itself are provided to access the data. A specific approach for distributed data storage and data structures administration is set in this layer. It provides developers with the ability to support multiple database servers by employing a class of functions that are always the same no matter which database is accessed. Typically, data can be stored in any type of storage, starting from classical relational data store passing to some hybrids stores and going even to storage in a cloud.

All three layers can be developed and maintained as independent components and most often on separate platforms. However, further on, we set our focus in this thesis only on the middle – business logic – level, as this level contains the core of a solution that allows the intelligent assistance development.

7.3.2. Business logic layer

The business logic layer is composed of four presented layers aimed at integrating the functionalities that are required for an implementation of an intelligent personal assistant. Each part helps to easier control a specific aspect of the intelligent assistance system and to clearer structure the final solution. A number of important decisions must be addressed in all layers to meet the design and development of dynamic business processes. We then describe each layer and will stress their key responsibilities.

The **Application services layer** provides an infrastructure to use the functionalities proposed by various atomic services already built into the system or proposed by third parties. Each service implements some part of a business functionality. This layer holds contracts to the operations provided by services, the messages required to interact with each operation, and the patterns for messages exchange. A service contract specifies the behaviour of a service and the messages required for the interaction with this service. The interaction part implements the translation between data outside the service and data within the service. Individual services are described using typical techniques used in service oriented architectures: WSDL, XSD, WS-Policy, WS-MetadataExchange, and related specifications.

The **Generic services layer** serves to create composed services that respond to the complex user needs. This layer builds interfaces that ensure loosely coupling. Reusability is a major benefit of this layer. Existing orchestrated services are set to a composed service and so could be reused in constructing higher-lever user-specific workflows. Composed services (when accomplished) are exposed to the process via the service interface so that they can be found and included into the business process.

The **Orchestration layer** is responsible for handling the context-awareness of the system. It provides the final user-specific business service for the system user. This layer consists in *business process management techniques* and in a *business rules' engine* that compounds rules to guide a concrete workflow by invoking the required services. The workflow is driven by applying ontologies and rule models, which shall be instantiated according to the resources existing in the resources layer. The ability to invoke services in an asynchronous manner is vital in achieving the required reliability, scalability, and adaptability in this layer. Using asynchronous services allows invoking services concurrently rather than sequentially and so enhances business process performance. Orchestrated services manage exceptions and transactional integrity, when an error happens or when invoked service does not respond in a given time. Transaction integrity management is taken in account in this layer: resources cannot be locked in a transaction that runs over a long period of time. Finally, services orchestration is dynamic, flexible, and adaptable to meet the changing business needs. Flexibility is achieved by providing a clear separation between the process logic and used services. A business rules engine is taking the role to map the latter two and ensures automated system adaptability at runtime. It handles the overall process flow, calls the appropriate services and determines the next step in order to complete a given task. Like that, services can be swapped in the process flow. Orchestration is based on knowledge models and contextual data and can perform two types of services management: process driven and event driven.

The **Resources layer** supports the system with in advance *defined descriptions of data acquisition* and with *knowledge models* containing domain-specific knowledge related to the problem that the system is designed to solve. This layer is based on the principals of model-driven architecture. The knowledge models here should be of a right dimension in order to be easily controlled and used when the business rules engine or the related business services are created. Knowledge engineers should take in account the nature of what is feasible in the model and its size. In systems that use a great number of knowledge models, the models hierarchy might be applied to increase the transparency of the system and to help to manage the knowledge. Multiple loosely coupled models can represent one large model. In the future vision this layer may also provide interfaces for domain experts or expert groups to model and manage their knowledge related to the specific domain.

7.3.3. Conclusion

The main advantage of our proposed architecture is the possibility to use data and knowledge models in order to enhance the typical SOA orchestration and to address the personalisation issues. Such an enhancement in the orchestration level enables to take contextual data in account at system runtime and to set a workflow that is compliant with both the specific environment of the user and with the domain-related data processing and analysis limitations. Hence, the system is capable to deploy the dynamic business process and so to provide a service of an intelligent assistant. Hereof, we named our architecture a Data and Model driven Service Oriented Architecture – DM-SOA. The goal of DM-SOA is to make business applications more intelligent and more responsive to the user. Applications based on DM-SOA then should be finally capable to put into practice automatically personalised complex business processes.

As a proof of concept we want to apply this framework to develop an intelligent assistant that helps a non-professional user to choose a sensor-system modality size in order to record a medical quality ECG. This is a typical decision support system that needs to take in account user specifics and complex domain-related knowledge in order to process user-specific input data and to select the most suitable sensor-system modality. To be able to model such a system, we first of all will specify the general framework of DM-SOA to a generic DM-SOA for decision support systems in section 7.4.

7.4. Generic DM-SOA for decision support systems

Decision support system modelling is a typical example of a multidisciplinary, expert knowledge requiring and situation-aware demand. The decision-making applications based on DM-SOA for decision support systems are intended to assist a user by facilitating its decision in a domain where the user has not the required knowledge to process his personal data or personal preferences in order to identify the most suitable modality of the needed object. The typical example on which we ground this thesis is taken from the cardiology domain. However, we could list other examples for an application of DM-SOA for decision support from almost all domains, such as:

E-consulting: a client wants to buy a new car, but he has not a profound understanding about cars, different engine types, brands, price/quality ratio and others. The same approach is valid for selecting a new mobile phone or a new personal computer.

E-learning: a pupil wants to get knowledge on some specific subject and we want to adapt the training to his profile (current knowledge, the speed of learning, type of memory and others).

The potential solution for all these decision support systems is laid in the enhancement of the orchestration, which manages services that encapsulate business functionalities so that they can be automatically reconfigured at runtime according to the user context. To support this type of orchestration, we constructed an ontology modelling core-concepts of decision-making (Fig.II.7_2). This generic ontology facilitates the initial modelling of business process management of any intelligent assistant with a decision support task. These concepts are used to unify services interfaces, which will then allow invisible services' swapping and reconfiguring in order to meet the conditions imposed by changing contextual information.

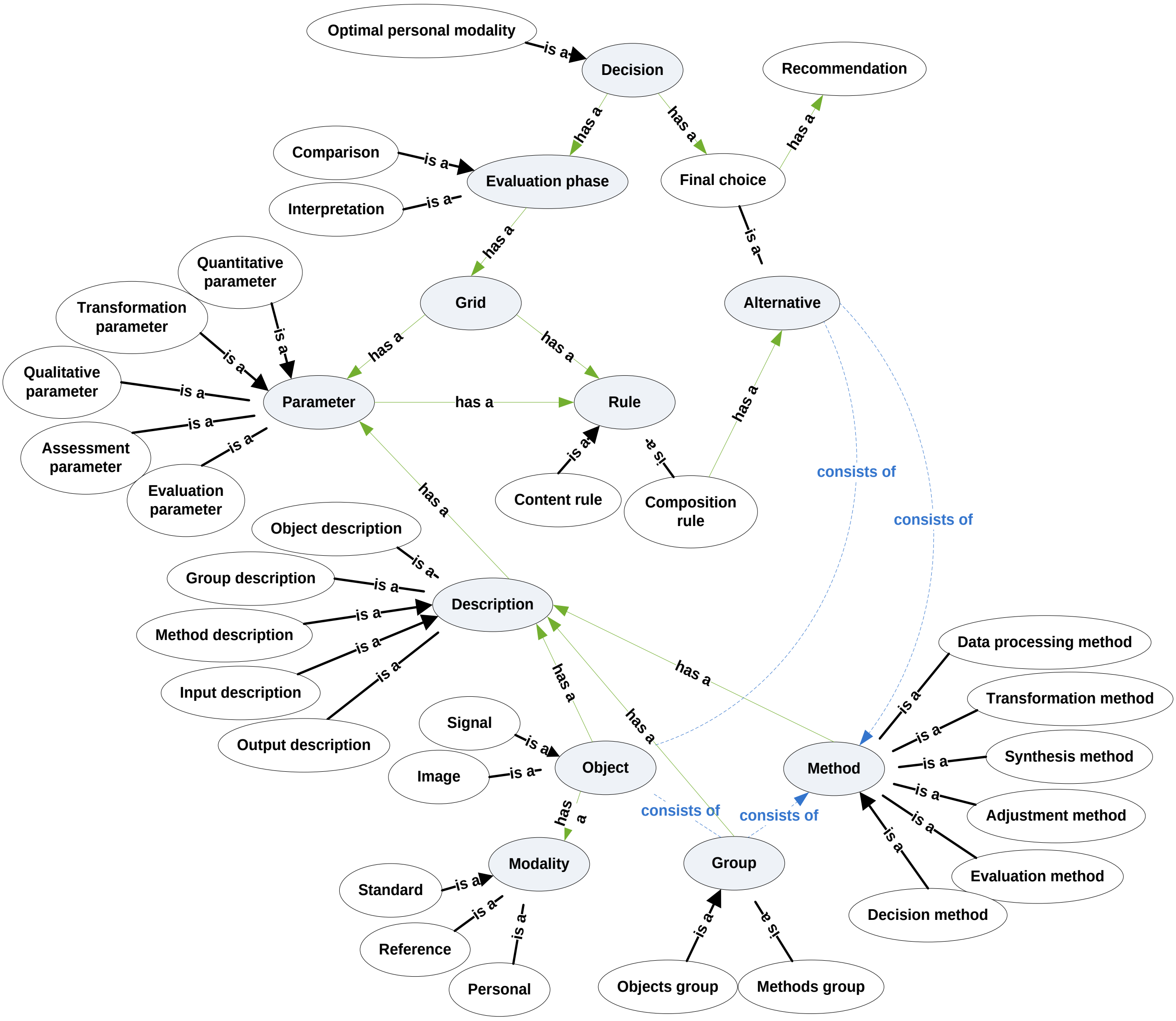


Fig.II.7_2 Decision support system process knowledge ontology empowering the management of dynamic decision-making business processes

The ontology-driven services orchestration approach raises specific requirements to the design of decision problem related knowledge models. Knowledge models should be constructed not only to compound the related domain and its context fundamentals, but also should be generalised via common definitions and should have identified relationships that would allow creating process management knowledge. To define the necessary knowledge models we took the core-ontology for decision support process knowledge modelling that we presented in section 6.6 (see Fig.II.6_5) and we described in terms of the generic ontology six knowledge models (Fig.II.7_3) containing business domain knowledge: *measurements*, *data processing*, *data transformation*, *information descriptors*, *descriptors fusion* and *decision-making strategy* models. These models are intended to ground not only the management of dynamic business processes, but also the implementation of atomic services, which represent pieces of business functionalities.

Measurements knowledge model: *What is needed to be measured?* This model is set to determine entity's characteristics which shall be measured in order to answer a given problem. A unified description of each entity group is settled.

Data processing knowledge model: *How to measure?* This model defines methods which serve to measure entity's characteristics. It classifies suitable data processing algorithms into groups and models each algorithm's operational fundamentals. A unified methods' description containing parameters such as execution time, preciseness and others is also assigned to each data processing algorithm group. Also static input and output definitions are determined for each data processing algorithms' group following which each algorithm in the group adapts their input data set and then provided results' set. These definitions then can be used in service contracts.

Data transformation knowledge model: *How to obtain a relevant set of measurements?* This model is intended to map the entity's description to the required parameters set and if necessary to synthesise additional information about the entity. The model contains definitions of data objects, transformation parameters and transformation methods which are used to transform and synthesise a source data object to a target data object. Knowledge in this model is mainly applied for a dual purpose: either to obtain transformation parameters for a repetitive use, or to obtain a transformed data entity.

Information descriptors knowledge model: *Which are important direct and/or derived and/or composite measurements necessary to make a decision?* This model forms a meaningful parameters set (or sets), which expresses decision-related knowledge. These descriptors can contain simply adjusted entity characteristics or derived and then if needed adjusted characteristics. Adjustment and classification methods are also designed in this model.

Descriptors fusion knowledge model: *How are decision-related possibilities represented in a pointful expression (usually digital)?* The goal is to define potential alternatives, which can be described also via new data structures. Each alternative represents a composition involving one or several entities evaluated using descriptors sets.

Decision-making strategy knowledge model: *How is the automatic recommendation produced for the system user determined?* This last model designs evaluation parts and possible decision strategies. It compounds alternatives' evaluation methods, which are identified following common interoperability support principles. Unified structures, describing a static input and a static output for these methods are defined. The model also defines possible decision-making strategies to select among several already evaluated alternatives. For instance, weighted decision rules, fuzzy logic, case-based reasoning, and trend analysis can be used to get a final choice and consequently to provide user-specific recommendations.

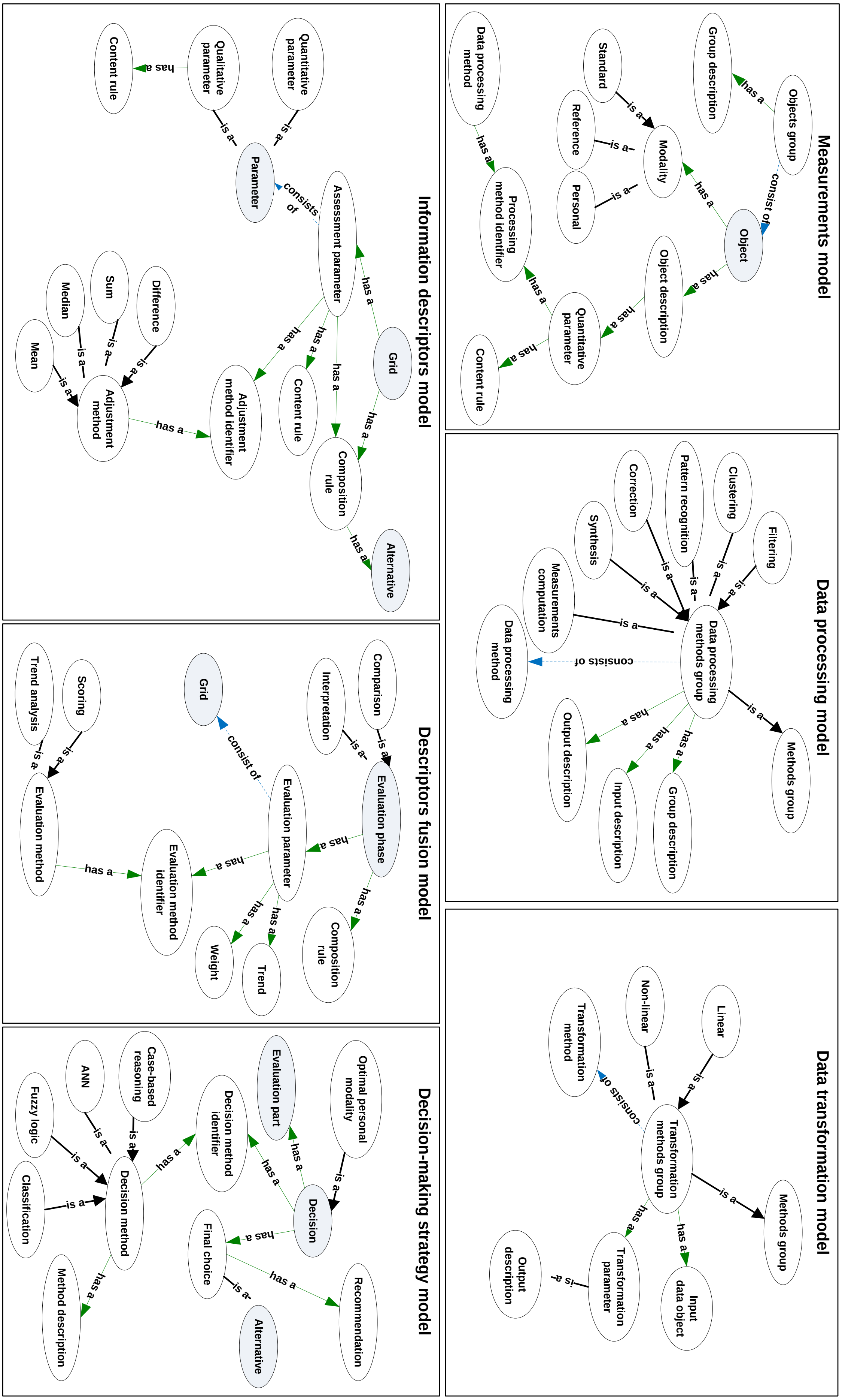


Fig.II.7_3 Schematic ontologies for each knowledge model

Having described the different decision-making supporting knowledge models, we present hereafter, layer by layer, the DM-SOA we propose for decision support systems development (Fig.II.7_4). This DM-SOA is based on the general architecture model displayed in Fig.II.7_1.

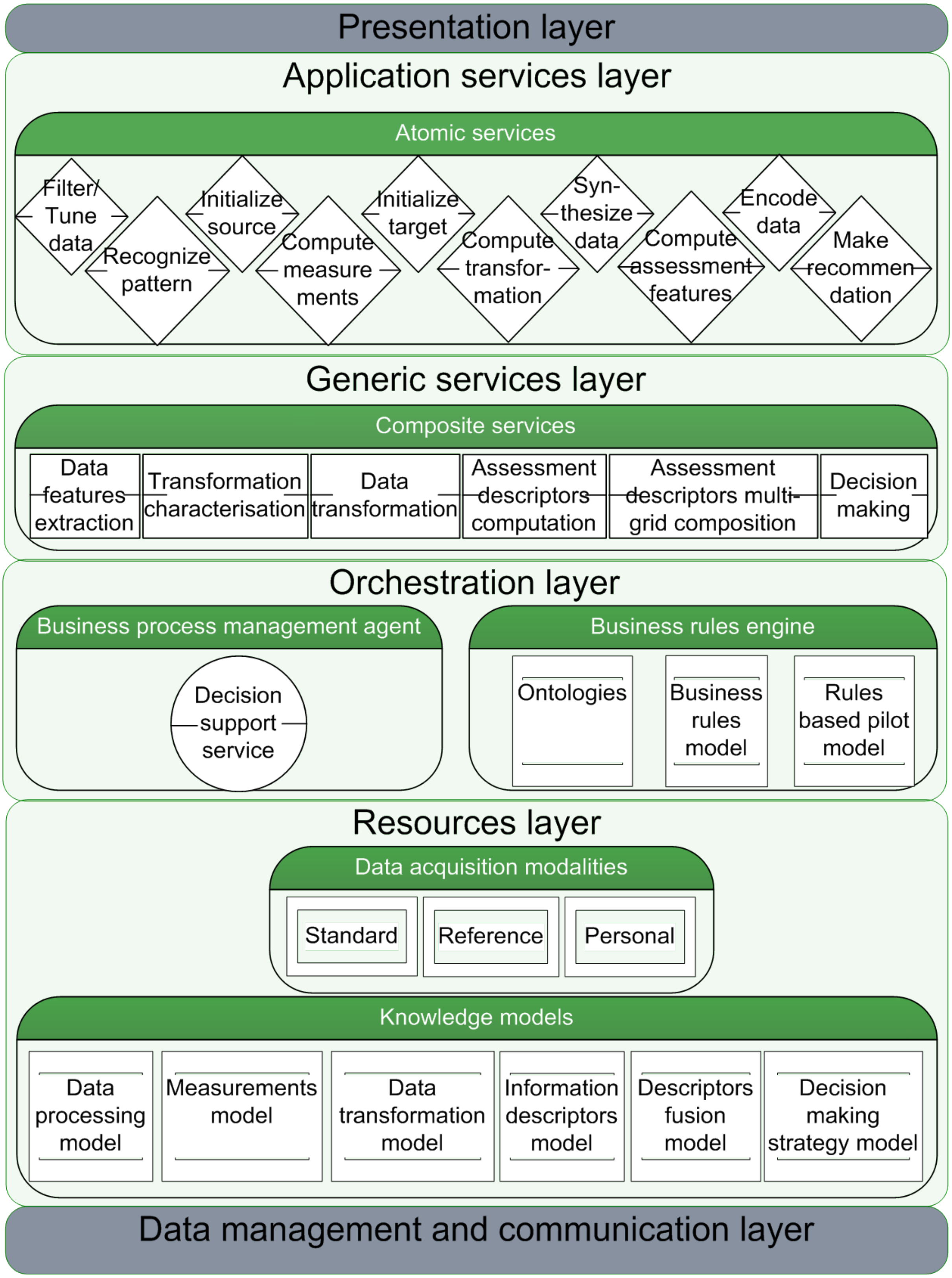


Fig.II.7_4 Data and model driven service oriented architecture for decision-making support

The **Application services layer** contains typical data processing and data analysis atomic services, such as: *filter data*, *tune data*, *recognise pattern*, *compute measurements*, *initialise source*, *initialise target*, *compute transformation*, *synthesise data*, *compute assessment features*, *encode data*, *make recommendation*.

The **Generic services layer** is specified into six composite services, which represents classical stages of decision-making: *data features extraction*, *transformation characterisation*, *data transformation*, *assessment descriptors computation*, *assessment descriptors multi-grid composition*, and *decision-making*. As in the next chapter we will give an example of these services orchestration we do not describe them here in details.

The **Orchestration layer** is composed of two main components: the *business rules engine* and the *business process management agent*. The *business rules engine* component includes:

1) a business rules model, which explains how rules for this system shall be constructed,
2) a rules based pilot model, which explains how the business rules engine is functioning,
3) the process knowledge ontology (ontologies).

The *business process management agent* extracts business rules according to the predefined model. These business rules are built using concepts defined in the process knowledge ontology and are targeted at managing a dynamic business process. The agent uses the rules to instantiate the dynamic business process by selecting available and suitable services in order to set-up a user-specific workflow.

The **Resources layer** is based on the core-ontology proposed in section 6.6 and consists in three data acquisition modalities and six knowledge models. We have already specified the knowledge models using the generic decision-making ontology (Fig.II.7_3). Here we only list the three proposed data acquisition modalities: *standard*, *personal*, and *reference*.

- The **Standard modality** allows obtaining data entities which are commonly used in the domain and widely accepted by professionals.
- The **Personal modality** allows obtaining simplified data entities which are operated by end users.
- The **Reference modality** allows obtaining data entities which take an intermediate position. These entities contain characteristics from the entities described in the standard modality as well as from entities described in the personal modality. Entities obtained with a reference modality can often be seen as extended entities, which were obtained with the personal modality. A reference modality could be seen as a target for the personal modality.

At this level of the DM-SOA conceptualisation we prepared the ground for an intelligent assistant aiming at decision support. In the following section we specify the final architecture of a domain-specific DM-SOA for a decision support system that is intended to optimise a personal ECG recording system selection.

7.5. Domain-specific DM-SOA for decision support system

Conclusively, we present the domain-specific DM-SOA (Fig.II.7_5) specified for a concrete decision-making demand, which stands for a selection of an optimal personal ECG recording modality that is one concrete business service of the Personal Cardio Assistant system (PCAs). This business service displays of how an intelligent assistant implements a dynamic business process, which adapts to user circumstances, his morphology and takes in

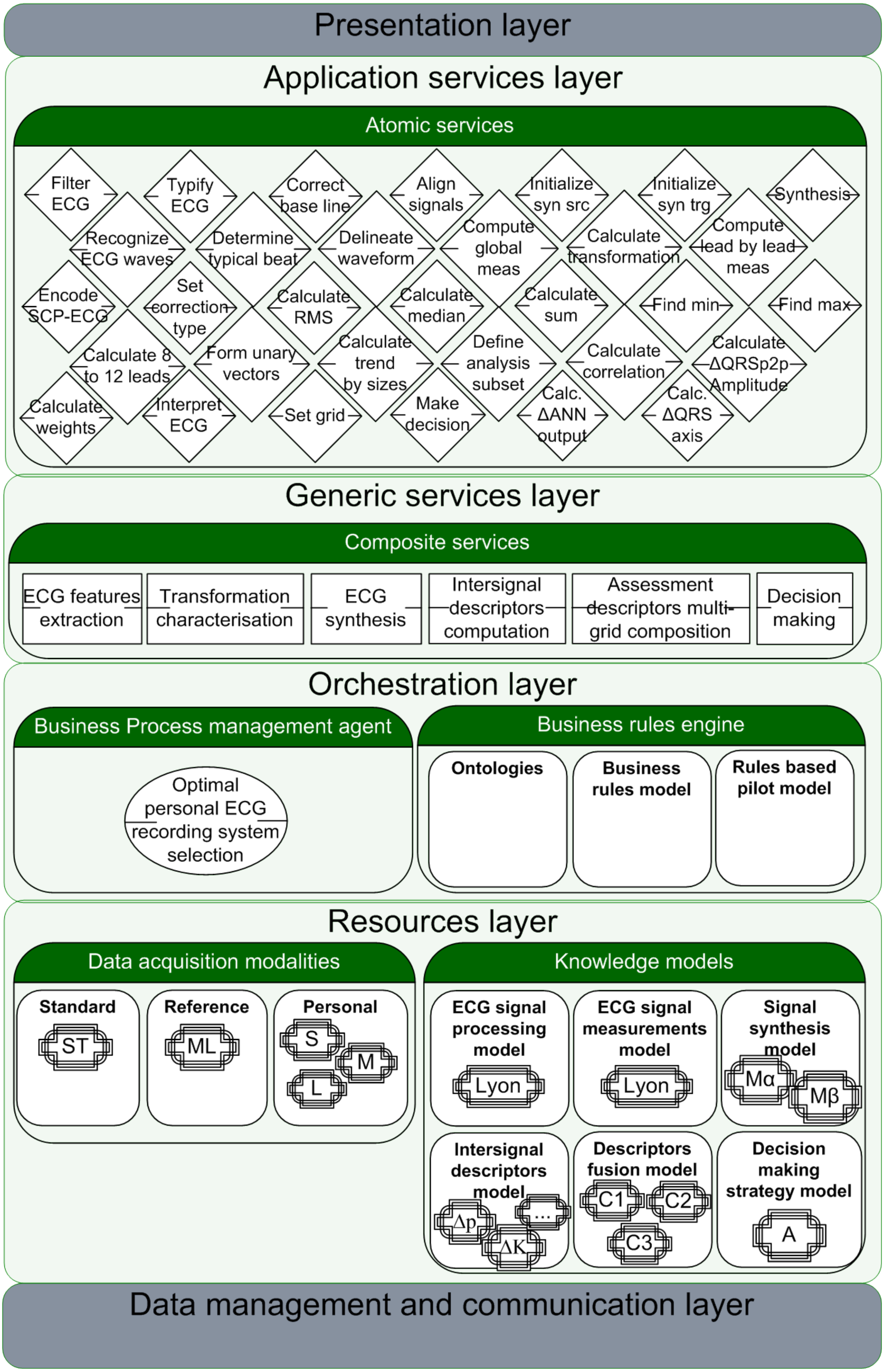


Fig.II.7_5 Domain-specific data and model driven service oriented architecture (DM-SOA) account the clinical history of the user. This concrete architecture is extendable to contain all other services of the PCAs.

The **Application services layer** compounds algorithms and methods that are used in the electrical heart signal processing domain. It is an explicitly rich domain proposing a variety of different techniques already presented in section 2.5. Fig.II.7_5 displays the corresponding atomic services. How these services are composed and which functionality they should implement we demonstrate in chapter 8 in the orchestration of these services.

The **Generic services layer** compounds six composite services that are based on composite services from a generic DM-SOA for decision support systems. These composite services include: *ECG features extraction*, *transformation characterisation*, *ECG synthesis*, *intersignal descriptors computation*, *assessment descriptors multi-grid composition*, and *decision-making*. These services orchestration is presented in chapter 8.

The **Orchestration layer** is set with a *business process management agent*, which targets to select a personal ECG recording modality that provides the best medical quality ECG signal for this user out of the given 3 sensor-system modalities: S, M and L (defined in chapter 4). This agent takes in account a set of *business rules* and uses available contextual information: the available types of recordings, methods that have to be used to process the signal, each type of signal recording, possible comparison techniques and others in order to end with an optimal, user-specific choice. We demonstrate how this orchestration is performed in chapter 8.

The **Resources layer** consists in 5 data acquisition modalities and 6 knowledge models:

Standard modality: ST (a standard 12-lead ECG)

Reference modality: ML (a 12-lead ECG obtained using Mason-Likar electrodes' positioning)

Personal modalities: S, M, L sizes (3-lead ECGs).

ECG signal processing model: the ECG signal typically passes at least six phases – signal filtering, ECG beats detection, typification, determination of the median beat, waveforms delineation and base line correction – in order to be pre-processed. Each of these phases contains specific heuristic knowledge, which may result in a variety of algorithms for one phase (therefore several services containing different heuristics can be implemented). This different logic shall not be lost. It shall be transparent and modelled as process knowledge that would help to ensure these algorithms compatibility. We used algorithms that were developed by the Lyon team and set into the Lyon program in order to keep this compatibility.

ECG signal measurements model: the ECG signal has global measurements, and specific lead-by-lead measurements. The Lyon program is compatible with the SCP-ECG protocol and provides 13 global measurements and 312 lead by lead measurements.

Signal synthesis model: data transformation model of a 3-lead ECG into a 12-lead ECG. We used 2 signal synthesis strategies: $\mathcal{M}_\alpha$, $\mathcal{M}_\beta$ that we explained in section 4.6. However, the signal synthesis model may be extended to other signal synthesis strategies.

Intesignal descriptors model: our defined model contains 7 intersignal descriptors defined in section 4.6 (see table I.4.3).

Descriptors fusion model: we set three possible descriptors fusion strategies: C1, C2 and C3 that are described in section 4.6. The user context defines how many of them and how many times each of them will be performed.

Decision making strategy model: we have proposed a decision strategy for data evaluation. However, as it should be tested on large clinical data sets, other decision strategies could be used in parallel in order to obtain a final recommendation.

The final set of operations that are performed while using these knowledge models is demonstrated within the services orchestration description in detail in chapter 8.

This conceptual architecture for a personal cardio assistant system (PCAs) is capable to deliver a personalised recommendation on the basis of heterogeneous data continuously provided by the citizen. All architectural layers complement each other and as a whole enable the development and implementation of the dynamic business service, which we explain further on in the next chapter.

The main purpose of our proposed conceptual DM-SOA was to draw the attention at an appropriate decomposition of the intelligent assistant system. We identified key concepts and showed the required components. With this architecture we did however not propose nor specify the underlying technologies. Indeed, the technological aspects depend on the system development environment and only then can be clearly articulated.

Chapter 8 – Enhanced services orchestration

8.1. Introduction 153

8.2. Use-case scenarios 153

8.2.1. Anomaly suspicion scenario 153

8.2.2. Follow-up scenario 156

8.2.3. Emergency scenario 158

8.2.4. PCAs general conceptual model 158

8.3. Optimal personal ECG recording system selection: Service Choreography 159

8.4. Optimal personal ECG recording system selection: Service Orchestration 161

8.4.1. ECG features extraction 162

8.4.2. Transformation characterisation 163

8.4.3. ECG synthesis 165

8.4.4. Intersignal descriptors computation 166

8.4.5. Asssessment descriptors multi-grid composition 166

8.4.6. Decision-making 168

8.5. Conclusion 170

8.1. Introduction

In this final chapter on intelligent assistant we aim at demonstrating a set of possible scenarios for a PCAs (Personal Cardio Assistant system), which is a typical example of systems for creating ambient intelligence, and at showing an example of realisation of one dynamic business process, an "Optimal personal ECG recording system selection", in terms of enhanced services orchestration. This dynamic business process aims at helping non-professional users to record in ubiquitous environments their electrical heart signal, which has the best diagnostic quality. In order to understand how this representative service operates in the global context of PCAs, we first will describe three scenarios: anomaly suspicion, follow-up and emergency. These use cases shall also serve for a validation of our proposed approach and architecture (in chapters 6 and 7). More specifically, we want to demonstrate how all architectural components function in order to enable the execution of a dynamic business process, which allows system personalisation at runtime.

8.2. Use-case scenarios

Situations where citizens that felt no clear pain in the chest area are likely to delay the visit to their general practitioner in order to do physical examination, are happening every day. These situations are dangerous and lead to increased mortality and partial or complete impairment. They mostly appear due to limitations of the currently proposed medical services, which have only two options: a medical visit or a waiting for clearer clinical signs. To fill this existing gap and to improve the current medical practice we propose PCAs, which could better meet the lifestyle of our rushing and stressful society. The PCAs aims at providing automated expert activities and at enabling citizens' self-care anywhere and at anytime. Further on, we have defined three healthcare scenarios that are based on automated services and could be used in the prevention or the follow-up of cardiovascular diseases.

8.2.1. Anomaly suspicion scenario (Fig.II.8_1)

This scenario is intended for any adult, who suspects having a disorder of his cardiovascular system. To do so, a citizen shall acquire a sensor-system at some specialised point-of-sale, for instance in a pharmacy. The pharmacist helps him to select an optimal sensor-system modality adapted to the person's morphology in order to record ECGs suitable for digital diagnosis. The citizen connects his sensor-system modality to his chosen monitoring device (a separate device, a mobile phone, or a personal computer) and personalises this device by creating his personal profile, which contains his demographic data and medical information. After the system is set-up, the citizen can record his personal ECG whenever he wants and get an assessment of his cardiovascular system.

This scenario includes the PCA system and implies two actors: a person himself and a pharmacist or any other certified person. The main goal of this PCAs is to evaluate person's heart status and to provide a user-specific recommendation. On the basis of this recommendation, the user can then make his decision of how to proceed with his healthcare. He can contact a doctor or can continue to follow up his health state based on the ECG signal analysis performed by the PCAs. This scenario proposes 5 use-cases that can be operated by actors and includes 6 tentative use-cases that can be included in the system in order to perform the required tasks. Let us review them in the following.

Create and modify personal profile (Create/Modify pProfile): This use case helps the

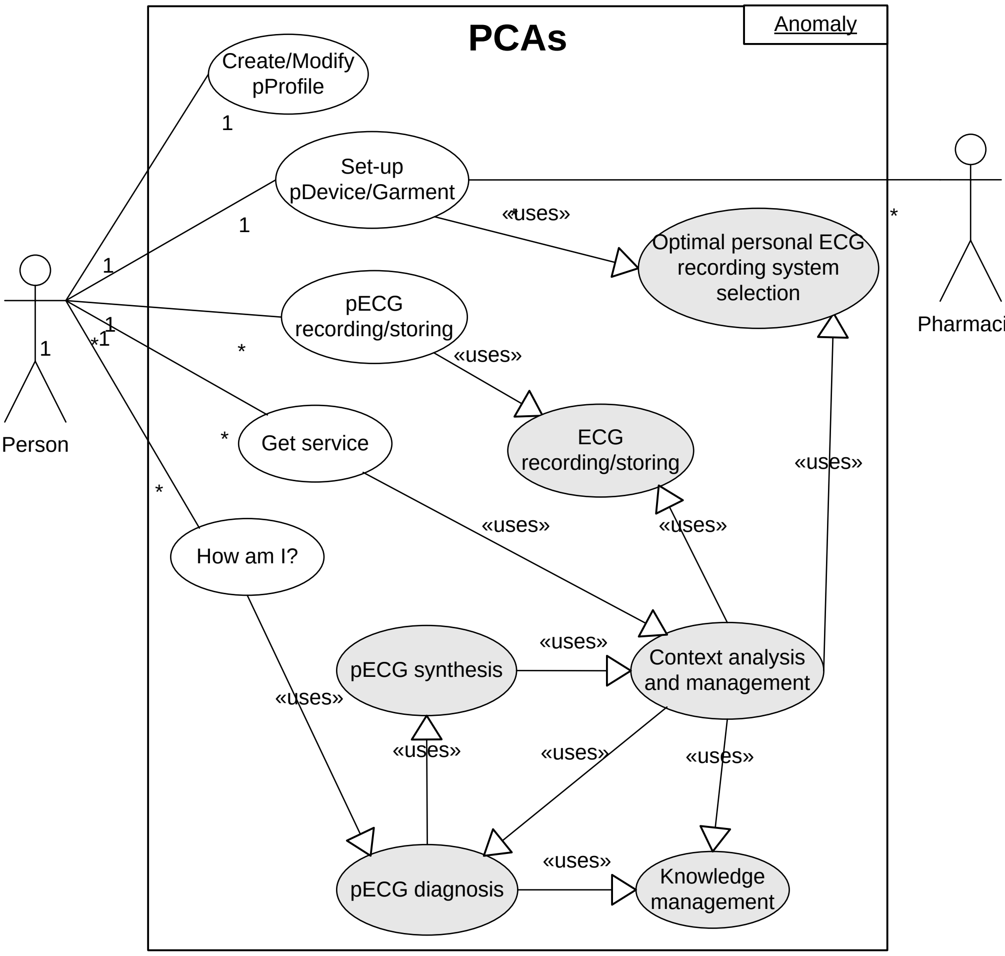


Fig.II.8_1 Use case diagram of a Personal Cardio Assistant system (PCAs) for an anomaly suspicion scenario ("p" stands for personal)

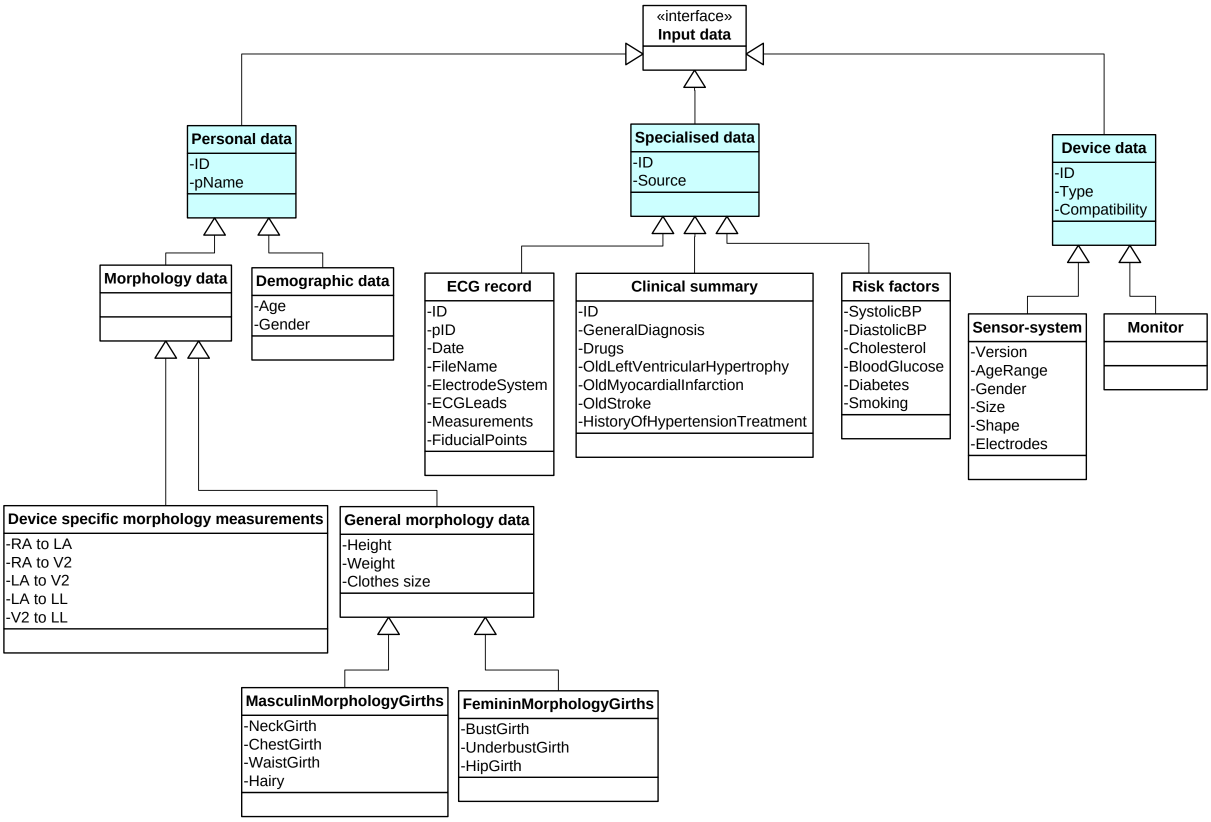


Fig.II.8_2 Class diagram of the PCAs input data system user to enter the required data specifying his personal information. We propose to describe the input data in three blocks: personal, specialised, and device data (Fig.II.8_2). The personal information typically contains person's demographic data, such as person name, age, gender and others. The block containing specialised data concerns the clinical history of the user and includes a clinical summary, risk factors (cholesterol level, smoking, diabetes, obesity, and others), and the recorded ECGs. Finally, the block for the device data is intended to describe a sensor-system used to record ECGs and a monitor, which contains system logic and performs all required computations. As we already demonstrated *[Krupaviciute'10b]*, the morphology of each subject is unique and a special sensor-system adaptation in order to record an ECG of a diagnostic quality is needed. We absolutely support the idea that the sensor-system used to record a personal ECG shall be integrated into garments in order to ensure easy electrodes' positioning and repeatable electrodes placement that is needed to obtain comparable ECGs. Therefore, we propose to describe this sensor-system in two dimensions: body dimension and product dimension. This results with the suggestion of complementing personal data that typically include only demographic data with the person's morphological data. We refer to the European standard for labelling clothes sizes EN 13402:2004 *[BS'04]* and for clothes of the upper body we propose to take into account for females: body mass, height, bust girth, under bust girth, waist girth and hip girth, while for male: body mass, height, neck girth, chest girth, waist girth. For specifying our concrete three modalities proposed sensor-system (presented in section 4.3), we suggest to measure five lengths corresponding to Mason-Likar positioning between 4 active electrodes: RA, LA, V2, and LL. Meantime, the product (the sensor-system) dimension shall include sensor-system size and type (form/shape, electrodes characteristics and compatibility with personal devices). All input data provided by the user are taken into account by **context analysis and management service** as contextual user information and are used to ensure higher system accuracy and for system personalisation purposes.

Set-up personal device and personal sensor-system (Set-up pDevice/Garment): In part I of this thesis, we concluded that the sensor-system shall be accurately selected to each PCAs user in order to provide a diagnostic quality ECG, which is crucial seeking to ensure a truthworth automated signal analysis and interpretation. The **optimal personal ECG recording system selection service** will help for that purpose. The execution of this use-case may necessitate some professional data to improve the preciseness of this service. Therefore we introduced a helping actor – a specially certified person, for instance a pharmacist – who enters into the system the required input data. In our proposed design of this service (that we will present in details in section 8.3), the minimum of input data consists in 3-lead ECGs recorded with 3 sensor-system modalities: S, M and L (presented in section 4.3). Our service concludes with proposing one modality of the sensor-system. This sensor-system modality shall then be used by the user to record his personal ECGs by himself anywhere and anytime.

Record personal ECG (pECG recording/storing): When the current data concerning the chosen personal ECG monitor and the proper electrode garment (sensor-system) are set to the personal profile (pProfile), the system user can record his personal ECG. Recorded ECGs are saved in the user's clinical summary using **the ECG recording/storing service** and data concerning the recording system used to obtain this signal are set automatically.

"How am I?": This use-case represents one of the services that is intended to be the start of each already configured PCAs. That is, when the system user switches on his PCAs, the first what he should get is a service that evaluates his current health state. Therefore, we called this service – how am I? The aim of this service is to provide an advice about myocardial infarction risk level: cardiac status is stable, minor risk, medium risk, and major risk for the system user. Following the advice, a person can decide himself of how to take care about his health. This service is related to a set of internal system services, for instance a **personal ECG diagnosis**, a **personal ECG synthesis**, **context analysis and management** and domain-specific **knowledge management** in order to perform his task. Typically, the assessment of the person's current heart status is done on the most recently recorded personal ECG. If there exist previous ECGs, the comparison results can be also shown to the system user in a form of recommendation.

Get service: The PCAs may include other services, such as alerts, follow-up scenarios, specific data storage and management. However, for an individual system and for a concrete context, not all of them will be available and not all of them can be recommended. The **context analysis and management service** shall propose the ones that are relevant to the system user.

8.2.2. Follow-up scenario (Fig.II.8_3)

The second scenario is designed for users who have already had cardiovascular problems and together with their physicians are involved in following the progress of healing. The patient that needs a constant supervision acquires a monitor device and a sensor-system that he can use in his daily routine. The selection of both these components is performed in collaboration with a professional. The monitor device and the sensor-system are specially tuned according to the user, all required contextual data are defined and patient-specific follow-up scenarios are set in the system. The professional can follow the progress of healing based on obtained data during regular visits or can get an alert if his patient state has become critical.

This scenario includes the PCA system and implies three actors: a person himself, a pharmacist or any other certified person, and a physician. The implementation of this scenario aims at allowing the patient to leave the hospital or a care center earlier (for example after a Myocardial Infarction) and still to feel secure and cared as his PCAs would continue to collect data about his heart status and, in case of some changes, would send an alert to medical care center or would provide a user-specific recommendation. This scenario proposes 10 use-cases that can be operated by actors and includes 6 tentative use-cases that can be included in the system in order to perform the required tasks. Most of the use-cases are the same as in the previous scenario, but the remaining ones we overview hereafter are mainly performed by the physician (cardiologist).

Create and modify medical profile (Create/Modify mProfile): We state that the patient and the doctor must be unambiguously identified in the system for security and auditing purposes. For the patient we have already defined a personal/patient profile (pProfile), which is accessible using a unique login and password. This profile according to the wish of the user may be related to one or more follow-up doctor profiles (mProfile). The system shall ensure that the patient himself can decide to whom he would show his clinical data. The doctor profile (mProfile) shortly describes the physician or the cardiologist with his last name, first name, institution he is working at and position he holds.

Store a reference ECG (rECG storing): It's a use-case that allows the doctor to store a reference ECG (a standard 12-lead ECG or a 12-lead ECG taken with electrodes in Mason-Likar positioning) in the PCAs. The reference ECG helps to improve the preciseness and reliability of other system services, such as optimal **person ECG recording system selection** or **personal ECG diagnosis**. Both the doctor and the pharmacist can help to choose the op-

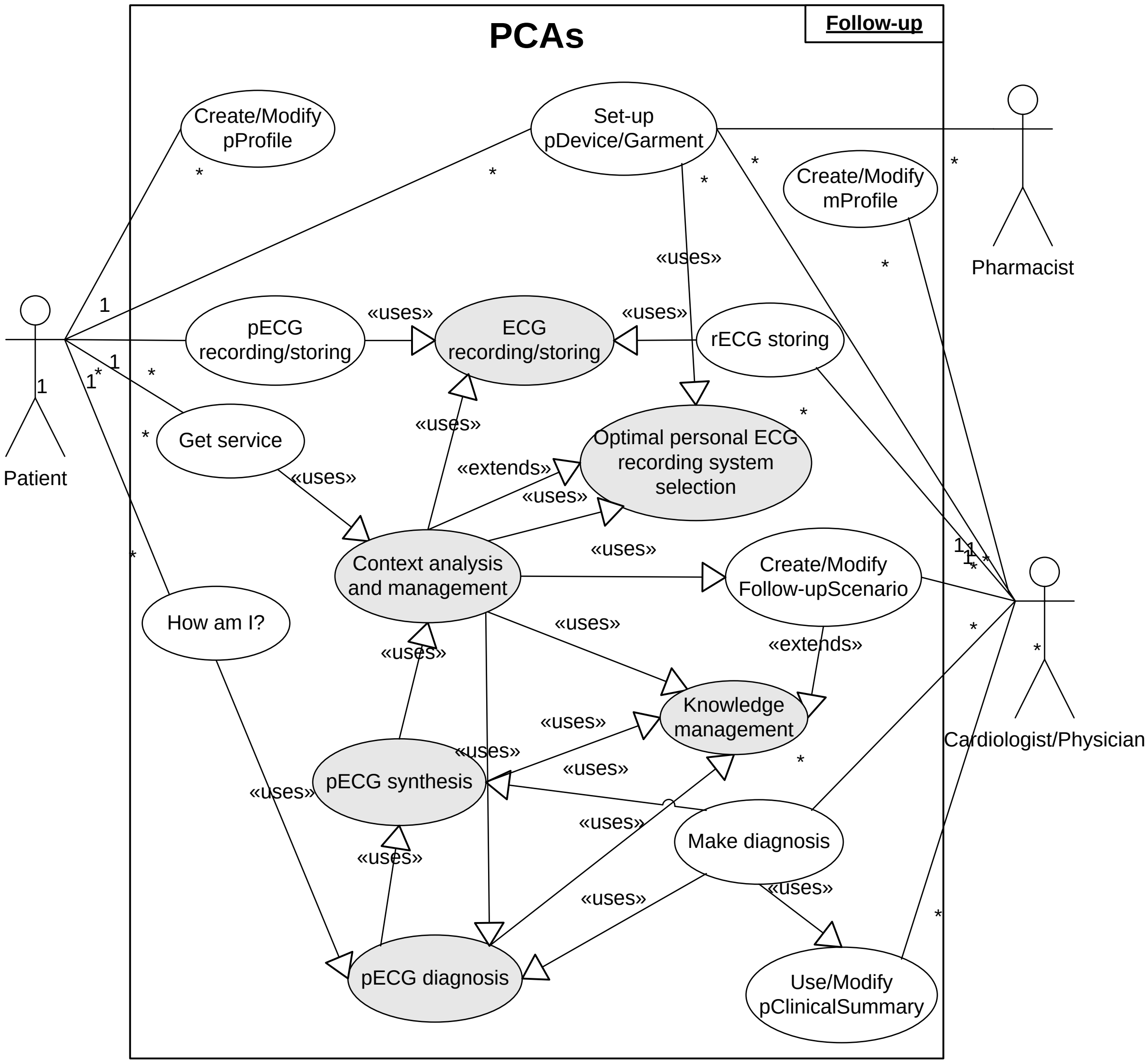


Fig.II.8_3 Use case diagram of a Personal Cardio Assistant system (PCAs) for a follow-up scenario

Timal sensor-system and a monitor (pDevice/Garment) that allow recording good quality personal ECGs (pECG).

Create and modify follow-up scenario (Create/Modify Follow-upScenario): The doctor should have tools to create a follow-up scenario, which is adapted to the concrete situation of the patient. The follow-up scenario is intended to guide the user during his healing on the daily basis activities. The PCAs with an integrated follow-up scenario can demand the patient to record regularly his pECGs or perform some activity. Additionally, the follow-up scenario may include business rules of how best to process obtained personal data or when some alerts need to be launched.

Get scenario: This use-case in this scenario is replacing the *get service* use case from the previous scenario. The patient that is followed-up with his login to the PCAs activates the service of **context analysis and management**, which searches for some patient-specific follow-up scenarios. If such scenario is not set, the system might provide a list of general follow-up scenarios that are based on medical knowledge about some patient groups. If there is an emergency, the patient can use the “how am I?” service to quickly make a decision.

Make diagnosis: The doctor in this use case has the possibility to make his diagnosis on the recorded personal or reference ECGs and to **modify patient's clinical summary** (Use/Modify pClinicalSummary). He can also take in account the digital interpretation that was performed by the **personal ECG diagnosis service** (pECG diagnosis).

8.2.3. Emergency scenario (Fig.II.8_4)

This scenario is intended to be followed by a relative or by an emergency crew on a person who is suspected to have cardiovascular problems. The relative or a member of the emergency crew puts the sensor-system that he possesses on the victim, records the ECG and gets the recommendation based on the digital assessment of this ECG. This scenario allows gathering important cardiovascular system related information that can help make fast decision.

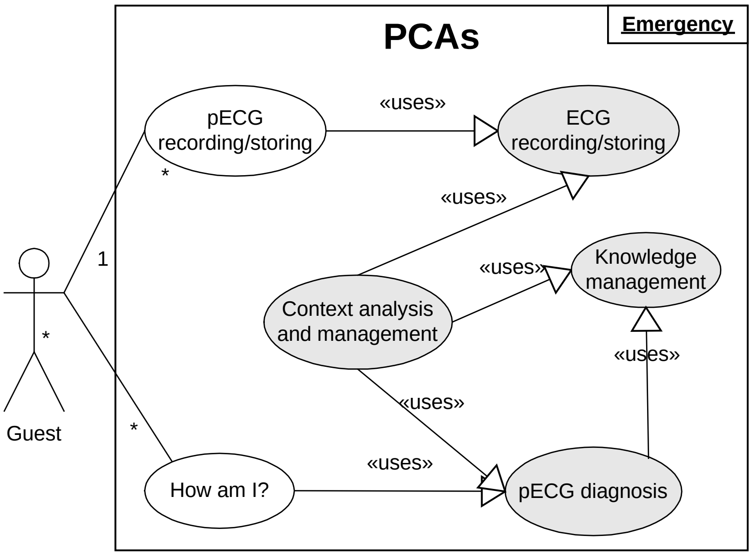


Fig.II.8_4 Use case diagram of a Personal Cardio Assistant system (PCAs) for an emergency scenario

This scenario includes the PCA system and implies only one actor: a victim, which as a guest gets PCAs assistance. We named the system user a guest to show that the PCAs does not contain any clinical history, demographic data and is entirely based on generic data processing techniques.

This scenario implements two use-cases: **record a personal ECG** and **make diagnosis** on this recorded ECG in a form of a recommendation. This diagnosis may lack of precision and even reliability of the person is not identified in the system and the personal sensor-system modality is not chosen. Thus, being not adapted to the victim's morphology, the electrodes positioning may distort the ECG signal. However, the recommendations that the PCAs would be capable to provide on the basis of generic data processing techniques in order to evaluate the victim's condition still can be invaluable.

8.2.4. PCAs general conceptual model

Each use case scenario described some central functionality that the PCAs shall contain. We present in Fig.II.8_5 a rough class diagram of a PCA system.

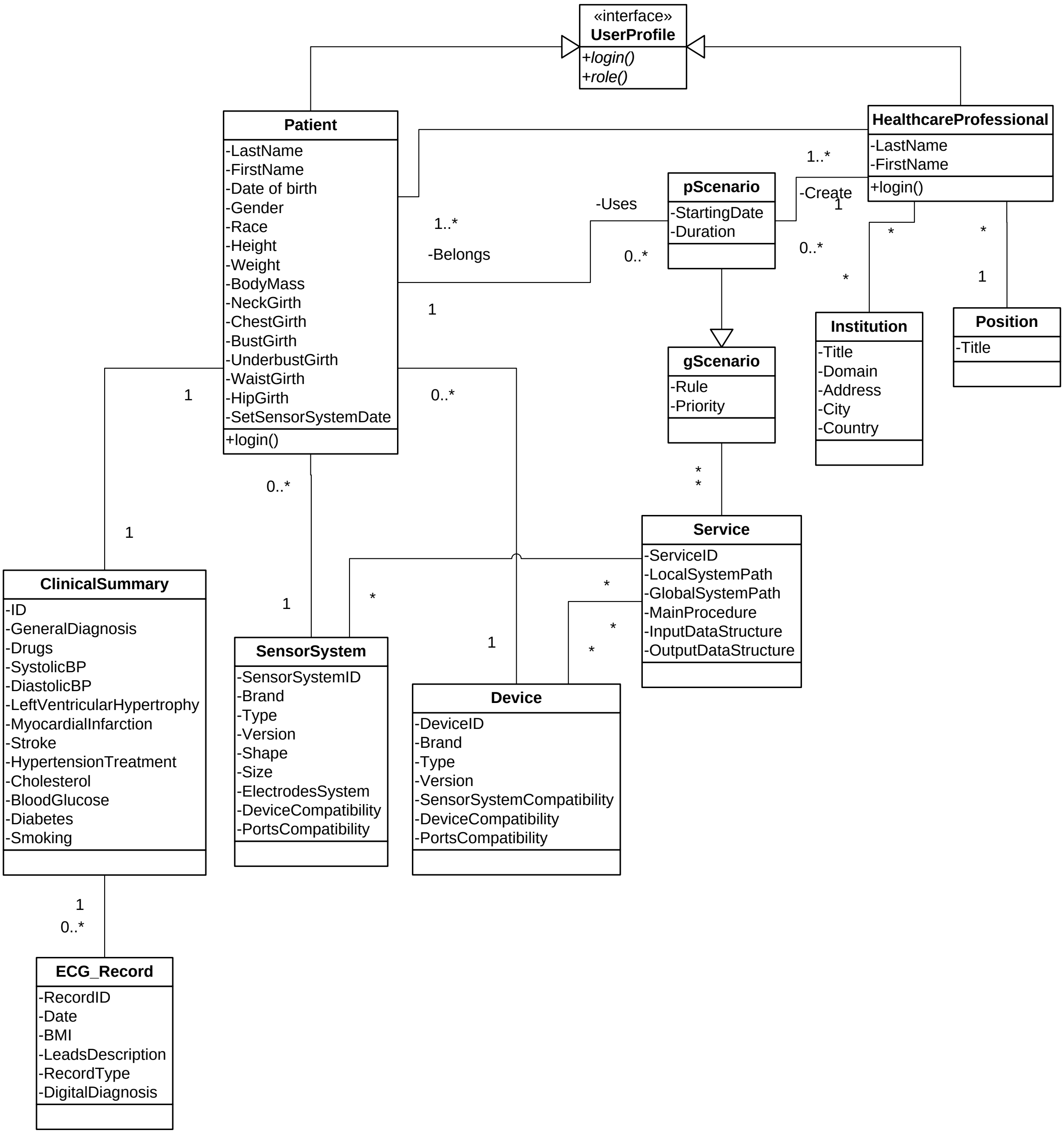


Fig.II.8_5 PCAs system class diagram

8.3. Optimal personal ECG recording system selection: Service Choreography

The functionalities of the PCAs include several very complex business services, for instance sensor-system selection or personal ECG diagnosis. However, in the context of this thesis we take only one business service – a personal sensor-system selection. The detailed example of how this business service shall be implemented will give a framework for other business services implementation. In order to better understand this service we specify it in terms of service choreography and service orchestration.

We will start with service choreography (Fig.II.8_6) as it explains a multi-party collaboration that describes externally observable interactions between composite services. Inhere, the final business service is seen from a user perspective and is easier understandable.

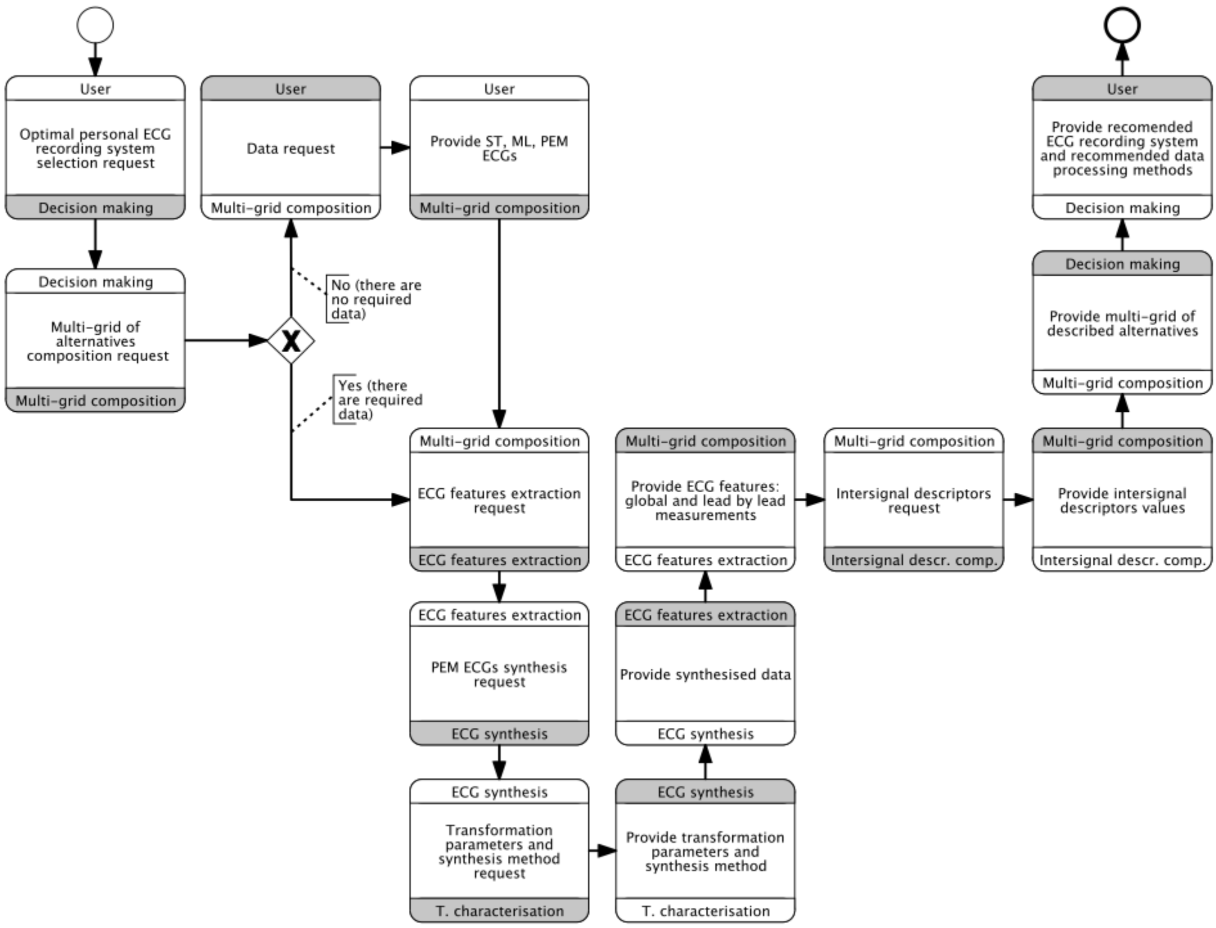


Fig.II.8_6 Choreography of the optimal personal ECG recording system selection. Legend: PEM – ECGs registered into the Personal ECG Monitor; T.characterisation – transformation characterisation; Multi-grid composition – assessment descriptors multi-grid composition; Intersignal descr. comp. – intersignal descriptors computation

We have identified 7 roles that participate in the activity of this business service. We used our proposed DM-SOA for decision making to set these roles: *user*, *decision making service*, *multi-grid composition service*, *intersignal descriptors computation service*, *ECG synthesis service*, *transformation characterisation service*, and *ECG features extraction service*. The message-based interactions between roles participating in the business service show global perspective of the interior of the business service. Actually, it is an algorithm that we can follow to understand how the collaboration among identified roles terminates by reaching the final business service objective:

1. *User* sends a request to the *Decision making service* to provide an optimal user-specific sensor-system modality.
2. *Decision making service* sends a request to the *Multi-grid composition service* to provide possible alternatives and evaluations related to these alternatives.
3. *Multi-grid composition service* checks the presence of user-specific information. In case this information is not present it sends a request to the *User* to provide the personal data needed for an optimal modality selection. In case such information is present it jumps to the step 5.
4. *User* provides the required information to the *Multi-grid composition service* (for example the standard and several personal ECGs obtained with different personal sensor-system modalities).

5. *Multi-grid composition service* sends a request to the *ECG features extraction service* to provide ECG features (measurements), which it could measure and evaluate in order to answer to the demand it got from the Decision making component.
6. *ECG features extraction service* sends a request to the *ECG synthesis service* to provide synthesised ECGs (reconstructed data) for the ECGs recorded using personal sensor-system modalities as data size and the number of leads are incompatible with standard ECG data.
7. *ECG synthesis service* sends a request to the *Transformation characterisation service* to provide transformation parameters and the data synthesis method.
8. *Transformation characterisation service* provides transformation parameters and the data synthesis method to the *ECG synthesis service*.
9. *ECG synthesis service* provides synthesised data to *ECG features extraction service*.
10. *ECG features extraction service* provides data to the *Multi-grid composition service*.
11. *Multi-grid composition service* sends a request to the *Intersignal descriptors computation service* to provide intersignal descriptors.
12. *Intersignal descriptors computation service* provides descriptors and their values to the *Multi-grid composition service*.
13. *Multi-grid composition service* provides evaluated alternatives to the *Decision making service*.
14. *Decision making service* provides to the user recommendations on ECG recording modality and on personal parameters settings, such as the description of the synthesis method to be used to get a standard ECG and the corresponding parameters.

We present hereafter in Fig.II.8._7) the body of a dynamic business process in 6 composite services for optimal personal ECG recording system selection – that is a stage (a) of the dynamic business process functioning, presented in section 6.5.

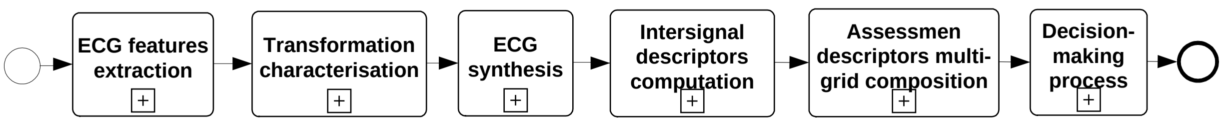


Fig.II.8_7 Schema of a dynamic business process for optimal personal ECG recording system selection

As soon as stage (a) is performed, we can deploy each of these composite services, that is we can go one level deeper to understand how each of these composite services is managed and how the contextual user data changes the behaviour of this service – perform stages (b), (c) and (d) of the dynamic business process activity (section 6.5). We demonstrate this via the orchestration of these services.

8.4. Optimal personal ECG recording system selection: Service Orchestration

Service orchestration relates to the execution of specific business processes. Service orchestration defines how (in what arrangement, in what order) services need to be executed by an orchestration engine.

8.4.1. ECG features extraction

The composite service of ECG features extraction consists in a process implemented via a set of atomic services (Fig.II.8_8) that aim to extract ECG signal parameters. Typically these parameters are extracted from one ECG beat (PQRST waves) that is selected as a representative beat for the whole record. To select this beat and then to measure the needed parameters this ECG features extraction process executes a typical ECG signal processing, which is performed in eight steps:

ECG signal filtering: The ECG signals, due to motion artefacts or to environmental disturbances causing various interferences, may be slightly distorted or "noisy". This noise is eliminated using an atomic filtering service.

ECG beats detection: A typical diagnostic ECG record is of the length of 10 seconds. If the patient heart rhythm is normal, i.e. ~60 beats/min, the record contains 9-10 beats. ECG beats detecting algorithms may apply various techniques to distinguish each beat, for example by searching for the highest peak (R wave) and taking a certain number of points before and after the peak.

ECG beats typification: When beats are detected (sometimes technical algorithms mix the typical beat detection and beats typification phases), they are classified according to their resemblance one to each other. The resemblance threshold can differ in different algorithms and so the number of classes can be different.

Typical/Median beat determination: The class having most of the beats is selected and a typical / median beat from that class is determined. This signal processing part can

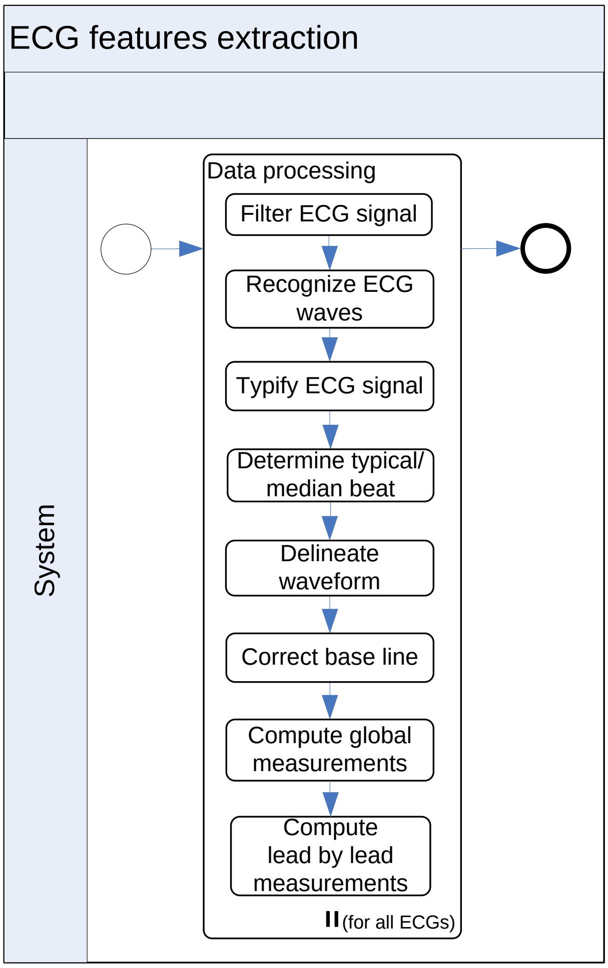


Fig.II.8_8 ECG features extraction process terminate with different beat selections when different algorithms (different atomic services) are used. Indeed, algorithms in this phase may apply different heuristics: make a fake beat that is a median of beats in the dominant class, or select the beat that resembles the most the median beat, and others.

Base line correction: The respiration or other artefacts can influence the recorded signal so that it would not keep the horizontal line – base line, therefore signal correction algorithms may shift the signal of the typical beat up or down, or may progressively change the signal to meet the base line.

Computation of global measurement: Once signal pre-processing is terminated, global measurements, that is measurements for the whole of the signal leads are calculated.

Computation of lead by lead measurements: If it is possible, measurements for each lead are taken.

To implement this composite service in practice we used atomic services based on the Lyon program *[Arnaud'00]*.

8.4.2. Transformation characterisation

The transformation characterisation process (Fig.II.8_9) is launched to calculate transformation parameters, which afterwards can be used upon an ECG obtained with a personal ECG recording modality to transform it into a 12-lead ECG in order to be comparable with a reference ECG. These transformation parameters are either generic ones (suitable to each user's ECG) or patient specific ones, typically deduced from one of the patient's standard ECGs. Then, the full process contains three principal steps: possible synthesis models design, their initialisation and, finally, transformation parameters calculation.

Synthesis model's design: We consider two of our already introduced (in section 4.6) synthesis models: $\mathcal{M}_\alpha$ and $\mathcal{M}_\beta$. The application of the $\mathcal{M}_\alpha$ strategy requires only a standard 12-lead ECG, while the application of the $\mathcal{M}_\beta$ strategy demands two reference ECGs: the standard 12-lead ECG (ST) and the 12-lead ECG recorded with electrodes positioned according to Mason-Likar placement (ML). We consider that not each system user has several 12-lead reference ECGs in advance. Then, depending on his clinical data, one (the $\mathcal{M}_\alpha$ strategy), or both (the $\mathcal{M}_\alpha$ and the $\mathcal{M}_\beta$ strategy) will be designed. In particular, we suppose that the ML-ECG is not mandatory required. At least one synthesis model shall be designed in order to proceed.

Models initialisation: The initialisation aims at setting source and target data for a transformation procedure with user-specific data in order to obtain user-specific transformation coefficients. A couple of source and target data may differ, depending on the used transformation strategy and patient context. The $\mathcal{M}_\beta$ strategy uses two different ECG signals, ST and ML, then before setting source and target data, these two signals need to be aligned. Consequently, the $\mathcal{M}_\alpha$ strategy uses only one ST-ECG to set the source and the target data, so alignment step is not needed. The available patient data will determine how many source and target couples will be initialised. The process will operate successfully even if the user possesses several ST and several ML ECGs. All possible combinations would then result in a source and target couple.

Transformation parameters computation: This final step in this business process checks for the available data synthesis methods (like Artificial Neural Nets - ANN and linear regression - REG) and taking couple by couple the source and target data produces, transformation parameters. Indeed, we have employed services that use ANN and REG methods, but other methods available in the user context and suitable with the data synthesis

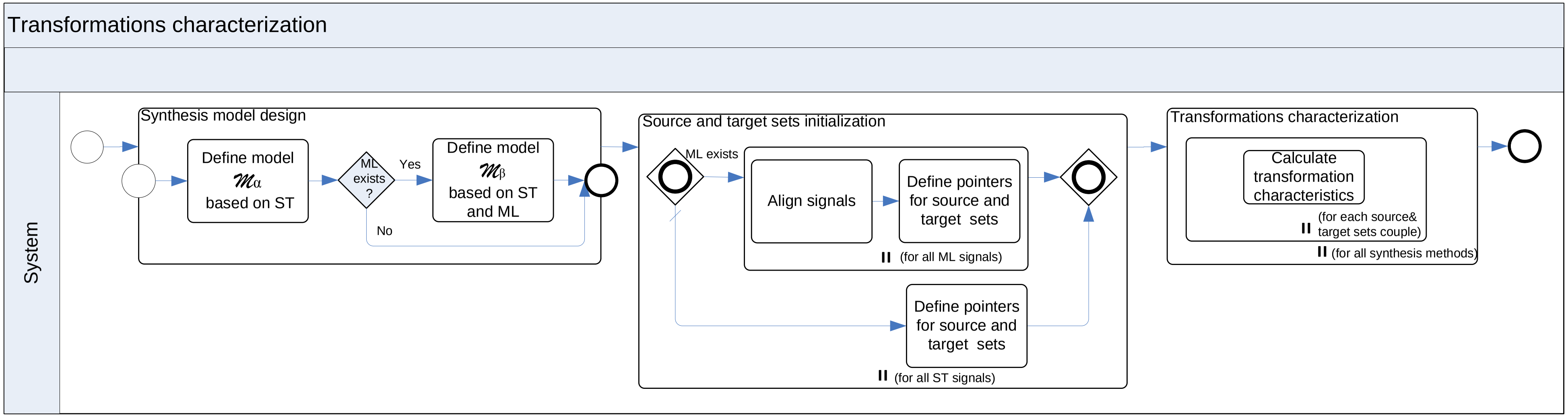


Fig.II.8_9 Transformation characterisation process. ST stands for standard 12-lead ECG and ML for Mason-Likar electrodes positioning based ECG

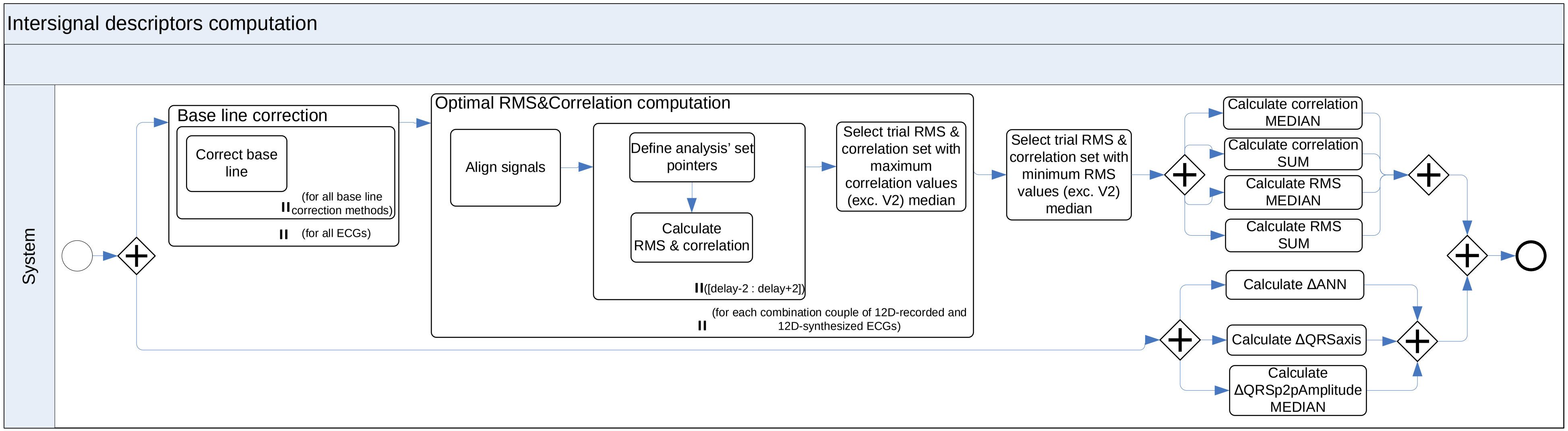


Fig.II.8_10 Intersignal descriptors computation process strategy can be used. All obtained transformation parameters sets are reusable and are applied when a new reduced-lead personal ECG is recorded with a personal ECG recording modality. That is, if the personal recording modality remains the same and there is no newer ST or ML ECGs, this process is not repeated.

8.4.3. ECG synthesis

All reduced lead ECGs can be synthesised (Fig.II.8_11) to correspond to a reference ECG if two initial steps were performed in advance: (1) the reduced lead ECG was preprocessed and (2) transformation parameters specific to the patient are calculated (or generic transformation parameters are available).

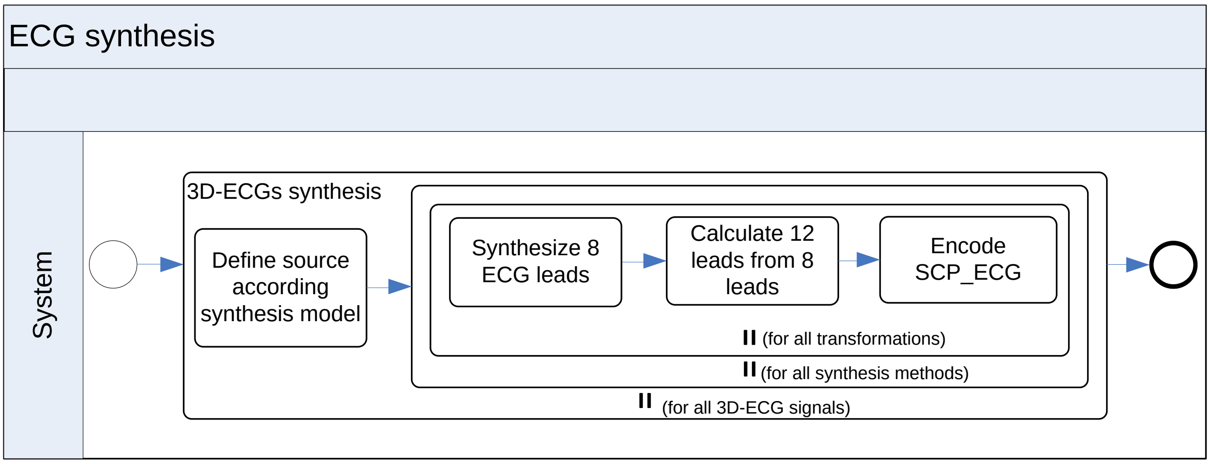


Fig.II.8_11 ECG synthesis process

Input initialisation: synthesis input initialisation is performed by taking the ECGs that need to be synthesised. Here we select our 3-lead ECGs recorded with one of the sensor-system modalities. The synthesis input initialisation step shall consider the synthesis strategies. As in our case both strategies use the same 3 leads (I, II, V2), data initialisation terminates with the preparation of three Ponset-Toffset signal segments from a 3-lead personal ECG.

Data synthesis: The peculiarity of data synthesis lies in the application of all available transformation parameters and using all available synthesis methods. The difficulty of the managing agent is to ensure that the adequate transformation parameters set will be applied on adequate input data with an adequate data synthesis method. According to the data synthesis strategy, either five leads ($\mathcal{M}_\alpha$) or seven leads ($\mathcal{M}_\beta$) are reconstructed. In a case of five synthesised leads, the three source leads are added to form a set of eight lead. In the other case, where seven leads are synthesised, only an originally recorded V2 lead is taken to form a set of eight leads. In the next step, from these eight principal leads the remaining four are derived and finally the signals of the full set of twelve leads are constructed and encoded according to the SCP_ECG format. The three last steps: construction of the 8 leads set, calculation of the 4 remaining leads in order to form 12 leads and 12-lead ECG encoding, are repeated till there are no available and suitable transformation parameters and data synthesis methods. Finally, the complete procedure is repeated with another 3-lead ECG.

8.4.4. Intersignal descriptors computation

We have defined in section 4.6 seven intersignal descriptors that are intended to evaluate the signal pattern quality and the medical content quality of an ECG. These parameters are obtained from two ECGs, one of them is a reference ECG and another is the one we expect to evaluate. In our case, a reference ECG is either the ST-ECG or the ML-ECG, and ECGs that we seek to evaluate are synthesised personal ECGs recorded with different modalities S, M and L. The aim of the intersignal descriptors computations is to express the difference among these two ECGs in a unique form. The values corresponding to these intersignal descriptors then can be assessed according to already determined thresholds or compared with other ECGs' descriptors' values and these evaluations' results can be used by the decision-making process.

Following the intersignal descriptors computation process (Fig.II.8_10), at first these descriptors are separated in two groups:

The first group consists of four descriptors that are based on cross-signal measurements via the correlation (r) and the RMS (M) values. These descriptors are calculated on the Qonset-Toffset signal segments from eight ECG leads: I, II, V1-V6 following the rigorous procedure that we explained in section 4.6. These four descriptors allow evaluating the signal deviations in comparison with the standard ECG, in particularly the quality of the signal synthesis.

The second group contains the remaining three descriptors: Δp (ΔANN), difference of the artificial neural net output for the reference and the synthesised ECGs; ΔQRSaxis, difference between the QRS axis of the reference and the synthesised ECGs; and ΔK (QRS peak to peak amplitude median), calculated by taking the difference between the medians of the lead by lead QRS peak to peak amplitudes from the original and the synthesised ECGs. For the accuracy, delta indicates that a defined value of a synthesised ECG was subtracted from a defined value of a reference ECG.

Δp (ΔANN): An artificial neural net (ANN) output value for each ECG is calculated. This calculation is based on a principal component analysis, during which an elaborated selection of specific ECG signal parameters is performed and the selected and pre-processed measurements are fed into an already prepared artificial neural nets committee *[Atoui'10]*. The answer that the ANNs provide is a number in a range between 0 and 1. This output value indicates probability of the risk of occurrence of myocardial infarction. Then, the delta ANN value (ΔANN) or, as we indicated, Δp explains a qualitative difference between the synthesised and the reference ECGs diagnosis interpretations.

ΔQRSaxis and (ΔQRSp2pamplitudeMedian): The delta QRS axis is calculated on a base of the QRS axis value of each ECG. The QRS axis values are taken from the global signal measurements sets and then the QRS axis values of the synthesised ECG is subtracted from the ones of the reference ECG. This descriptor as well as the third one, the median of the delta peak to peak QRS amplitudes calculated from the synthesised and reference ECGs lead by lead measurements, is used to characterise inadmissible signal deviations and in general terms express signal quality.

8.4.5. Asssessment descriptors multi-grid composition

In order to make a decision, the processed data shall be aggregated into a meaningful whole. This role is taken by the process of assessment descriptors multi-grid composition (Fig.II.8_12.). A new data structure, a grid, here is introduced seeking to manipulate mea-

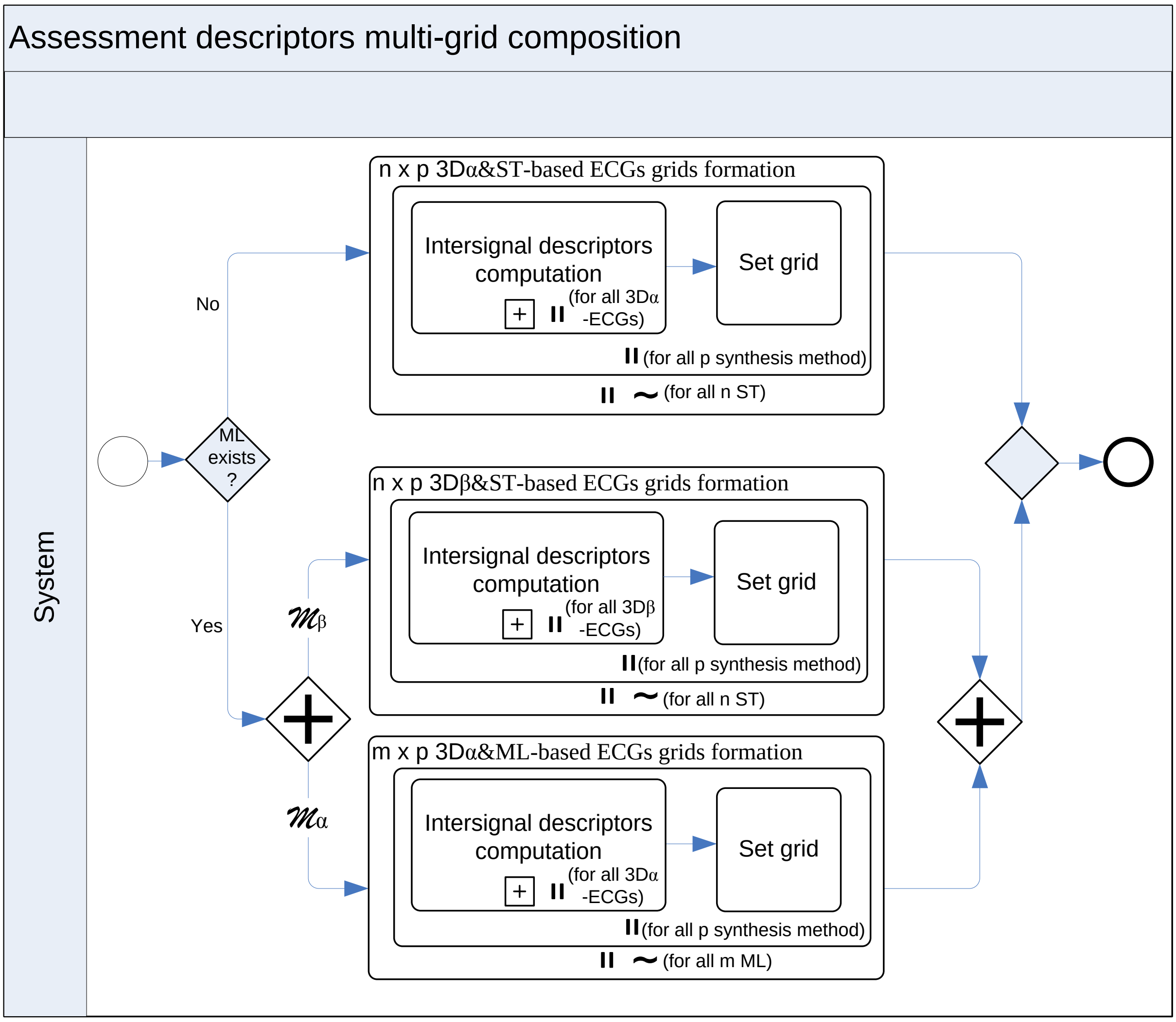


Fig.II.8_12 ECG features extraction process

ningful parameters in a manner which afterwards will be taken as an input for a decision making process. This multi-grid shall express the data through a unified data evaluation parameters set and also through a subject related contextual information.

Our here described business service aims at selecting an optimal ECG recording modality from a couple of personal sensor-system modalities and to recommend which data synthesis method (from the available ones) goes best with it for a concrete patient's case. Therefore, the contextual information we are interested in are data acquisition modalities and synthesis methods. For simplicity, let us operate already defined data acquisition modalities – S, M and L, and two already declared data synthesis methods: linear regression (REG) and ANN. The unified data evaluation parameters are represented by a seven descriptors set defined by the assessment descriptors computation process (section 8.4.4). Then, the concrete goal is to select which modality S, M or L and with which synthesis method REG or ANN the sensor-system is capable to produce the best quality ECG for the specific user.

According to these stated settings, one grid then will correspond to a matrix of values representing comparison values between a reference 12-lead ECG and three patient specific 3-lead ECGs (recorded with S, M and L sensor-system templates) synthesised with one of the available synthesis methods (Annex 4.2).

Besides the above mentioned grid structure, also the initial target related reference and synthesised ECGs comparisons should be set in order to construct a decision related multi-grid. For our representative case we have modelled a multi-grid from three principal levels. If a patient has only an ST type ECG as a reference ECG, only the first level multi-grid modelling will be performed. Though, if a patient also has an ML type ECG as an additional reference ECG all three levels of multi-grid modelling will be performed.

The first multi-grid modelling level sets the standard ST-ECG as a reference ECG and takes set by set all simplified ECGs synthesised with one method. Here then the number of available synthesis methods obtained from the patient's context will define the multi-grid's width. There also might happen that the patient for some reasons owns more than one ST, in that case after finishing comparisons with the first ST, the second ST will be set as a reference one and the comparisons' procedure with all synthesised ECGs ordered according to the synthesis method they were synthesised with will be repeated. As a result the multi-grid height will depend on how many ST type ECGs a patient owns.

The second and third multi-grid modelling levels are similar to the first level as the same set by set of all simplified ECGs synthesised with one method will be taken. Though, there exists some additional domain knowledge based management. The second and the third levels are influenced by the existing of another reference ECG type, the ML, what indicates that the synthesis strategy $\mathcal{M}_\beta$ was used and so there are more comparison options. We state that ECGs synthesised following the $\mathcal{M}_\beta$ model are by default closer to the ST, as not only the relationship between leads, but also the electrodes position changes were included into this model. Conversely, ECGs synthesised according to the $\mathcal{M}_\alpha$ model are closer to the ECG with electrodes positions similar with a data acquisition template, in our case that's an ML 12-lead ECG. Therefore we compare all simplified ECGs that were synthesised according to the $\mathcal{M}_\beta$ model with the ST type reference ECGs and we compare all simplified ECGs which were synthesised according to the $\mathcal{M}_\alpha$ model with the ML type reference ECG. Again the number of ST and ML ECGs will augment the height of the multi-grid, while the width will remain the same as in the first level as data synthesis methods remain the same.

8.4.6. Decision-making

Finally, the decision-making process (Fig.II.8_13) shall produce the evaluation of multi-grid values using a predefined decision strategy and provide a recommendation to user's problem. Various data evaluation methods can be included in this process and can be

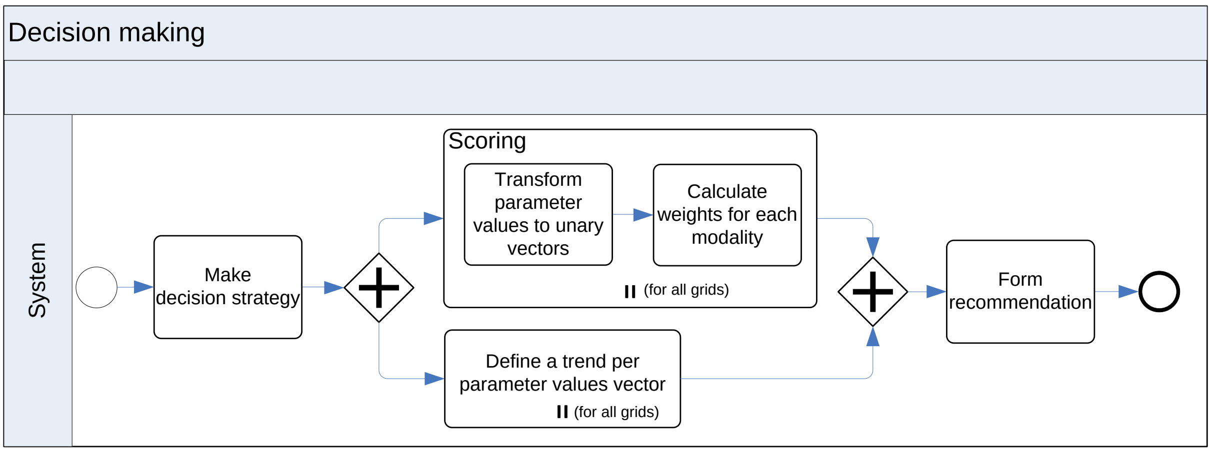


Fig.II.8_13 Decision making process

performed in parallel, one with each other. Though first of all a strategy of how to manage and weight these different evaluations and how to conclude shall be set. Then, the conclusion should be "translated" into a form which is understandable by a user and can be provided as a recommendation.

As an example, we chose a decision strategy to evaluate our multi-grid using two techniques: scoring and trend analysis. We first selected the data acquisition modality size and then using our developed technique, we selected the data synthesis method. Though, we state that additional techniques, such as case-based reasoning, artificial neural nets and others might only improve the preciseness of a final decision.

The aim of the scoring method is to express synthesised ECGs' values per one descriptor via a simplified score: 1 – the synthesised ECG descriptor is enough similar to the reference ECG or 0 - the deviation from a reference ECG is not acceptable. In the first step all multi-grid values are transformed into absolute values, as they all represent deviation from a reference ECG, no matter the deviation direction. Then, if we first want to select the template size, we treat all grids of the multi-grid independently. For each grid the minimal descriptor value is identified in a vector row representing all templates (S, M and L). If this value is a zero value we add 1 to all this descriptor vector values and then the minimal value for that descriptor is 1. Then we divide each vector with its minimal value to obtain as we call it a coefficients vector. At this point an expert knowledge can be introduced as we want to form unary vectors for each descriptor, where 1 corresponds to the ECG descriptor value which is enough similar to the reference ECG descriptor. This "enough" here should be expressed via a numerical value and so its determination is making influence on the final decision. Via this value a domain expert or expert's group opinion can guide the system. In our example we were rigorous and we took k=1, that is only the values producing the minimal deviation from the reference ECG can be treated as acceptable ones. Though for an interest we experimented as well with setting k=1.2, that's coefficient vector values less or equal to 1.2 will still be counted as acceptable ones (marked with 1) and values superior to 1.2 will be treated as inacceptable ones (marked with 0). Using this manner the grid of descriptor values becomes a unary grid. The next step is to calculate weights for each modality. There the whole multi-grid should be considered. First of all we separated parameters into two groups representing signal interpretation evaluation part and signal synthesis evaluation part (Fig.II.8_14). The interpretation part was represented by a ΔANN descriptor and the remaining six descriptors expressed data synthesis evaluation part. We set that interpretation and synthesis descriptors evaluation parts have an equal importance on deciding which sensor-system modality is providing the best quality record. Though as the synthesis evaluation part is represented by six descriptors, we grouped them by two and defined that descriptors expressing RMS and correlation sum have the highest priority, descriptors marking RMS and correlation median values have the middle priority and the two remaining parameters:

Evaluation parts:	Interpretation			Synthesis																		
Descriptors:				Sum						Median						QRS						
	a			b			c			d			e			f			g			
Priority:			12			3			3			2			2			1			1	
Trend	Mix			Mix			Mix			Mix			Mix			Mix			Mix			
Size:	S	M	L	S	M	L	S	M	L	S	M	L	S	M	L	S	M	L	S	M	L	
Result:	2	1	0	1	1	1	2	1	0	2	1	0	2	1	0	1	2	1	2	0	1	**ANS**
S			24	20																		**44**
M			12	12																		**24**
L			0	5																		**5**

Fig.II.8_14 Example of size selection delta QRS axis and median of the delta peak to peak QRS amplitude, which were used as instant error preventing parameters, are of the lowest priority.

The aim of the trend analysis is to support the scoring evaluation result. Using this method, the absolute multi-grid values are taken. At first, grids are processed independently and then, an evaluation of the whole set of grids (the multi-grid) is performed. The goal of the trend analysis is to indicate if the vector values have a tendency to increase (Aug) or to decrease (Des), otherwise their state is set as a Mix. The trend analysis idea is based on the peculiarity of data acquisition modalities, which shall be designed of different sizes. The size concept includes an idea of some tendency and so the deviation tendency would support one or another template selection.

The same scoring and trending techniques are used for choosing the best performing synthesis method. However, here the multi-grid is reduced to the only already selected size.

8.5. Conclusion

In order to overcome subjects' inter- and intra-variability, we have followed a systematic approach to the care delivery, grounded by our DM-SOA for decision support systems approach. It includes the management of interrelated processes, which are implemented via appropriate services. The latter are selected and invoked by a process-centred, knowledge based orchestration. We have successfully demonstrated this principle by orchestrating domain-specific services that implement the goal-oriented dynamic business process of the optimal sensor-system modality selection. This adaptive orchestration allows combining different sensor-system modalities, a number of data synthesis strategies, a set of data synthesis methods and grid-based comparisons. It successfully addresses the challenge of the spatiotemporal diagnostic ECG information retrieval and results in an optimal electrodes' placement for each system user. This intelligent data processing dynamically handles contextual data, demonstrates the adaptation to changing circumstances and provides accumulated domain-specific knowledge for a personalised assistance.

Conclusions and future work

The objective of Ambient Intelligence in HealthCare is to broaden the interaction between any citizen and professional care services by involving advanced technologies through the use of ubiquitous computing devices. The core problem in making Ambient Intelligence present in everyday life and enhancing current HealthCare procedures is not technological. It lays in understanding upon the functioning of physical, biological and technical reality and the ability to manage the awareness of information systems in order to offer new generation services that possess a purposeful user adaptive behaviour. This was the main problem we related our research work to and dedicated this thesis for. We are convinced that Ambient Intelligence has the potential to significantly impact the current prevailing HealthCare processes and the private life of each citizen.

Chapter 9 – Final conclusions and perspectives

9.1. Conclusion 175
9.2. Discussion 176
9.3. Limitations 177
9.4. Future work 177

9.1. Conclusion

We have proposed to formulate the problem of self-care in cardiology as a development problem of an information system providing business services that encapsulate automated expert activities. We named this information system the Personal Cardio Assistant system and viewed it as a cognitive agent operating in an Ambient Intelligence environment. In order to answer the challenges related to this type of systems development, we have first performed experiments in the application domain and then, on the basis of the gathered knowledge, we have chosen the contributing techniques of information structuring, processing and management as a foundation to construct our approach based on data and model driven services orchestration, which aims at automating sophisticated and personalised professional care activities.

The analysis of the application domain lead us to draw a scene of algorithms' and methods' networks that ECG signal shall pass in order to result with a decision and to find pathways that could be applied when using in home-based electronic equipment. We found out that the main blocking element disturbing wearable sensor technologies to use these pathways and to enable ubiquitous self-care is their disability to ensure a correct electrodes placement for different morphology subjects and therefore to record diagnostic quality signals. These findings let us conclude with the principal objective of this thesis saying that instructions on anatomically correct electrodes' location that professionals are now using shall be replaced with an automated assistance to non-professional users aiming at self-care. We pictured this assistance as a combination of a wearable sensor-system having several modalities in order to adapt to various morphologies and an intelligent assistant, which helps to apply this sensor-system.

To investigate this solution we designed an easy to apply, used on demand sensor-system of three different modalities. Then, using this new design we performed a set of experiments on eight healthy volunteers and observed that an only one size sensor-system or sizes of the wearable sensor-system modalities only corresponding to the clothes size of the subjects do not ensure recording a diagnostic quality ECG signal. The results of this experiment highlighted the importance in having different sensor-system modalities, using different types of signal recordings, applying several data transformation strategies and available data synthesis methods for the selection of the personalised placement of electrodes. We concluded that the optimal, user-specific sensor-system providing a reduced lead set ECG signal of the quality as much as possible adequate to a standard 12-lead ECG shall be chosen for each user on the basis of his contextual information. In order to automate this process and so to ensure clinically-viable monitoring systems foreseen for self-care, a complex data processing of user-specific heterogeneous information needs to be set. We projected that this data processing needs to be performed proactively by considering the current input data, morphological and sensor-system application changes, availability and suitability of related data processing and analysis techniques, so to adjust itself as a domain professional (an expert) would do.

To enable such high level automation we introduced a new, dynamic business process concept which enhances business process adaptation to its context in order to provide coherent user-specific results. The dynamic business process is capable of providing different workflows, where each is adapted to the context of a specific user, even so this context as a whole or in part is only known at runtime. This extends the typical business process development view, mainly focusing on services composition, with a business intelligence view which aims to support the design of complex, automatically reconfigurable services. In order to set-up this intelligent environment we investigated the potential integration of knowledge management (KM) and business process management (BPM) and we designed an approach that proposes a composition model that is based on process-oriented knowledge modelling, which maps a domain knowledge representation with a data-driven business process context. We suggested implementing such a model on the basis of a context-aware concepts' ontology of business processes and on related context handling rules, which sense and content would ensure the coordination and the correct execution of business process activities (tasks). Following this approach we created an ontology-based instrument for decision support systems' management. It's a generic method that enables to capture and to granulate the knowledge into a set of models that support dynamic business processes management.

Finally, to ground this approach we provided a Data and Model driven Service Oriented Architecture (DM-SOA) architectural framework to support the automation of expert activities by deploying the dynamic business process. In order to encourage the employment of this framework we specified it as a DM-SOA for decision support systems. At last, we demonstrated the soundness of the method and of the underlying decision-making techniques in the cardiology domain by showing the automatic selection of the most adequate subject-specific sensor-system, characterised by its ability to yield for an ECG signal of similar diagnostic content than a standard 12-lead ECG.

9.2. Discussion

This thesis represents a multidisciplinary research study that aimed at overall integration issues in order to provide a consistent solution for a concrete problem in the application domain. The main issue was to find out possible integration solutions for the related domains and to fill open gaps in order to deal with all problem-specific challenges and not to search for a best possible integration of available technical tools. Therefore, the selection of other suitable techniques that would contribute (maybe even better) to the final solution is possible and even expected as a future work.

Our study implies that the ontology-based enhancement of business services orchestration allows implementing a user-adaptive and dynamically reconfigurable business process. This approach puts the automation of business services orchestration, which keeps experts know-how and applies it as process logic, at the centre of business process modelling. We notice this capacity as a very important feature, which may lead automated business services orchestration (BSO) at becoming a major advancement in e-Business proposing e-Services.

We also want to underline the potential of systems' personalisation, which was a key aspect we studied under the scope of this thesis. This characteristic allows systems not only to contribute to the creation of user-friendly environments, but, in particular, to step towards the automation of activities that were long time devoted only to human-experts. New techniques and an intelligent environment developed for supporting this automation could change the whole paradigm of currently provided human-to-human interaction based professional care. These automations would allow non-expert users to autonomously use personalised professional services and would shift the role of domain professionals currently responsible for making diagnoses, providing an advice or teaching, to a role of a knowledge engineer.

9.3. Limitations

The limitations that we observe in the relation with this research study are largely originated to the specificity of the application domain this thesis has dealt with. The major requirement for information systems that are targeted for the medical domain is patients' safety. Therefore this application domain is error-intolerant and any conclusions need to be tested on a large database. For this reason, the results that we obtained and presented in part I need to be taken as experimental results.

9.4. Future work

This research study has a distinction of taking pieces from many research domains. Therefore the contribution it provides tends to influence several domains and so to trigger the prospective research work routes at the application level as well as at the conceptual level.

At the application level, we witnessed a lack of clinical research where patients' morphological data would be taken in consideration. We hope that our proposed sensor-system design would fill this gap and can be used as an utility to register these data and so will trigger an accumulation of such data in clinically reliable databases. Such databases then could be analysed invoking knowledge extraction techniques, for instance data mining, and could contribute with new decision-making strategies applied in self-care. This would greatly benefit intelligent assistance delivery to end-users.

At the conceptual level, we envisage several directions where future work could proceed. This would include support of intelligent assistant development, enhancement of underlying knowledge enabling intelligent assistance, and, finally, extension (new application/new employment) of our proposed techniques in different domains. In order to speed-up the development of intelligent assistants, new tools, such as a platform for dynamic business processes modelling or an engine which can execute rules from the compositional knowledge models, are needed. The meta-object facility (MOF) development for any intelligent assistant would provide a necessary terminology and underlying knowledge enabling a unified modelling of all intelligent assistants. A comparison and consequently the improvement of a set of intelligent assistants (for example related to one type of problems – decision-making) would provide an evidence-based data set that would allow the enhancement of various expertises. Finally, our proposed data and model driven business services orchestration might be applied in other domains, for instance to support autonomous resources management.

In summary, our dissertation is composed of two main parts.

In *part I* we have proposed the design of a new wearable sensor-system enabling any citizen to record its own ECG signal. This sensor-system has the advantage to take in account the persons' morphological differences in order to ensure the recording of a diagnostic quality ECG in self-care. It can be cheaply manufactured, and should therefore trigger a widespread use of personal ECG monitors *[Fayn'10]*. It preserves the day-to-day or month-to-month reproducibility of the patient's signals and the results are compliant with the minimum requirements for providing a digital cardiac risk stratification *[Atoui'10]*. Finally, our sensor-system is simple to use and we eliminated the constraint of continuous sensor-system wearing, which might be uncomfortable and even impractical in real application scenarios.

In *part II* we have proposed an adaptive web services based architecture, which enables the completely personalised and automated selection of an optimal sensor-system modality for each individual. The developed business service can be used by all citizens. Moreover, this architecture is generic and can be applied in other domains to support the non-professional users' decision-making. In order to use it in another domain, it is sufficient to create a process-oriented ontology related to the specific problem by following our proposed context-aware decision-making core-ontology, which will provide the needed knowledge models and operable data acquisition modalities in order to build the needed atomic services and to orchestrate them automatically thanks to our already specified composite services.

In conclusion, our contribution should facilitate the development of self-care in cardiology, and thus an early, ubiquitous detection of the main cause of mortality, which represents an important societal impact.

References

[Abbas'08], Abbas K., Flory F., Verdier C. Adapted and Contextual Information in Medical Information Systems. In: Proc. of the 10th intern. conf. on Enterprise Information Systems, 2008, Barcelona, Spain: Volume HCI, vol. 5, pp. 363-366.

[Abowd'99], Abowd G.D., Dey A.K., Brown P.J., et al. Towards a Better Understanding of Context and Context-Awareness. In: Proc. of the 1st intern. symposium on Handheld and Ubiquitous Computing, 1999, London, UK: Springer-Verlag, pp. 304–307.

[Adamopoulou'07], Adamopoulou P., Sakkopoulos E., Tsakalidis A., et al. Web service selection based on QoS knowledge management. Journal of Universal Computer Science, 2007, vol. 13, no. 9, pp. 1138–1156.

[Ahn'05], Ahn H.J., Lee H.J., Cho K., et al. Utilizing knowledge context in virtual collaborative work. Journal of Decision Support Systems, 2005, vol. 39, no. 4, pp. 563-582.

[ANSI'01], ANSI. Diagnostic electrocardiographic devices (ANSI/AAMI EC11:1991/(R)2001). Arlington, VA: Association for the Advancement of Medical Instrumentation, 2000, p.

[Arnaud'90], Arnaud P., Rubel P., Morlet D., et al. Methodology of ECG interpretation in the Lyon program. Methods of Information in Medicine, 1990, vol. 29, no. 4, pp. 393-402.

[Atoui'06], Atoui. Conception de systèmes intelligents pour la télémédecine citoyenne. Thesis Inserm ERM107. Lyon: Insa de Lyon, 2006, 184 p.

[Atoui'06a], Atoui H., Fayn J., Gueyffier F., et al. Cardiovascular Risk Stratification in Decision Support Systems: A Probabilistic Approach. Application to pHealth. In: Proc. of 33rd intern. conf. on Computers in Cardiology, 2006, Valencia, Spain: IEEE, pp. 281-284.

[Atoui'04], Atoui H., Fayn J., Rubel P. A neural network approach for patient-specific 12-lead ECG synthesis in patient monitoring environments. In: Proc. of 31st intern. conf. on Computers in Cardiology, 2004, Chicago, USA: IEEE, pp. 161-164.

[Atoui'10], Atoui H., Fayn J., Rubel P. A Novel Neural Network Model for Deriving Standard 12-lead ECGs from Serial 3-lead ECGs. Application to Self Care. IEEE Transactions on Information Technology in Biomedicine, 2010, vol. 4, no. 3, pp. 883-890.

[August'58], August T., Mazzeleni A., Wolff L. Positional and respiratory changes in precordial lead patterns simulating acute myocardial infarction. American Heart Journal, 1958, vol. 55, no. 5, pp. 706–714.

[Ayache'05], Ayache N., Boissel J.P., Brunak S., et al. Towards virtual physiological human: Multilevel modelling and simulation of the human anatomy and physiology [Online]. Available: http://ec.europa.eu/information_society/activities/health/docs/events/barcelona2005/ec-vph-white-paper2005nov.pdf. (Accessed: 15-Sep-2011).

[Baier'08], Baier J.A., Fritz C., Bienvenu M., et al. Beyond classical planning: procedural control knowledge and preferences in state-of-the-art planners. In: Proc. of the 23rd national conf. on Artificial intelligence (AAAI'08), 2008, Chicago, USA: AAAI Press, vol. 3, pp. 1509–1512.

[Baldauf'07], Baldauf M., Dudstdar S., Rosenberg F. A survey on context-aware systems. Ad Hoc and Ubiquitous Computing, 2007, vol. 2, no. 4, pp. 263-277.

[Barrett'06], Barrett R., Patcas L.M., Pahl C., et al. Model driven distribution pattern design for dynamic web service compositions. In: Proc. of the 6th intern. conf. on Web Engineering, 2006, Palo Alto, California, USA: ACM Press, pp. 129-136.

[Bartosik'95], Bartosik J., Pahlm O., J. Edenbrandt J., et al. Reconstruction of the standard 12-lead ECG from recordings using nonstandard activity-compatible proximal limb lead positions. Journal of Electrocardiology, 1995, vol. 28, no. 1, pp. 33-38.

[Beack'06], Beack T. SOA Maturity Model: Compass on the SOA Journey, SOA Institute, 2006 [Online]. Available: http://www.soainstitute.org/articles/article/article/soa-maturity-model-compass-on-the-soa-journey.html (Accessed: 31-Oct-2011).

[Benatallah'03], Benatallah Boualem, Dumas Marlon, Fauvet Marie-Christine, et al. Towards patterns of web services composition. In: Patterns and skeletons for parallel and distributed computing, UK: Gorlatch S., Rabhi F. Eds., Springer-Verlag, 2003 pp. 265-296. ISBN 1852335068.

[Benjamin'09], Benjamin P., Patki M., Mayer R. Using Ontologies for Simulation Modeling. In: Proc. of the 38th conf. on Winter Simulation (WSC'06), 2006, Monterey, CA, USA: IEEE Press, pp. 1151-1159.

[Berardi'04], Berardi D., Calvanese D., Giacomo G.D., et al. Synthesis of composite e-Services based on automated reasoning. In: Proc. of the 2nd intern. conf. on Service oriented computing, 2004, New York, USA: ACM Press, pp. 105-114.

[Beuscart'09], Beuscart R., McNair P., Darmoni S.F., et al. Patient Safety: Detection and Prevention of Adverse Drug Events. In: Adlassnig K.P., Blobel B., Mantas J., Masic I. Eds. Medical Informatics in a United and Healthy Europe – Proc. of the 22nd intern, congress of the European Federation for Medical Informatics, 2009, Sarajevo, Bosnia and Herzegovina, Studies in Health Technology and Informatics: IOS Press, pp. 968-971.

[Black'04], Black J.P., Segmuller W., Cohen N., et al. Pervasive Computing in Health Care: Smart Spaces and Enterprise Information Systems. In: Proc of the 2nd intern. MobiSys Workshop on Context Awareness, 2004, Boston, USA: Citeseer, pp. 1-6.

[Bloomberg'06], Bloomberg Jason, Schmelzer Ronald. Service Orient or Be Doomed!: How Service Orientation Will Change Your Business. 1st ed.: Wiley, 2006, 272 p. ISBN 9780471768586

[Bonner'72], Bonner R.E., Crevasse L., Ferrer M.I., et al. A new computer program for analysis of scalar electrocardiograms. Computers and Biomedical Research, 1972, vol. 5, no. 6, pp. 629-653.

[Borromeo'07], Borromeo S., Rodriguez-Sanchez C., Machado F., et al. A reconfigurable, wearable, wireless ECG system. In: Proc. of 29th annual intern. conf. on Engineering in Medicine and Biology Society (EMBC), 2007, Lyon, France: IEEE, pp. 1659-1662.

[Bortolan'93], Bortolan G., Willems J.L. Diagnostic ECG classification based on neural networks. Journal of Electrocardiology, 1993, suppl. 26, pp. 75-79.

[BS'03], British Standard. Medical electrical equipment. Particular requirements for safety. Particular requirements for safety, including essential performance, of recording and analysing single channel and multichannel electrocardiographs. BS EN 60601-2-51: BSI, 2003, 86 p. ISBN 0580427285

[BS'04], British Standard. Size Designation of Clothes. Measurements and intervals. BS EN 13402-3: BSI, 2004, 32 p. ISBN 0580 455718.

[Brodnick'00], Brodnick D. A method to locate electrode placement. Journal of Electrocardiology, 2000, suppl. 33, pp. 211-218.

[Callens'08], Callens L., Carrault G., Cordier M.O., et al. Intelligent adaptive monitoring for cardiac surveillance. In. Proc. of the 18th European conf. on Artificial Intelligence, 2008, Patras, Greece: IOS Press, pp. 653-657.

[Cardoso'03], Cardoso J., Sheth A. Semantic E-Workflow Composition. Journal of Intelligent Information Systems, 2003, vol. 21, no. 3, pp. 191-225.

[Castelli'08], Castelli G., Mamei M., Zambonelli F. Engineering contextual knowledge for autonomic pervasive services. Journal of Information and Software Technology, 2008, vol. 50, no. 1, pp. 36-50.

[Cauvet'08], Cauvet C., Guzelian G. Business Process Modeling: A Service-Oriented Approach. In: Proc. of the 41st annual intern. conf. on System Sciences (HICSS2008), 2008, Waikoloa, HI, USA: IEEE, pp. 98-98.

[CDC'11], Centers for disease control and prevention. Chronic Disease Overview [Online]. Available: http://www.cdc.gov/chronicdisease/overview/index.htm. (Accessed: 10-Sep-2011).

[Chappell'02], Chappell David A., Jawell Tyler. Java Web Services. 1st ed.: O'Reilly Media, Inc., 2002, 276 p. ISBN 9780596002695

[Chen'03], Chen H., Finin T., Joshi A. An ontology for context-aware pervasive computing environments. The Knowledge Engineering Review, 2003, vol. 18, pp. 197–207.

[Ciuksys'07], Ciuksys D., Caplinskas A. Reusing Ontological Knowledge about Business Processes in IS Engineering: Process Configuration Problem. Journal of Informatica, 2007, vol. 18, no. 4, pp. 585-602.

[Cornfield'73], Cornfield J., Dunn R.A., Batchlor C.D., et al. Multigroup diagnosis of electrocardiograms. Computers and Biomedical Research, 1973, vol. 6, no. 1, pp. 97-120.

[Czarnecki'00], Czarnecki Krysztof, Eisenecker Ulrich. Generative Programming: Methods, Tools, and Applications. 1st ed.: Addison-Wesley Professional, 2000, 864 p. ISBN 9780201309775

[De Luca'04], De Luca G., Suryapranata H., Ottervanger J.P., et al. Time Delay to Treatment and Mortality in Primary Angioplasty for Acute Myocardial Infarction. Circulation, 2004, vol. 109, no. 10, pp. 1223 -1225.

[Debenham'89], Debenham John K. Knowledge Systems Design. Australia: Prentice-Hall, 1989, 583 p. ISBN 9780135171295

[Rienzo'07], Di Rienzo M., Rizzo F., Meriggi P., et al. MagIC: a textile system for vital signs monitoring. Advancement in design and embedded intelligence for daily life applications. In: Proc. of the 29th annual intern. conf. on Engineering in Medicine and Biology Society (EMBS), 2007, Lyon, France: IEEE, pp. 3958-3961.

[Donnelly'06], Donnelly M., Nugent C.D., Finlay D., et al. Optimal electrode placement for the identification of old MI and LVH. In: Proc. of the 33rd intern. conf. on Computers in Cardiology, 2006, Valencia, Spain: IEEE, pp. 437-440. ISBN 9781424425327.

[Dower'84], Dower G.E. The ECGD: a derivation of the ECG from VCG leads. J Electrocardiol, 1984, vol. 17, no. 2, pp. 189-191.

[Dzenowagis'05], Dzenowagis J. Connecting for health: global vision, local insight. Report for the World Summit on the Information Society. World Health Organization 2005 [Online]. Available: http://www.who.int/kms/resources/WSISReport_Connecting_for_Health.pdf. (Accessed: 30-Sept-2011).

[Endsley'00], Endsley Mica R. Theoretical underpinnings of situation awareness: A critical review. In: Situation awareness analysis and measurement, 1st ed., Mahwah, USA: Endsley M.R., Garland D.J. Eds., CRC Press, 2000, pp. 3–32. ISBN 9780805821345

[Fahkri'11], Fahkri M., Cauvet C. Méthodes d'ingénierie de systèmes d'information à base de services. Ingénierie des Systèmes d'Information, 2011, vol. 16, no. 2, pp. 43-66.

[Fahy'08], Fahy C., Davy S., Boudjemil SZ., et al. Towards an Information Model That Supports Service-Aware, Self-Managing Virtual Resources. In: van der Meer S., Burgess M., Denazi S.G. Eds. Lecture Notes in Computer Science on Modelling Autonomic Communications Environment, Proc. of the 3rd intern. workshop on Modelling Autonomic Communications Environments (MACE2008), 2008, Samos Island, Greece: Springer, pp. 102-107.

[Fayn'11], Fayn J. A Classification Tree Approach for Cardiac Ischemia Detection Using Spatiotemporal Information From Three Standard ECG Leads. IEEE Trans Biomed Eng, 2011, vol. 58, no. 1, pp. 95-102.

[Fayn'07], Fayn J., Macfarlane P.W., Rubel P. Can the lessons learned from the assessment of automated electrocardiogram analysis in the Common Standards for quantitative Electrocardiology study benefit measurement of delayed contrast-enhanced magnetic resonance images? J Electrocardiol, 2007, vol. 40, no. 3, pp. 246-250.

[Fayn'88], Fayn J., Rubel P. CAVIAR: a serial ECG processing system for the comparative analysis of VCGs and their interpretation with auto-reference to the patient. Journal of Electrocardiology, 1988, vol. 21, pp. 173-176.

[Fayn'07a], Fayn J., Rubel P., Pahlm O., et al. Improvement of the detection of myocardial ischemia thanks to information technologies. Int J Cardiol, 2007, vol. 120, no. 2, pp. 172-80.

[Fayn'10], Fayn J., Rubel P. Toward a Personal Health Society in Cardiology. IEEE Transactions on Information Technology in Biomedicine: A Publication of the IEEE Engineering in Medicine and Biology Society, 2010, vol. 14, pp. 409-409.

[Feuerlicht'09], Feuerlicht G. Design of Composable Services. In: Proc. of intern. workshop on Service-Oriented Computing, 2008, Sydney, Australia: Springer-Verlag, pp. 15-27.

[Feild'02], Field D.Q., Feldman C.L., Horacek B.M. Improved EASI coefficients: their derivation, values, and performance. Journal of Electrocardiology, 2002, vol. 35, pp. 23-33.

[Finlay'08], Finlay D., Nugent C.D., Donnelly M.P., et al. Optimal electrocardiographic lead systems: practical scenarios in smart clothing and wearable health systems. IEEE Transactions on Information Technology in Biomedicine, 2008, vol. 12, no. 4, pp. 433-41.

[Flach'96], Flach J.M Situation awareness: In search of meaning. CSERIAC Gateway, 1996, vol. 6, no. 6, pp. 1–4.

[Forrester'11], Forrester E.C., Buteau B.L., Shrum S. CMMI for Services: Guidelines for Superior Service, 2nd ed.: Addison-esley, 2011, 800 p. ISBN 0-321-71152-1.

[Gadomski'09], Gadomski A.M, Zimny T.A. Application of IPK (Information, Preferences, Knowledge) Paradigm for the Modelling of Precautionary Principle Based Decision-Making. In: Lecture Notes in Computer Science on Critical Information Infrastructure Security: Springer, 2009, vol. 5508, pp. 319–327.

[Gadomski'89], Gadomski A.M. From Know-How to How-to-Know: An approach to knowledge ordering for Specification of complex problems (TOGA methodology). In: Materials of intern. symposium on Computational Intelligence, 1989, Milan, Italy.

[Garcia'98], Garcia J., Lander P., Sornmo L., et al. Comparative study of local and Karhunen-Loève-based ST-T indexes in recordings from human subjects with induced myocardial ischemia. Computers and Biomedical Research, 1998, vol. 31, no. 4, pp. 271-292.

[Gartner'10], Gartner: Gartner Reveals Five Business Process Management Predictions for 2010 and Beyond. 2010, UK [Online]. Available: http://www.gartner.com/it/page.jsp?id=1278415. (Accessed: 10-Jan-2011)

[Garvey'06], Garvey J.L., MacLeod B.A., Sopko G., et al. Pre-hospital 12-lead electrocardiography programs: a call for implementation by emergency medical services systems providing advanced life support--National Heart Attack Alert Program (NHAAP) Coordinating Committee; National Heart, Lung, and Blood Institute (NHLBI); National Institutes of Health. Journal of the American College of Cardiology, 2006, vol. 47, no. 3, pp. 485-491.

[Gemperle'98], Gemperle F., Kasabach C., Stivortic J., et al. Design for wearability. In: Proc of the 2nd intern. symposium on Wearable Computers, 1998, Pittsburgh, PA, USA: IEEE, pp. 116-122.

[GTWM'96], Georgia Tech Wearable Motherboard. The Intelligent Garment for the 21st Century [Online]. Available: http://www.gtwm.gatech.edu/ (Accessed: 30-Sep-2011).

[Gibbons'10], Gibbons Michael C., Bali Rajeev, Nilmini Wickramasinghe. Perspectives of Knowledge Management in Urban Health (Healthcare Delivery in the Information Age). 1st ed.: Springer, 2010, 231 p. ISBN 9781441956439.

[Goodson'06], Goodson E. Technology, Healthcare, and You. HealthNex, 2006 [Online]. Available: httphttp://healthnex.typepad.com/web_log/2006/06/personal_health.html (Accessed: 31-Oct-2011).

[Green'07], Green M., Ohlsson M., Forberg J.L., et al. Best leads in the standard electrocardiogram for the emergency detection of acute coronary syndrome. J Electrocardiol, 2007, vol. 40, no. 3, pp. 251-256.

[Gripay'08], Gripay Y., Laforest F., Petit J., et al. Management of data sources and services in pervasive environments. In : Proc. of the 8th intern. conf. on NOuvelles TEchnologies de la REpartition - NOTERE, 2008, Lyon, France: ACM press, pp. 34-39.

[Gruninger'05], Gruninger M., Hull R., McIlraith S. A First-Order Ontology for Semantic Web Services. In: Position paper at W3C Workshop on Frameworks for Semantic Web Services, 2005 [Online]. Available: http://www.w3.org/2005/04/FSWS/Submissions/59/w3c05.pdf. (Accessed 12-Jan-2011)

[Gulrajani'88], Gulrajani R.M. The forward and inverse problems of electrocardiography. Engineering in Medicine and Biology Magazine, 1998, vol. 17, no. 5, pp. 84–101.

[Hadzievski'04], Hadzievski L., Bojovic B., Vukcevic V., et al. A novel mobile transtelephonic system with synthesized 12-lead ECG. Information Technology in Biomedicine, 2004, vol. 8, no. 4, pp. 428-438

[Han'09], Han K.H., Park J.W. Process-centered knowledge model and enterprise ontology for the development of knowledge management system. Journal of Expert Systems with Applications, 2009, vol. 36, no. 4, pp. 7441-7447.

[Havey'05], Havey Michael. Essential business process modeling. 1st ed.: O'Reilly Media, Inc., 2005, 352 p. ISBN 9780596008437

[Healy'06], Healy J.C. Editorial of the special health and the Internet for all. International Journal of Medical Informatics, 2006, vol. 75, no. 1, pp. 5-7.

[Herbsleb'97], Herbsleb J., Zubrow D., Goldenson D., et al. Software Quality and the Capability Maturity Model. Communications of the ACM, 1997, vol. 40, no. 6, pp. 30-40.

[Herman'91], Herman M.V., Ingram D.A., Levy J.A., et al. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clinical Cardiology, 1991, vol. 14, no. 6, pp. 469-476.

[Hernandez'09], Hernandez A.I, Le Rolle V., Defontaine A., et al. A multiformalism and multiresolution modelling environment: application to the cardiovascular system and its regulation. Philos Transact A Math Phys Eng Sci, 2009, vol. 367, no. 1908, pp. 4923-4940.

[Horacek'00], Horacek B.M., Warren J.W., P. Stovicek P., et al. Diagnostic accuracy of derived versus standard 12-lead electrocardiograms. Journal of Electrocardiology, 2000, vol. 33, pp. 155-160.

[Horacek'02], Horacek B.M., Warren J.W., Penney C.J., et al. Optimal electrocardiographic leads for detecting acute myocardial ischemia. Journal of Electrocardiology, 2001, vol. 34, pp. 97-111.

[Horacek'08], Horacek B.M., Warren J.W., Wang J.J. On designing and testing transformations for derivation of standard 12-lead/18-lead electrocardiograms and vectorcardiograms from reduced sets of predictor leads. Journal of Electrocardiology, 2008, vol. 41, no. 3, pp. 220-229.

[Horan'78], Horan L.G. The quest for optimal electrocardiography. American Journal of Cardiology, 1978, vol. 41, no. 1, pp. 126-129.

[Hu'73], Hu K.C., Francis D.B., Gau G.T., et al. Development and performance of Mayo-IBM electrocardiographic computer analysis programs (V70). Mayo Clinic, 1973, vol. 48, no. 4, pp. 260-268.

[Hurst'07], Hurst A., Hudson S.E., Mankoff J. Dynamic Detection of Novice vs. Skilled Use Without a Task Model. In: Proc. of the SIGCHI conference on Human factors in computing systems, 2007, San Jose, California, USA: ACM Press, pp. 271-280.

[ISO'05], ISO/DIS 11073-91064.2007 and EN 1064:2005. Medical Informatics—Standard Communication Protocol—Computer-assisted Electrocardiography (SCP-ECG). ISO document, 2005 [Online]. Available: http://www.centc251.org/ (Accessed: 31-Oct-2011).

[ISO'03], ISO/IEC 15504: Process assessment. ISO document, 2003-2011 [Online]. Available: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=38932 (Accessed: 31-Oct-2011).

[Jafari'05], Jafari R., Dabiri F., Brisk P., et al. Adaptive and fault tolerant medical vest for life-critical medical monitoring. In: Proc. of the 2005 ACM symposium on Applied computing, 2005, Santa Fe, New Mexico, USA: ACM Press, pp. 272-279.

[Jardim-Goncalves'07], Jardim-Goncalves Ricardo, Muller Jorg P., Mertins, et al. Eds. Enterprise Interoperability II: New Challenges and Approaches, 1st ed.: Springer, 2007, 912 p. ISBN 9781846288579

[Josuttis'07], Josuttis Nicolai M. SOA in Practice: The Art of Distributed System Design. 1st ed.: O'Reilly Media, Inc., 2007, 352 p. ISBN 9780596529550.

[Kadish'01], Kadish A.H., Buxton A.E., Kennedy H.L., et al. ACC/AHA clinical competence statement on electrocardiography and ambulatory electrocardiography: a report of the ACC/AHA/ACP-ASIM Task Force on Clinical Competence (ACC/AHA Committee to Develop a Clinical Competence Statement on Electrocardiography and Ambulatory Electrocardiography). Circulation, 2001, vol. 104, pp. 3169–3178.

[Kearney'07], Kearney K., Thomas C., McAdams E. Quantification of motion artifact in ECG electrode design. In: Proc. of 29th annual intern. conf. on Engineering in Medicine and Biology Society (EMBC), 2007, Lyon, France: IEEE, pp. 1533-1536.

[Klug'07], Klug T., Muhlhauser M. Modeling human interaction resources to support the design of wearable multimodal systems. In: Proc. of the 9th intern. conf. on Multimodal interfaces, 2007, Nagoya, Aichi, Japan: ACM Press, pp. 299-306.

[Klusmeier'76], Klusmeier S., Klinkers H., FriedrichsP., et al. Comparison of different algorithms for P-onset and P-offset recognition in ECG programs. Advances in Cardiology, 1976, vol. 16, pp. 221-226.

[Knight'05], Knight F., Schwirtz A., Psomadelis F., et al. The design of the SensVest. Personal Ubiquitous Comput., 2005, vol. 9, no. 1, pp. 6-19.

[Ko'08], Ko J.M., Kim C.O., Kwon I. Quality-of-service oriented web service composition algorithm and planning architecture. Journal of Systems and Software, 2008, vol. 81, no. 11, pp. 2079-2090.

[Kocaballi'07] Kocaballi A.B., Kocyigit A. Granular best match algorithm for context-aware computing systems. Journal of Systems and Software, 2007, vol. 80, no. 12, pp. 2015-2024.

[Kokar'09], Kokar M.M., Matheus C.J., Baclawski K. Ontology-based situation awareness. Journal on Information Fusion, 2009, vol. 10, no. 1, pp. 83-98.

[Komatsoulis'08], Komatsoulis G.A., Warzel D.B., Hartel F.W., et al. caCORE version 3: Implementation of a model driven, service-oriented architecture for semantic interoperability. J. of Biomedical Informatics, 2008, vol. 41, no. 1, pp. 106-123.

[Kors'86], Kors J.A., Talmon J.L., van Bemmel J.H. Multilead ECG analysis. Computers and Biomedical Research, 1986, vol. 19, no. 1, pp. 28-46

[Kors'02], Kors J.A., van Herpen G. How many electrodes and where? A "poldermodel" for electrocardiography. Journal of Electrocardiology, 2002, Suppl., no 35, p. 7-12.

[Koutkias'09], Koutkias V., Lazou K., Kilintzis V., et al. On Intelligent Procedures in Medication for Patient Safety: The PSIP Approach. In: Proc. of the 9th intern. conf. on Intelligent Systems Design and Applications, 2009, Pisa, Italy: IEEE Computer Society, pp. 363-366.

[Krupaviciute'09], Krupaviciute A., Fayn F., Rubel P., et al. Information System Architecture for Wearable Cardiac Sensors Personalization. In: Proc. of 14th intern. conf. on Engineering of Complex Computer Systems, 2009 Potsdam, Germany: IEEE press, pp. 265-272.

[Krupaviciute'10], Krupaviciute A., Fayn J. Expert Activities Automation through Enhanced Business Services Orchestration. In: Babar M.A., Gorton I. Eds. Lecture Notes in Computer Science on Software Architecture, 2010, Copenhagen, Denmark: Springer Berlin / Heidelberg, pp. 497-500.

[Krupaviciute'10a], Krupaviciute A., Fayn J. Ontology driven approach enhancing services orchestration. In: Proc. of the 2nd intern. conf. on Computational Intelligence, Communication Systems and Networks, 2010, Liverpool, UK: IEEE press, pp. 344-348.

[Krupaviciute'10b], Krupaviciute A., Fayn J., McAdams E., et al. Personal sensor-system modalities evaluation for simplified ECG recording in self-care. In: Proc. of 37th intern. conf. on Computing in Cardiology, 2010, Belfast, NI, UK: IEEE press, pp. 541-544.

[Kurlkarni'07], Kulkarni P., Ozturk Y. Requirements and design spaces of mobile medical care. SIGMOBILE Mob. Comput. Commun. Rev., 2007, vol. 11, no. 3, pp. 12-30.

[Lateef'03], Lateef F., Nimbkar N., Da Z.K., et al. Vertical displacement of the precordial leads alters electrocardiographic morphology. Indian Heart Journal, 2003, vol. 55, no. 4, pp. 339-43.

[Laukkanen'04], Laukkanen M., Helin H. Composing Workflows of Semantic Web Services. In: Extending Web Services Technologies, New York: Springer-Verlag, 2004, pp. 209-228.

[Lee'08], Lee J., Wyner G., Pentland B. Process Grammar as a Tool for Business Process Design. Journal of MIS Quarterly, 2008, vol. 32, no. 4, pp. 757-778.

[Lee'00], Lee T.H., Goldman L., Evaluation of the patient with acute chest pain. New England Journal of Medicine, 2000, vol. 342, no. 16, pp. 1187-1195.

[Leonardi'07], Leonardi G., Panzarasa S., Quaglini S., et al. Interacting agents through a web-based health serviceflow management system. Journal of Biomedical Informatics, 2007, vol. 40, no. 5, pp. 486-499.

[Leyking'07], Leyking K. Service-oriented Knowledge Architectures - Integrating Learning and Business Information Systems. In: Proc. of the EC-TEL 2007 PROLEARN Doctoral Consortium, 2007, Crete, Greece: CEUR-WS, pp. 34-42.

[Liebowitz'01], Liebowitz Jay. Knowledge Management: Learning from Knowledge Engineering. 1st ed.: CRC Press Inc, 2001, 152 p. ISBN 9780849310249

[Lin'02], Lin Tsau Y., Yao Yiyu, Zadeh Lotfi A. Data mining, rough sets, and granular computing. 1st ed: Physica-Verlag HD, 2002, 536 p. ISBN 9783790814613

[Luepker'05], Luepker R.V. Delay in acute myocardial infarction: why don’t they come to the hospital more quickly and what can we do to reduce delay?. American Heart Journal, 2005, vol. 150, no. 3, pp. 368-370.

[Luprano'06], Luprano J. European Projects in Smart Fabrics, Interactive Textiles: sharing Opportunities and Challenges. In: Workshop on Wearable Technology and Intelligent Textiles, 2006 [Online]. Available: http://csnej106.csem.ch/sfit/pdf/eu_sfit_projects_opportunities_and_challenges.pdf (Accessed: 31-Oct-2011).

[Luprano'06a], Luprano J., Sola J., Dasen S., et al. Combination of body sensor networks and on-body signal processing algorithms: the practical case of MyHeart project. In: Proc. of intern. workshop on Wearable and Implantable Body Sensor Networks, 2006, Cambridge, MA, USA: IEEE, pp. 75-79.

[Lymberis'06], Lymberis A., Gatzoulis L. Wearable Health Systems: from smart technologies to real applications. In: Proc. of the 28th annual intern. conf. on Engineering in Medicine and Biology Society (EMBC), 2006, New York, NY, USA: IEEE, pp. 6789-6792.

[Lymberis'08], Lymberis A., Paradiso R. Smart Fabrics and Interactive Textile Enabling Wearable Personal Applications: R&D State of the Art and Future Challenges. In: Proc. of 30th annual intern. conf. on Engineering in Medicine and Biology Society (EMBC), 2008, Vancouver, BC: IEEE, pp. 5270-5273.

[Macfarlane'03], Macfarlane P.W., Colaco R., Stevens K., et al. Precordial electrode placement in women. Netherlands Heart Journal, 2003, vol. 11, no. 3, pp. 118-122.

[Macfarlane'89], Macfarlane Peter W. Comprehensive Electrocardiology: Vols 1-3: Theory and Practice in Health and Disease, 1st ed. New York: Lawrie Weitch T.D., Pergamon Press, 1989, 1900 p. ISBN 9780080355689

[Macfarlane'10], Macfarlane Peter W., van Oosterom Adriaan, Pahlm Olle, et al. Comprehensive Electrocardiology. 2nd ed. : John Camm, Springer, 2010, 2310 p. ISBN 9781848820456

[Mackey'04], Mackey Judith, Mensah George A., Mendis Shanthi., et al. The atlas of heart disease and stroke. World Health Organization, 2004, 112 p. ISBN 9789241562768

[Macris'09], Macris A., Papakonstantinou D., Malamateniou F., et al. Using ontology-based knowledge networks for user training in managing healthcare processes. International Journal of Technology Management, 2009, vol. 47, no. 1, pp. 5-21.

[Maglaveras'05], Maglaveras N., Chouvarda i., Koutkias V.G., et al. The Citizen Health System (CHS): a modular medical contact center providing quality telemedicine services. IEEE Transactions on Information Technology in Biomedicine, 2005, vol. 9, no. 3, pp. 353-362.

[Manzi'10], Manzi de Arantes W., Verdier C. Defining quality-measurable medical alerts from incomplete data through fuzzy linguistic variables and modifiers. IEEE Transactions on Information Technology in Biomedicine, 2010, vol. 14, no. 4, pp. 916-922.

[Martin'04], Martin T., Jones M., Edmison J., et al. Modeling and simulating electronic textile applications. In: Proc. of 2004 ACM SIGPLAN/SIGBED conference on Languages, compilers, and tools for embedded systems, 2004, Washington, DC, USA: ACM Press, vol. 39, no 7, pp. 10-19.

[Martin'00], Martin T., Jovanov E, Raskovic D. Issues in wearable computing for medical monitoring applications: a case study of a wearable ECG monitoring device. In: Proc. of 4th intern. symposium on Wearable Computers, 2000, Atlanta, USA: IEEE Comput. Soc, pp. 43-49.

[Martinez'07], Martinez J.P., Laguna P., Olmos S., et al. Assessment of QT-measurement accuracy using the 12-lead electrocardiogram derived from EASI leads. Journal of Electrocardiology, 2007, vol. 40, no. 2, pp. 172-179.

[Martinez-Costa'09], Martinez-Costa C., Menarguez-Tortosa M., Fernandez-Breis J.T., et al. A model-driven approach for representing clinical archetypes for Semantic Web environments. Journal of Biomedical Informatics, 2009, vol. 42, no. 1, pp. 150-164.

[Mathers'08], Mathers Colin, Fat Doris M., Boerma J.T. The global burden of disease: 2004 update. World Health Organization, 2008, 146 p. ISBN 9789241563710

[Matthews'07], Matthews R., McDonald N., Hervieux P., et al. A Wearable Physiological Sensor Suite for Unobtrusive Monitoring of Physiological and Cognitive State. In: Proc. of 29th annual intern. conf. on Engineering in Medicine and Biology Society (EMBC), 2007, Lyon, France: IEEE, pp. 5276-5281.

[McAdams'96], McAdams E., Jossinet J., Lackermeier A., et al. Factors affecting electrode-gel-skin interface impedance in electrical impedance tomography. Medical & biological engineering & computing, 1996, vol. 34, no. 6, pp. 397-408.

[McAdams'11], McAdams E., Krupaviciute A., Gehin C., et al. Wearable Sensor Systems: The Challenges. In: Proc. of 33rd annual intern. conf. on the Engineering in Medicine and Biology Society (EMBC), 2010, Boston, USA: IEEE press, pp. in press

[McAdams'94], McAdams E., Lackermeier A., Jossinet J. AC impedance of the hydrogel-skin interface," presented at the Engineering in Medicine and Biology Society. In: Proc. of the 16th annual intern. conf. on Engineering Advances: New Opportunities for Biomedical Engineers, 1994, Baltimore, USA: IEEE, vol. 2, pp. 870-871. ISBN 0780320506

[McAdams'11a], McAdams Eric., Krupaviciute Asta., Gehin Claudine, et al. Wearable Electronic Systems: Applications to Medical Diagnostics/Monitoring. In: Wearable Monitoring Systems, 1st ed., New York: Annalisa Bonfiglio and Danilo Emilio de Rossi, Springer, 2011, pp. 179-203. ISBN 9781441973832

[McCann'07], McCann K., Holdgate A., Mahammad R., et al. Accuracy of ECG electrode placement by emergency department clinicians. Emergency Medicine Australasia, 2007, vol. 19, no. 5, pp. 442-448.

[McGinn'05], McGinn A.P., Rosamond W.D., Goff D.C, et al. Trends in prehospital delay time and use of emergency medical services for acute myocardial infarction: experience in 4 US communities from 1987-2000. American Heart Journal, 2005, vol. 150, no. 3, pp. 392-400.

[MDA'11], Model Driven Architecture Home Page. OMG, 2011 [Online]. Available: http://www.omg.org/mda/ (Accessed: 31-Oct-2011).

[Mestrovic'07], Mestrovic M., Helmer R., Kyratzis L., et al. Preliminary study of dry knitted fabric electrodes for physiological monitoring. In: Proc. of the 3rd intern. Conf. on Intelligent Sensors, Sensor Networks and Information, 2007, Melbourne, Australia: IEEE, pp. 601-606.

[Michelson'06], Michelson B.M. Event-driven architecture overview. Event-Driven SOA Is Just Part of the EDA Story. Patricia Seybold Group, 2006 [Online]. Available: http://www.omg.org/soa/Uploaded%20Docs/EDA/bda2-2-06cc.pdf. (Accessed 20-Apr-2011)

[Mirvis'07], Mirvis D.M., Pahlm O., Rautaharju P., et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology. A scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology. J Am Coll Cardiol, 2007, vol. 49, no. 10, pp. 1109-27.

[Modre'06], Modre R., Seger M., Fischer G., et al. Cardiac anisotropy: is it negligible regarding noninvasive activation time imaging?. IEEE Transactions on Biomedical Engineering, 2006, vol. 53, no. 4, pp. 569–580.

[Moore'88], Moore R.L, Rosenhof H., Stanley G. Process control using the G2 Real-Time Expert System. In: Conference Record of the 1989 IEEE Industry Applications Society Annual Meeting, 1989, San Diego, CA, USA: IEEE, vol. 2, pp. 1452–1456.

[Morlet'90], Morlet D., Rubel P., Willems J.L. Value of scatter-graphs for the assessment of ECG computer measurement results. Methods of Information in Medicine, 1990, vol. 29, no. 4, pp. 413-423.

[Mortara'77], Mortara D.W. Digital filters for ECG signals. In: Proc. of the 4th intern. conf. on Computers in Cardiology, 1977, Rotterdam, The Netherlands: Ostrow H.G., Ripley K.L. Eds., IEEE Comput Soc, 1977, pp. 511-514.

[Moulton'98], Moulton D. Space Age T-Shirt Senses Vital Signs, Wound Location. The Medical Post, 1998, vol. 34, no. 33, pp. 63-63.

[Mundt'05], Mundt C.W., Montgomery K.N., Udoh U.E., et al. A multiparameter wearable physiologic monitoring system for space and terrestrial applications. IEEE Transactions on Information Technology in Biomedicine, 2005, vol. 9, no. 3, pp. 382-391.

[Mustapha'06], Mustapha S.M.F.D.S. Relation-Based Case Retrieval Approach for Web Services Selection. In: Proc. of the intern. conf on Web Intelligence, 2006, Honk Kong : IEEE, pp. 644-648.

[Naone'09], Naone E. MIT Technology Review 2009 Special Report on 10 emerging technologies. 2009. [Online]. Available: http://www.technologyreview.com/read_article.aspx?ch=specialsections&sc=tr10&id=22117. (Accessed 12-Jan-2011).

[Nelwan'08], Nelwan S.P., Finlay D., van Dam T.B., et al. Improved 12-lead ECG reconstruction from lead sub sets by dynamic selection of frontal leads. In Proc. of 35th intern. conf. on Computers in Cardiology, 2008, Bologna, Italy: IEEE, pp. 213–216.

[Nelwan'04], Nelwan S.P., Kors J.A., Meij S.H., et al. Reconstruction of the 12-lead electrocardiogram from reduced lead sets. Journal of Electrocardiology, 2004, vol. 37, no. 1, pp. 11-18.

[Nonaka'95], Nonaka Ikujiro, Takeuchi Hirotaka. The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. USA: Oxford University Press, 1995, 304 p. ISBN 978019509269

[OMG'10], OMG. Business Process Model and Notation (BPMN), version 2.0. OMG document, 2010 [Online]. Available: http://www.omg.org/cgi-bin/doc?dtc/10-06-04. (Accessed: 10-Jan-2011).

[Orszag'08], Orszag P. Congressional Budget Office Testimony: Growth in Health Care Costs. Delivered before the Committee on the Budget, United States Senate, January 31, 2008 [Online]. Available: http://www.cbo.gov/ftpdocs/89xx/doc8948/01-31-HealthTestimony.pdf. (Accessed: 10-Sep-2011).

[Laguna'96], P. Laguna P., Jane R., Olmos S., et al. Adaptive estimation of QRS complex wave features of ECG signal by the Hermite model. Medical and Biological Engineering and Computing, 1996, vol. 34, no. 1, pp. 58–68.

[Pacelli'06], Pacelli M., Loriga G., Taccini N., et al. Sensing Fabrics for Monitoring Physiological and Biomechanical Variables: E-textile solutions. In: Proc. of the 3rd International Summer School and Symposium on Medical Devices and Biosensors, 2006, Cambridge, MA, USA: IEEE, pp. 1-4.

[Pahl'07], Pahl C. Semantic model-driven architecting of service-based software systems. Journal of Information and Software Technology, 2007, vol. 49, no. 8, pp. 838-850.

[Paluska'08], Paluska J.M, Pham H., Saif U., et al. Structured decomposition of adaptive applications. Journal of Pervasive and Mobile Computing, 2008, vol. 4, no. 6, pp. 791-806.

[Pan'08], Pan J., Chen K., Hsu W. Self risk assessment and monitoring for cardiovascular disease patients based on service-oriented architecture. In: Proc. of the 35th intern. conf. on Computers in Cardiology, 2008, Bologna, Italy: IEEE, pp. 637-640.

[Papouchado'87], Papouchado M., Walker P.R., James M.A., et al. Fundamental differences between the standard 12–lead electrocardiograph and the modified (Mason—Likar) exercise lead system. European Heart Journal, 1987, vol. 8, no. 7, pp. 725-733.

[Paradiso'05], Paradiso R., Belloc C., Loriga G., et al. Wearable healthcare system, new frontier on e-textile. Studies in Health Technology and Informatics, IOS Press, 2005, vol. 117, pp. 105-113.

[Park'06], Park C., Chou P.H., Bai Y., et al. An ultra-wearable, wireless, low power ECG monitoring system. In: Proc. of conf on Biomedical Circuits and Systems, 2006, London, UK: IEEE, pp. 241–244.

[Pawar'07], Pawar T., Chaudhuri S., Duttagupta S.P. Body movement activity recognition for ambulatory cardiac monitoring. IEEE Transactions on Biomedical Engineering, 2007, vol. 54, no. 5, pp. 874–882.

[Pedrycz'07], Pedrycz W. Granular computing-the emerging paradigm. Uncertain Systems, 2007, vol. 1, no. 1, pp. 38–61.

[Peltz'03], Peltz C. Web services orchestration and choreography. Computer, 2003, vol. 36, no. 10, pp. 46-52.

[Phillips-Wren'09], Phillips-Wren G., Mora M., Forgionne G., et al. An integrative evaluation framework for intelligent decision support systems. European Journal of Operational Research, 2009, vol. 195, no. 3, pp. 642-652.

[Pipberger'61], Pipberger H.V., Arms R.J., Stallmann F.W. Automatic Screening of Normal and Abnormal Electrocardiograms by Means of a Digital Electronic Computer. Experimental Biology and Medicine, 1961, vol. 106, no. 1, p. 130.

[Pola'07], Pola T., Vanhala J. Textile Electrodes in ECG Measurement. In: Proc. of the 3rd intern. Conf. on Intelligent Sensors, Sensor Networks and Information, 2007, Melbourne, Australia: IEEE, pp. 635-639.

[Pollehn'02], Pollehn T., Brady W., Perron A., et al. The electrocardiographic differential diagnosis of ST segment depression. Emergency Medicine Journal, 2002, vol. 19, no. 2, pp. 129-135.

[Preuveneers'04], Preuveneers D., Van den Bergh J., Wagelaar D., et al. Towards an extensible context ontology for ambient intelligence. Ambient Intelligence, 2004, vol. 3295, pp. 148–159.

[Protégé'11], Protégé Home Page. Stanford Center for Biomedical Informatics Research, 2011 [Online]. Available: http://protege.stanford.edu/ (Accessed: 31-Oct-2011).

[Rajaganeshan'08], Rajaganeshan R., Ludlam C.L., Francis D.P., et al. Accuracy in ECG lead placement among technicians, nurses, general physicians and cardiologists. International Journal of Clinical Practice, 2008, vol. 62, no. 1, pp. 65-70.

[Ralyte'08], Ralyte J., Jeusfeld M.A., Backlund P., et al. A knowledge-based approach to manage information systems interoperability. Information Systems, 2008, vol. 33, no. 7, pp. 754-784.

[Rasmussen'94], Rasmussen Jens, Pejtersen Annelise M., Goodstein L.P. Cognitive Systems Engineering, 1st ed.: Wiley-Interscience, 1994, 396 p. ISBN 9780471011989

[Rautaharju'98], Rautaharju P.M., Park L., Rautaharju F.S., et al. A standardized procedure for locating and documenting ECG chest electrode positions: Consideration of the effect of breast tissue on ECG amplitudes in women. Journal of Electrocardiology, 1998, vol. 31, no. 1, pp. 17-29.

[Rautaharju'81], Rautaharju P.M., Warren J.W., Jain U., et al. Cardiac infarction injury score: an electrocardiographic coding scheme for ischemic heart disease. Circulation, 1981, vol. 64, no. 2, pp. 249-256.

[Robinson'08], Robinson J., Wakeman I., Chalmers D. Composing software services in the pervasive computing environment: Languages or APIs?. Journal of Pervasive and Mobile Computing, 2008, vol. 4, no. 4, pp. 481-505.

[Romhilt'68], Romhilt D.W., Estes E.H. A point-score system for the ECG diagnosis of left ventricular hypertrophy. American Heart Journal, 1968, vol. 75, no. 6, pp. 752-758

[Rubel'86], Rubel P., Ayad B. The true boundary recognition power of multidimensional detection functions. An optimal comparison. In: Computer ECG analysis: towards standardization. Amsterdam, The Netherlands: Willems J.L., van Bemmel J.H., Ziwietz C., Elsevier Sciences Publishers, 1986, pp. 97–103.

[Rubel'86a], Rubel P., Fayn J., Arnaud P., et al. New strategies in serial VCG comparison and trend analysis. In: Willems J.L., Van Bemmel J.H., and Zywietz C, Eds., Computer ECG Analysis: Towards Standardization, Amsterdam: North- Holland, 1986, pp. 323-328.

[Rubel'91], Rubel P., Benhadid I., Fayn J. Quantitative assessment of 8 different methods for synthesizing Frank VCGs from simultaneously recorded standard ECG leads. J Electrocardiol, 1991, suppl. 24, pp. 197-202.

[Rubel'05], Rubel P., Fayn J., Nollo G., et al. Toward personal eHealth in cardiology. Results from the EPI-MEDICS telemedicine project. Journal of Electrocardiology, 2005, vol. 38, no. 4, pp. 100-106.

[Sakkopoulos'09], Sakkopoulos E., Adamopoulou P., Tsakalidis A.K., et al. Personalized selection of web services for mobile environments: the m-scroutz solution. In: Proc. of the intern. conf. on Management of Emergent Digital EcoSystems, 2009, Lyon, France: ACM, pp. 254-260.

[Sasa'08], Sasa A., Juric M., Krisper M. Service-Oriented Framework for Human Task Support and Automation. IEEE Transactions on Industrial Informatics, 2008, vol. 4, no. 4, pp. 292-302.

[Scherer'89], Scherer J.A., Jenkins J.M., Nicklas J.M. Synthesis of the 12-lead electrocardiogram from a 3-lead subset using patient-specific transformation vectors. An algorithmic approach to computerized signal synthesis. Journal of Electrocardiology, 1989, vol. 22, pp. 128-136.

[Scherer'92], Scherer J.A., Rubel P., Fayn J., et al. Quantitative assessment of 12-lead ECG synthesis using CAVIAR. Journal of Electrocardiology, 1992, vol. 25, no. 1, pp. 137-142.

[Schijvenaars'01], Schijvenaars B.J., Kors J.A., van Herpen G., et al. Employment of intra-individual variability to improve computerized ECG interpretation. Studies in health technology and informatics, 2001, vol. 84 (Pt 1), pp. 513–517.

[Schijvenaars'97], Schijvenaars R.J., Kors J.A., van Herpen G., et al. Effect of electrode positioning on ECG interpretation by computer. Journal of Electrocardiology, 1997, vol. 30, pp. 247-256.

[Schijvenaars'96], Schijvenaars R.J., Kors J.A., van Herpen G., et al. Use of the standard 12-lead ECG to simulate electrode displacements. Journal of Electrocardiology, 1996, suppl. 29, pp. 5-9.

[Scilingo'05], Scilingo E.P., Gemignani A., Paradiso R., et al. Performance evaluation of sensing fabrics for monitoring physiological and biomechanical variables. IEEE Transactions on Information Technology in Biomedicine, 2005, vol. 9, no. 3, pp. 345-352.

[Sejersen'06], Sejersten M., Pahlm O., Pettersson J., et al. Comparison of EASI-derived 12-lead electrocardiograms versus paramedic-acquired 12-lead electrocardiograms using Mason-Likar limb lead configuration in patients with chest pain. Journal of Electrocardiology, 2006, vol. 39, pp. 13-21.

[Sejersten'07], Sejersten M., Wagner G.S., Pahlm O., et al. Detection of acute ischemia from the EASI-derived 12-lead electrocardiogram and from the 12-lead electrocardiogram acquired in clinical practice. Journal of Electrocardiology, 2007, vol. 40, no. 2, pp. 120-126.

[Simon-Chautemps'04], Simon-Chautemps L., Adami S., Arod S., et al. System Revisions, EPI-MEDICS IST-2000-26164 European Project Deliverable 7.1, 2004, 139 p.

[Simonet'11], Simonet A. Automatic Production of an Operational Information System from a Domain Ontology Enriched with Behavioral Properties. In: Bellatreche L., Pinto F.M. Eds. Lecture Notes in Computer Science on Model and Data Engineering – Proc. of the 1st intern. conf. on Model and Data Engineering 2011, Obidos, Portugal: Springer, pp. 4-17.

[Sirin'03], Sirin E., Hendler J., Parsia B. Semi-automatic Composition of Web Services using Semantic Descriptions. In: Proc. of modelling, architecture and infrastructure workshop on Web Services, 2003, Angers, France: ICEIS Press, pp. 17-24.

[Sohrabi'09], Sohrabi S., Prokoshyna N., McIlraith S. Web Service Composition via the Customization of Golog Programs with User Preferences. In: Borgida A., Chaudhri V.K., Giorgini P., Yu E.S.K. Eds. Lecture Notes in Computer Science on Conceptual Modeling: Foundations and Applications, 2009: Springer Berlin / Heiderberg, pp. 319-334.

[Sowa'1999], Sowa John F. Knowledge Representation: Logical, Philosophical and Computational Foundations. 1st ed., Pacific Grove, CA, USA: Brooks Cole Publishing Co., 1999, 512 p. ISBN 9780534949655

[Soylu'09], Soylu A., Causmaecker P.D., Desmet P. Context and Adaptivity in Pervasive Computing Environments: Links with Software Engineering and Ontological Engineering. Journal of Software, 2009, vol. 4, no. 9, pp. 992–1013.

[Stallmann'61], Stallmann F.W., Pipberger H.V. Automatic recognition of electrocardiographic waves by digital computer. Circulation Research, 1961, vol. 9, no. 6, pp. 1138–1143.

[Strassner'07], Strassner J., O'Sullivan D., Lewis D. Ontologies in the Engineering of Management and Autonomic Systems: A Reality Check. Journal of Network and Systems Management, 2007, vol. 15, no. 1, pp. 5-11.

[Strassner'09], Strassner J., Souza J.N., Raymer D., et al. The design of a novel context-aware policy model to support machine-based learning and reasoning. Cluster Computing, 2009, vol. 12, no. 1, pp. 17-43.

[Such'07], Such O. Motion tolerance in wearable sensors--the challenge of motion artifact. In: Proc. of 29th annual intern. conf. on Engineering in Medicine and Biology Society (EMBC), 2007, Lyon, France: IEEE, pp. 1542-1545.

[Sullivan'07], Sullivan T., Deiss S., Cauwenberghs G. A Low-Noise, Non-Contact EEG/ECG Sensor. In: Proc. of conf. on Biomedical Circuits and Systems, 2007, Montreal, Canada: IEEE, pp. 154-157.

[Swenne'73], Swenne C.A., van Bemmel J.H., Hengeveld S.J., et al. Pattern recognition for ECG-monitoring: An interactive method for the classification of ventricular complexes. Computers and Biomedical Research, 1973, vol. 6, no. 2, pp. 150–160.

[Swenson'10], Swenson Keith D. Mastering the Unpredictable: How Adaptive Case Management Will Revolutionize the Way That Knowledge Workers Get Things Done. 1st ed.: Meghan-Kiffer Press, 2010, 354 p. ISBN 9780929652122

[Taccardi'63], Taccardi B. Distribution of Heart Potentials on the Thoracic Surface of Normal Human Subjects. Circ Res, 1963, vol. 12, no. 4, pp. 341-352.

[TFR'98], Task Force Report. The pre-hospital management of acute heart attacks. Recommendations of a Task Force of The European Society of Cardiology and The European Resuscitation Council. European Heart Journal, 1998, vol. 19, no. 8, pp. 1140-1164.

[Temal'08], Temal L., Dojat M., Kassel G., et al. Towards an ontology for sharing medical images and regions of interest in neuroimaging. Journal of Biomedical Informatics, 2008, vol. 41, no. 5, pp. 766-778.

[Teo'06], Teo C.B., Gay R.K.L. A knowledge-driven model to personalize e-learning. Journal on Educational Resources in Computing, 2006, vol. 6, no. 1, pp. 3-18.

[Tighiouart'03], Tighiouart B., Rubel P., Bedda M. Improvement of QRS boundary recognition by means of unsupervised learning. In: Proc of the 30th intern. conf. on Computers in Cardiology, 2003, Thessaloniki, Greece: IEEE, pp. 49-52.

[Tisseyre'99], Tisseyre René-Charles. Knowledge management: Théorie et pratique de la gestion des connaissances. : Hermes Sciences Publicat., 1999, 185 p. ISBN 9782746200692

[Traverso'04], Traverso P., Pistore M. Automated composition of semantic web services into executable processes. In: McIlraith S.A., Plexousakis D., van Harmelen F. Eds. Lecture Notes in Computer Science of 3rd intern. conf. on Semantic Web, 2004, Hiroshima, Japan: Springer, pp. 380–394.

[Tsai'08], Tsai W.T., Sun X., Huang Q., et al. An ontology-based collaborative service-oriented simulation framework with Microsoft Robotics Studio. Journal of Simulation Modelling Practice and Theory, 2008, vol. 16, no. 9, pp. 1392-1414.

[Umar'09], Umar A., Zordan A. Enterprise Ontologies for Planning and Integration of Business: A Pragmatic Approach. IEEE Transactions on Engineering Management, 2009, vol. 56, no. 2, pp. 352-371.

[Bemmel'90], van Bemmel J.H., Willems J.L Standardization and validation of medical decision-support systems: the CSE project. Methods of Information in Medicine, 1990, vol. 29, no. 4, pp. 261-262.

[Varshney'07], Varshney U. Pervasive healthcare and wireless health monitoring. Mob. Netw. Appl., 2007, vol. 12, no. 2, pp. 113-127.

[Weigand'09], Weigand H., Johannesson P., Andersson B., et al. Value-Based Service Modeling and Design: Toward a Unified View of Services. In: Proc. of the 21st intern. conf. on Advanced Information Systems Engineering, 2009, Amsterdam, The Netherlands: Springer-Verlag Berlin, pp. 410-424.

[Wenger'96], Wenger W., Kligfield P. Variability of precordial electrode placement during routine electrocardiography. Journal of Electrocardiology, 1996, vol. 29, no. 3, pp. 179-184.

[Wijesiriwardana'04], Wijesiriwardana R., Mitcham K., Dias T. Fibre-meshed transducers based real time wearable physiological information monitoring system. In: Proc. of the 8th intern, symposium on Wearable Computers, 2004, Arlington, Virginia, USA: IEEE, pp. 40-47.

[Willems'91], Willems J.L., Abreu-Lima C., Arnaud P., et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. New England Journal of Medicine, 1991, vol. 325, no. 25, pp. 1767-1773.

[Willems'86], Willems J.L. A review of computer ECG analysis: time to evaluate and standardize. CRC Critical Reviews in Medical Informatics, 1986, vol. 1, no. 2, pp. 165-207.

[Willems'90], Willems J.L., Arnaud P., van Bemmel J.H., et al. Common standards for quantitative electrocardiography: goals and main results. Methods of information in medicine, 1990 vol. 29, no. 4, pp. 263-271.

[Willems'72], Willems J.L., Pipberger H.V. Arrhythmia detection by digital computer. Computers and Biomedical Research, 1972, vol. 5, no. 3, pp. 273-278.

[Wilson'34], Wilson F.N, Johnston F.D, Macleod A.G., et al. Electrocardiograms that represent the potential variations of a single electrode. American Heart Journal, 1934, vol. 9, no. 4, pp. 447-458.

[Wolf'72], Wolf H., MacInnis P., Stock S., et al. Computer analysis of rest and exercise electrocardiograms. Computers and Biomedical Research, 1972, vol. 5, no. 4, pp. 329-346.

[Wright'08], Wright A., Sittig D.F. SANDS: a service-oriented architecture for clinical decision support in a National Health Information Network. Journal of Biomedical Informatics, 2008, vol. 41, no. 6, pp. 962-981.

[Zaunseder'07], Zaunseder S., Fischer W., Netz S., et al. Prolonged Wearable ECG Monitoring - a Wavelet Based Approach. In: Proc. of the 6th conf. on Sensors, 2007, Atlanta, GA, USA: IEEE, pp. 1197-1200.

[Zipes'06], Zipes D.P, Camm A.J., Borggrefe M., et al. ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. Circulation, 2006, vol. 114, pp. 1088-1132.

[Zywietz'90a], Zywietz C., van Bemmel J.H., Degani R. Evaluation of ECG interpretation systems: Signal analysis. Meth. Inform. Med, 1990, vol. 29, no. 4, pp. 298-307.

[Zywietz'90b], Zywietz C., Willems J.L., Arnaud P., et al. Stability of computer ECG amplitude measurements in the presence of noise. Computers and Biomedical Research, 1990, vol. 23, no. 1, pp. 10–31

Table of figures

Fig.I.2_1 ECG beat corresponding to one electrical cycle 27
Fig.I.2_2 Electrodes positioning for: a) standard 12-lead b) Mason-Likar 12-lead and c) 3-lead ECGs 28
Fig.I.2_3 Electrodes positioning for: a) EASI 3-lead ECG b) Frank XYZ-lead VCG *[Macfarlane'89]* c) BSPM. 32
Fig.I.2_4 Body artefacts sources affecting the ECG signal's quality. 36
Fig.I.3_1 Examples of wearable sensor-systems: a) Wealthy suit, b) Philips HeartCycle t-shirt, c) VivoMetrics Vivo Responder belt, d) Exmovere BT2 watch (Images a) and b) are extracted from *[Goodson'06]*, *[Luprano'06]* and c) and d) reprinted from *[McAdams'11]* with permission of the authors) 47
Fig.I.4_1 a) a circle type element b) electrode displacement map 61
Fig.I.4_2 Electrodes' positioning while moving a set of 4 electrodes 62
Fig.I.4_3 Design of the sensor system modalities 64
Fig.I.4_4 EPI-MEDICS project based sensor-system placement at Mason-Likar positioning for 8 subjects superposed at V2 and a brief description of each of the subject 66
Fig.I.4_5 Two different models for the calculation of patient-specific transformations. Here S1 indicates a standard 12-lead ECG and ML a Mason-Likar 12-lead ECG 68
Fig.I.4_6 Optimal correlation (Cor) and root mean square (RMS) values computation method for typical beats of the synthesised and the reference ECGs comparison 72
Fig.I.4_7 Normalised comparison results between a 12-lead Mason-Likar positioning ECG (ML-ECG) and 3 simplified ECGs for 8 healthy volunteers: 4 women and 4 men 74
Fig.I.4_8 Normalised comparison results between a first standard 12-lead ECG (S1-ECG) and 3 simplified ECGs for 8 healthy volunteers: 4 women and 4 men 77
Fig.I.4_9 Normalised comparison results between the second standard 12-lead ECG (S2-ECG) and the three 12-lead ECGs derived from the 3 simplified ECGs for 8 healthy volunteers: 4 women and 4 men. 80
Fig.I.4_10 Normalised comparison results between a second standard 12-lead ECG (S2-ECG) and 3 sets of synthesised 12-lead ECGs reconstructed from 3 simplified ECGs recorded at a few weeks interval for the 1st subject 84
Fig.I.4_11 Normalised comparison results between a second standard 12-lead ECG (S2-ECG) and 3 sets of synthesised 12-lead ECGs reconstructed from 3 simplified ECGs recorded at a few weeks interval for the 3rd subject 84
Fig.I.4_12 Normalised comparison results between a second standard 12-lead ECG (S2-ECG) and 3 sets of synthesised 12-lead ECGs reconstructed from 3 simplified ECGs recorded at a few weeks interval for the 4th subject 84
Fig.II.6_1 Approach of data and model driven services orchestration 124
Fig.II.6_2 Process knowledge modelling schema 127
Fig.II.6_3 Schematic dynamic business process activity 128
Fig.II.6_4 Symbolic representation of orchestrated atomic and composite services 129
Fig.II.6_5 Core-ontology empowering decision making services orchestration 131
Fig.II.7_1 Data and model driven Service Oriented Architecture 139
Fig.II.7_2 Decision support system process knowledge ontology empowering the management of dynamic decision-making business process 142
Fig.II.7_3 Schematic ontologies for each knowledge model 144
Fig.II.7_4 Data and model driven service oriented architecture for decision-making support 145
Fig.II.7_5 Domain-specific data and model driven service oriented architecture (DM-SOA) 147
Fig.II.8_1 Use case diagram of Personal Cardio Assistant system (PCAs) for an anomaly suspicion scenario ("p" stands for personal) 154
Fig.II.8_2 Class diagram of the PCAs input data 154
Fig.II.8_3 Use case diagram of Personal Cardio Assistant system (PCAs) for a follow-up scenario 157
Fig.II.8_4 Use case diagram of Personal Cardio Assistant system (PCAs) for an emergency scenario 158
Fig.II.8_5 PCA system class diagram 159
Fig.II.8_6 Choreography of the optimal personal ECG recording system selection. Legend: PEM – ECGs registered into the Personal ECG Monitor; T.characterisation – transformation characterisation; Multi-grid composition – assessment descriptors multi-grid composition; Intersignal descr. comp. – intersignal descriptors computation 160
Fig.II.8_7 Schema of a dynamic business process for optimal personal ECG recording system se- 161

lection . . . .
Fig.II.8_8 ECG features extraction process 162
Fig.II.8_9 Transformation characterisation process ST stands for 12-lead ECG and ML for Mason-Likar electrodes positioning based ECG 164
Fig.II.8_10 Intersignal descriptors computation process 164
Fig.II.8_11 ECG synthesis process 165
Fig.II.8_12 ECG features extraction process 167
Fig.II.8_13 Decision making process 168
Fig.II.8_14 Example of size selection 169
Fig.A.2.1_1 Schematic heart's electrical system 197
Fig.A.4.3_1 Normalised comparison results between a second standard 12-lead ECG (S2-ECG), not used for the synthesis models determination, and the 3 sets of 12-lead ECGs reconstructed from the 3 simplified ECGs recorded at 3 different dates, for the 1st subject 204
Fig.A.4.3_2 Normalised comparison results between a second standard 12-lead ECG (S2-ECG), not used for the synthesis models determination, and the 3 sets of 12-lead ECGs reconstructed from the 3 simplified ECGs recorded at 3 different dates, for the 3rd subject 205
Fig.A.4.3_3 Normalised comparison results between a second standard 12-lead ECG (S2-ECG), not used for the synthesis models determination, and the 3 sets of 12-lead ECGs reconstructed from the 3 simplified ECGs recorded at 3 different dates, for the 4th subject 206

List of tables

Table I.3.1 A brief overview on *continuous* and on *on demand* monitoring types via 5 common monitoring steps . . . 57
Table I.4.1 Four stripes measurements . . . 64
Table I.4.2 Five distances taken among electrodes placed in their anatomically correct positions for each subject . . . 66
Table I.4.3 The seven final descriptors used to compare original and synthesized ECGs . . . 71
Table I.4.4 C1: optimal sensor-system modality and ECG synthesis method suggestion for each subject . . . 76
Table I.4.5 C2: optimal sensor-system modality and data synthesis method suggestion for each subject . . . 79
Table I.4.6 C3: optimal sensor-system modality and data synthesis method suggestion for each subject . . . 82
Table I.4.7 Paired t-tests results taking into account all descriptors, f – female, m – male . . . 82
Table I.4.8 Paired t-test results for 3 descriptors, f – female, m – male . . . 83
Table I.4.9 Paired t-test results for the 3 sets of data per subject . . . 85
Table II.6.1 Summary of procedure-based activities and expert activities overview . . . 123
Table A.4.1.1 Interval and duration of measurements (in milliseconds). The first cut indicates the mean ± standard deviation of 265 males ECGs in that age range. The second cut takes the 96% range . . . 199
Table A.4.1.2 Results (in milliseconds) for the global ECG intervals measurements of the 1st subject for 22 different ECG records obtained with slightly different electrodes positions (measurements marked in bold and underlined go out of given limits) . . . 199
Table A.4.1.3 Results for the 2nd subject (measurements marked in bold and underlined go out of given limits) . . . 200
Table A.4.1.4 Amplitude of the P, Q, R, S, T+ and T- waves measurements (in microvolts). The first cut indicates the mean ± standard deviation of 265 males ECGs in that age range. The second cut takes the 96% range . . . 201
Table A.4.1.5 Lead II amplitude measurement results (in microvolts) for the 1st subject (measurements marked in bold and underlined go out of given limits) . . . 201
Table A.4.1.6 Lead II measurements results for the 2nd subject (measurements marked in bold and underlined go out of given limits) . . . 202
Table A.4.2.1 Values of the seven descriptors that were obtained in the context of the C1 study in function of the S, M and L sensor-system size modalities using two data reconstruction methods: artificial neural nets (ANN) and linear regression (Regression). r – correlation, M – RMS value, K – QRS peak-to-peak amplitude, p – risk of an acute Myocardial Infarction . . . 203

Annexes

Annex 2.1 – Principles of the heart's electrical system

Electrical heart system

The heart's electrical system includes three main parts (Fig.A.2.1_1): the SA node (sinoatrial node) - located in the right atrium of the heart, the AV node (atrioventricular node) - in an area between the atria and the ventricles, and the His-Purkinje system, which fibres are located in the inner ventricular walls of the heart.

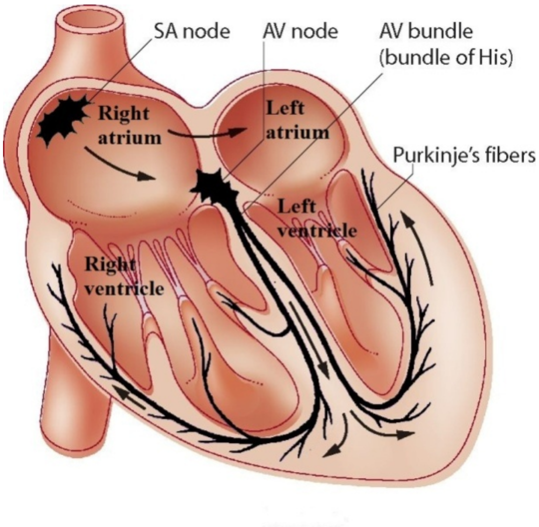


Fig.A.2.1_1 Schematic heart's electrical system

Each cardiac cycle, normally, starts in a bundle of specialized cells in the right atrium – SA node, which creates the electrical signal and, therefore, is often called a natural pacemaker. Typically, 60 to 80 electrical impulses are created per minute while resting and define our heart rate. The created electrical impulse is propagated across both atria stimulating their myocardium to contract and to push blood into the ventricles. This conduction is seen in the **P wave**, which assessment gives an idea of where the atrial depolarisation begins and whether or not the atria are enlarged. When the electrical signal reaches the AV node it pauses giving the time for ventricles to be filled with blood, otherwise they would contract at the same time as atria and so the blood would not be pumped effectively. This pause appears due to the AV node, which acts like an electrical bridge between the atria and the ventricles, where except in rare conditions, no other cells between the atria and ventricles allow the electricity flow. This delay of electricity in AV node forms much of **PR segment** on the ECG, which length could explain some types of slow heart rhythms (bradycardias). Afterwards, the signal reaches the His-Purkinje system, which spreads electrical signals throughout the ventricles to make them contract and to push blood out to lungs and body. This electricity flow through the ventricular myocardium produces the **QRS complex** on the ECG. Ventricles muscle mass is higher than the atria, therefore the QRS complex amplitude is larger than the P wave and as the conduction of Purkinje fibres is fast, it results to the rapid cells depolarisation, what is reflected into the short and "sharp" QRS complex. Clinically, the complex may indicate a ventricular hypertrophy (the increase in the volume of organ/tissue/cell), though patients' weight and age should be taken in account. QRS complex correlates a little with patients' weight and age, it tends to be smaller for higher weight or older patients. Also, observation of changes in QRS complex through simultaneously recorded ECG leads can reflect infarctions, or some injury. Finally, the cardiac cycle terminates with the repolarisation of ventricles. The **ST segment** occurs when the ventricular depolarisation has ended and the repolarisation has not yet begun. This segment is normally going slightly upwards and is of a particular interest for cardiac disease identification. The flat or down sloping ST segment may indicate coronary ischemia, the elevation of ST segment can be related to myocardial infarction, whilst the depression of ST segment may be associated with hypokalemia (abnormally low level of potassium) or poisoning due to excessive doses of digoxin (digitalis) – a plant based ingredient used for drug preparation. When the repolarisation starts, the **T wave** (and occasionally the **U wave**) occurs on the ECG. Since repolarisation is slower than depolarisation, the T wave is broad and has a slow upstroke. The T wave can be observed for indications of ischemia or electrolyte disturbances. The U wave, which may be seen in some leads, is associated with metabolic disturbances, typically, hypokalemia and hypomagnesemia (abnormally low level of magnesium).

This, even very schematic, explanation on electrical heart cycle demonstrates how well the ECG reflects both the structure of the heart and of its electrical system and of the ionic behaviour of the cells.

Annex 4.1 – Initial simulation results

Global measurement comparison

Normal limits for 18-29 years old Caucasian males for global ECG measurements *[Macfarlane'89 vol. 3, page 1442]*:

Parameter / Cut	Pdur		QRSdur		QT		QTc_Baz	
	from	till	from	till	from	till	from	till
1st cut	*88.80*	*117.20*	*87.80*	*105.00*	*356.60*	*414.40*	*390.80*	*437.00*
2nd cut	***72.00***	***128.00***	***80.00***	***114.00***	***336.00***	***442.00***	***370.00***	***463.00***

Table A.4.1.1 Interval and duration of measurements (in milliseconds). The first cut indicates the mean ± standard deviation of 265 males ECGs in that age range. The second cut takes the 96% range

ECG	ID	Pdur	QRSdur	QT	QTc_Baz
StandardECG	**2**	*112*	*88*	*366*	*397.9*
ReferenceECG	**3**	*92*	*88*	***354***	***380.8***
Rec_20000	**4**	***88***	***86***	*360*	***372.5***
Rec_40000	**5**	***88***	***86***	*362*	***<u>367.9</u>***
Rec_60000	**6**	*92*	***86***	*372*	*394.3*
Rec_80000	**7**	***84***	***86***	*366*	***384.5***
Rec_02000	**8**	*112*	***80***	*360*	***375.9***
Rec_04000	**9**	***88***	***86***	*366*	*402.7*
Rec_06000	**10**	*96*	***84***	*372*	*399.7*
Rec_08000	**11**	*104*	***84***	*362*	*395.9*
Rec_00200	**12**	***76***	***84***	*372*	*390.8*
Rec_00400	**13**	*96*	***86***	*364*	***384.1***
Rec_00600	**14**	*100*	***86***	*374*	***390.3***
Rec_00800	**15**	***120***	***82***	*368*	*394.3*
Rec_46082	**16**	***84***	*90*	*374*	*398.0*
Rec_82046	**17**	***80***	***84***	*378*	*399.3*
Rec_28064	**18**	*108*	***86***	*374*	***381.5***
Rec_64028	**19**	*96*	*88*	*372*	***386.4***
Rec_37037	**20**	*108*	*88*	*382*	*407.7*
Rec_73073	**21**	*92*	***84***	*372*	***378.5***
Rec_11055	**22**	*104*	***84***	*366*	***377.3***
Rec_55011	**23**	*100*	***86***	*370*	*412.4*

Table A.4.1.2 Results (in milliseconds) for the global ECG intervals measurements of the 1st subject for 22 different ECG records obtained with slightly different electrodes positions (measurements marked in bold and underlined go out of given limits)

ECG	ID	Pdur	QRSdur	QT	QTc_Baz
StandardECG	25	96	**108**	374	430.1
ReferenceECG	26	**124**	102	380	405.8
Rec_20000	27	116	102	374	403.8
Rec_40000	28	**120**	100	370	419.5
Rec_60000	29	108	100	380	409.8
Rec_80000	30	**124**	**106**	368	413.8
Rec_02000	31	92	104	386	416.5
Rec_04000	32	92	**106**	376	404.7
Rec_06000	33	112	102	380	426.7
Rec_00200	35	**120**	**106**	376	404.7
Rec_00400	36	112	100	374	413.5
Rec_00600	37	112	**106**	372	415.4
Rec_00800	38	**128**	100	374	435.4
Rec_46082	39	112	104	382	405.6
Rec_82046	40	**160**	104	372	420.1
Rec_28064	41	**140**	102	374	433.6
Rec_64028	42	**136**	98	382	436.5
Rec_37037	43	**140**	**108**	384	**454.8**
Rec_73073	44	**128**	100	368	435.5
Rec_11055	45	**116**	102	366	409.5
Rec_55011	46	112	94	386	436.2

Table A.4.1.3 Results for the 2nd subject (measurements marked in bold and underlined go out of given limits)

Lead-by-lead measurements comparison

Normal limits for 18-29 years old Caucasian males for lead-by-lead ECG measurements *[Macfarlane'89, vol. 3, pp. 1446-57]*:

Parameter / Lead	Qamp		Ramp		Samp		T+amp		T-amp	
	from	till	from	till	from	till	from	till	from	till
I (1st cut)	-96	20	468	1097	-366	-96	236	458	-	-
I (2nd cut)	**-223**	**0**	**206**	**1495**	**-526**	**0**	**141**	**609**	**-**	**-**
II (1st cut)	-156	-34	975	1857	-348	-92	298	576	-	-
II (2nd cut)	**-240**	**0**	**577**	**2394**	**-496**	**0**	**179**	**766**	**-**	**-**
III (1st cut)	-251	-27	223	1393	-999	-487	48	248	-162	-22
III (2nd cut)	**-370**	**0**	**35**	**2167**	**-999**	**0**	**0**	**406**	**-238**	**0**
aVR (1st cut)	-1194	-722	14	186	-1364	-796	-	-	-495	-285
aVR (2nd cut)	**-1288**	**0**	**0**	**357**	**-1617**	**0**	**-**	**-**	**-628**	**-185**
aVL (1st cut)	-173	-19	36	600	-670	-128	66	242	-76	-12
aVL (2nd cut)	**-267**	**0**	**28**	**1243**	**-1053**	**0**	**0**	**364**	**-86**	**0**
aVF (1st cut)	-163	-37	560	1608	-346	-38	153	395	-59	-17
aVF (2nd cut)	**-239**	**0**	**43**	**2211**	**446**	**0**	**49**	**565**	**-78**	**0**
V1 (1st cut)	-1339	0	188	626	-1872	-854	58	360	-170	-48
V1 (2nd cut)	**-1339**	**0**	**77**	**1019**	**-2723**	**510**	**0**	**552**	**-121**	**0**

V2 (1st cut)	-2556	156	465	1269	-3071	-1479	630	1244	-6	0
V2 (2nd cut)	**-2716**	**0**	**189**	**1989**	**-4082**	**771**	**367**	**1596**	**-6**	**0**
V3 (1st (cut)	-112	-18	571	1629	-2027	-639	627	1191	-	-
V3 (2nd cut)	**-187**	**0**	**348**	**2512**	**-2953**	**157**	**341**	**1524**	**-**	**-**
V4 (1st cut)	-177	-11	1319	2745	-1041	-225	455	1021	-	-
V4 (2nd cut)	**285**	**0**	**800**	**3481**	**-1780**	**0**	**248**	**1428**	**-**	**-**
V5 (1st cut)	-187	-27	1337	2577	-585	-73	342	816	-	-
V5 (2nd cut)	**-338**	**0**	**1018**	**3560**	**-966**	**0**	**177**	**1142**	**-**	**-**
V6 (1st cut)	-187	-41	1135	2143	-348	-46	270	620	-31	-31
V6 (2nd cut)	**-302**	**0**	**784**	**3005**	**-602**	**0**	**147**	**914**	**-31**	**-31**

Table A.4.1.4 Amplitude of the P, Q, R, S, T+ and T- waves measurements (in microvolts). The first cut indicates the mean ± standard deviation of 265 males ECGs in that age range. The second cut takes the 96% range

ECG	ID	Lead Name	Q amplitude	R amplitude	S amplitude	T + amplitude
StandardECG	2	II	-83	1017	**-43**	**173**
ReferenceECG	3	II	-145	1339		**194**
Rec_20000	4	II	-105	1363	**-47**	**252**
Rec_40000	5	II	-166	1465		**219**
Rec_60000	6	II	-143	1331		**169**
Rec_80000	7	II	-105	1290		**188**
Rec_02000	8	II	-108	1344		**205**
Rec_04000	9	II	-124	1329		**190**
Rec_06000	10	II	-125	1378		**219**
Rec_08000	11	II	-142	1380		**203**
Rec_00200	12	II	-102	1335		**222**
Rec_00400	13	II	-105	1355		**234**
Rec_00600	14	II	-105	1373		**223**
Rec_00800	15	II	-122	1372		**217**
Rec_46082	16	II	-157	1431		**209**
Rec_82046	17	II	-86	1256	**-57**	**194**
Rec_28064	18	II	-123	1328		**200**
Rec_64028	19	II	-125	1324		**208**
Rec_37037	20	II	-150	1328		**201**
Rec_73073	21	II	-121	1305		**215**
Rec_11055	22	II	-121	1320	**-39**	**215**
Rec_55011	23	II	-167	1372		**175**

Table A.4.1.5 Lead II amplitude measurement results (in microvolts) for the 1st subject (measurements marked in bold and underlined go out of given limits)

ECG	ID	Lead Name	Q amplitude	R amplitude	S amplitude	T + amplitude
StandardECG	25	II	**-178**	1048	-199	**202**
ReferenceECG	26	II	**-248**	1324	-276	**257**
Rec_20000	27	II	**-243**	1391	-269	**271**
Rec_40000	28	II	**-271**	1382	-266	**246**
Rec_60000	29	II	**-225**	1363	-265	**276**
Rec_80000	30	II	**-222**	1319	-256	**230**
Rec_02000	31	II	**-235**	1338	-272	**252**
Rec_04000	32	II	**-240**	1331	-258	**245**
Rec_06000	33	II	**-242**	1341	-258	**227**
Rec_00200	35	II	**-245**	1315	-287	**225**
Rec_00400	36	II	**-236**	1321	-281	**233**
Rec_00600	37	II	**-227**	1318	-276	**245**
Rec_00800	38	II	**-246**	1291	-293	**207**
Rec_46082	39	II	**-246**	1348	-300	**265**
Rec_82046	40	II	**-226**	1279	-265	**200**
Rec_28064	41	II	**-248**	1354	-264	**195**
Rec_64028	42	II	**-232**	1327	-312	**256**
Rec_37037	43	II	**-252**	1327	-294	**233**
Rec_73073	44	II	**-229**	1335	-273	**212**
Rec_11055	45	II	**-238**	1287	-243	**193**
Rec_55011	46	II	**-246**	1332	-313	**243**

Table A.4.1.6 Lead II measurements results for the 2nd subject (measurements marked in bold and underlined go out of given limits)

N.B. Due to the thesis manuscript limitations, the data concerning the other leads as well as the data and evaluation results of the synthesised leads are available on demand.

Annex 4.2 – Example of a subject specific data grid

1st subject descriptors' profile (before scaling)

Subject 1				
C1		S	M	L
ANN	1. $1-\underset{\{leads\}}{Median}(r)$	*0.003*	*0.012*	*0.014*
	2. $\underset{\{leads\}}{Median}(M)$	*30.2*	*59.8*	*62.4*
	3. $\sum_{\{leads\}}(1-r)$	*0.043*	*0.099*	*0.108*
	4. $\sum_{\{leads\}} M$	*252.9*	*438.1*	*463.8*
	5. Δp	*0.05*	*0.02*	*0.01*
	6. ΔQRSaxis	*0*	*0*	*2*
	7. $\Delta \underset{\{leads\}}{Median}(K)$	*4.7*	*29.6*	*58.3*
Regression	1. $1-\underset{\{leads\}}{Median}(r)$	*0.007*	*0.018*	*0.022*
	2. $\underset{\{leads\}}{Median}(M)$	*57.4*	*73.4*	*77.8*
	3. $\sum_{\{leads\}}(1-r)$	*0.155*	*0.217*	*0.224*
	4. $\sum_{\{leads\}} M$	*496.8*	*611.9*	*626.0*
	5. Δp	*0.135*	*0.122*	*0.215*
	6. ΔQRSaxis	*0*	*0*	*2*
	7. $\Delta \underset{\{leads\}}{Median}(K)$	*5.7*	*42.6*	*156.5*

Table A.4.2.1 Values of the seven descriptors that were obtained in the context of the C1 study in function of the S, M and L sensor-system size modalities using two data reconstruction methods: artificial neural nets (ANN) and linear regression (Regression). r – correlation, M – RMS value, K – QRS peak-to-peak amplitude, p – risk of an acute Myocardial Infarction

N.B. Due to the thesis manuscript limitations, the data concerning the other leads as well as the data and evaluation results of the synthesised leads are available on demand.

Annex 4.3 – Repeatability test results

1st subject

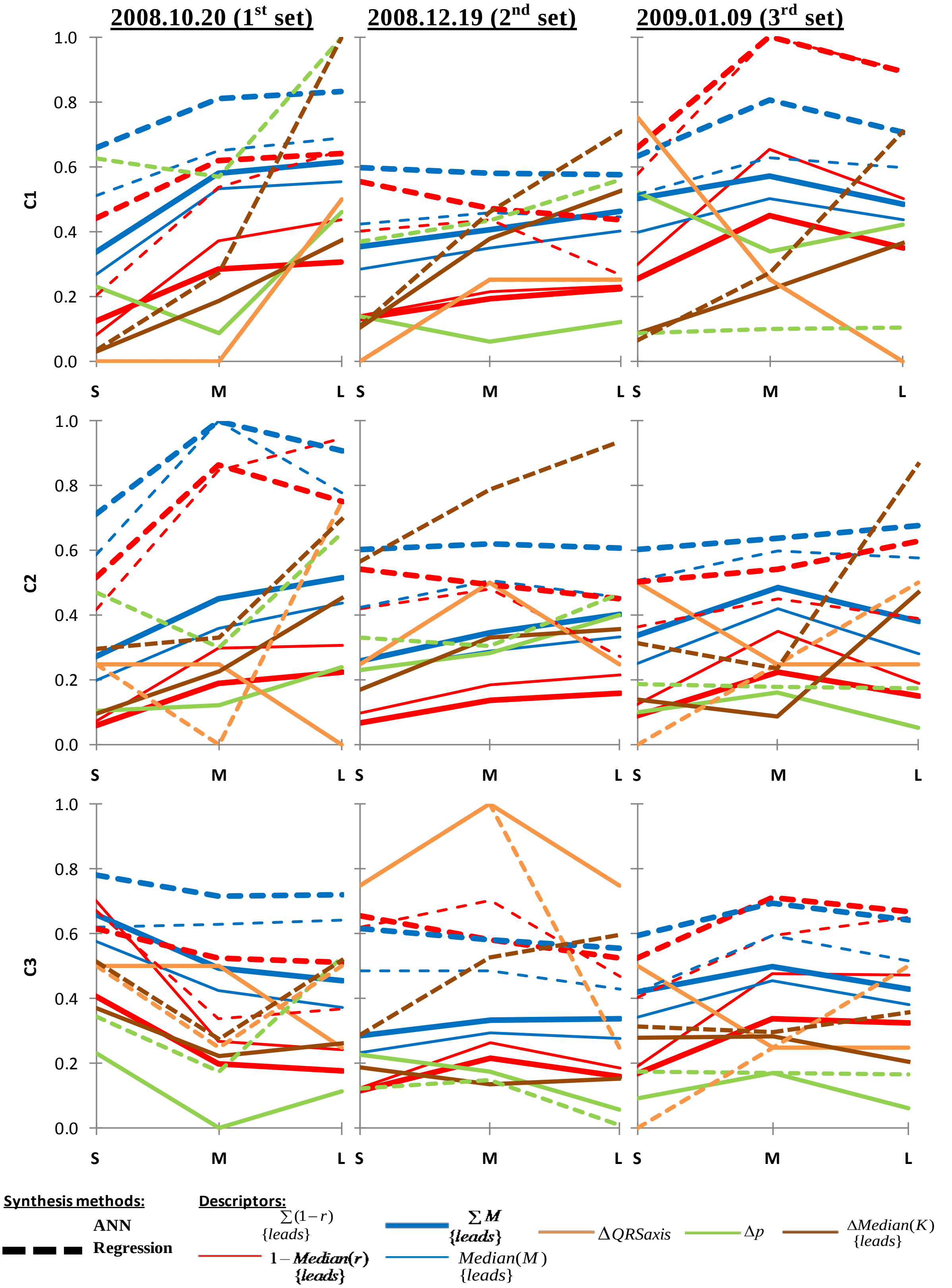


Fig.A.4.3_1 Normalised comparison results between a second standard 12-lead ECG (S2-ECG), not used for the synthesis models determination, and the 3 sets of 12-lead ECGs reconstructed from the 3 simplified ECGs recorded at 3 different dates, for the 1st subject

3rd subject

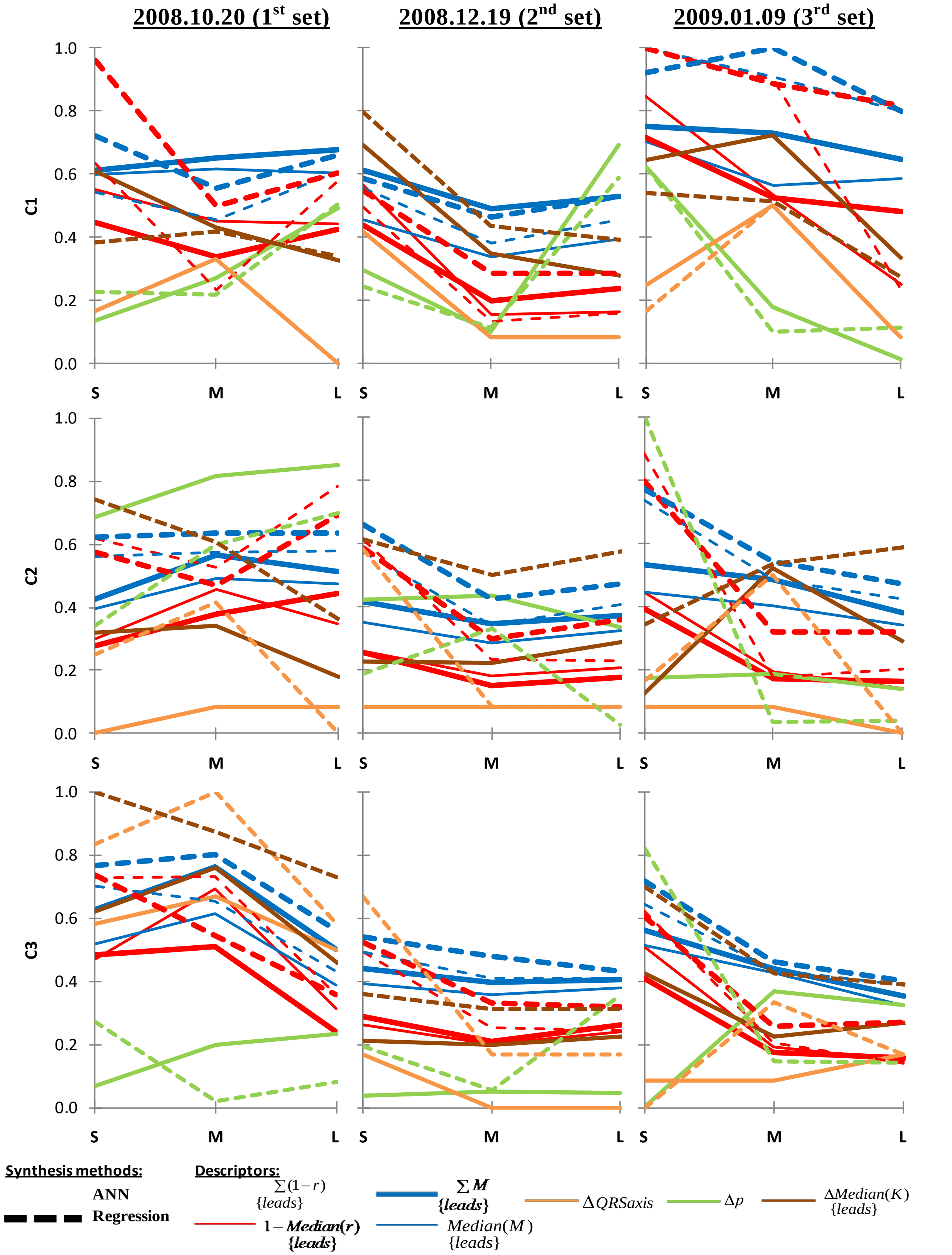


Fig.A.4.3_2 Normalised comparison results between a second standard 12-lead ECG (S2-ECG), not used for the synthesis models determination, and the 3 sets of 12-lead ECGs reconstructed from the 3 simplified ECGs recorded at 3 different dates, for the 3rd subject

4^{th} subject

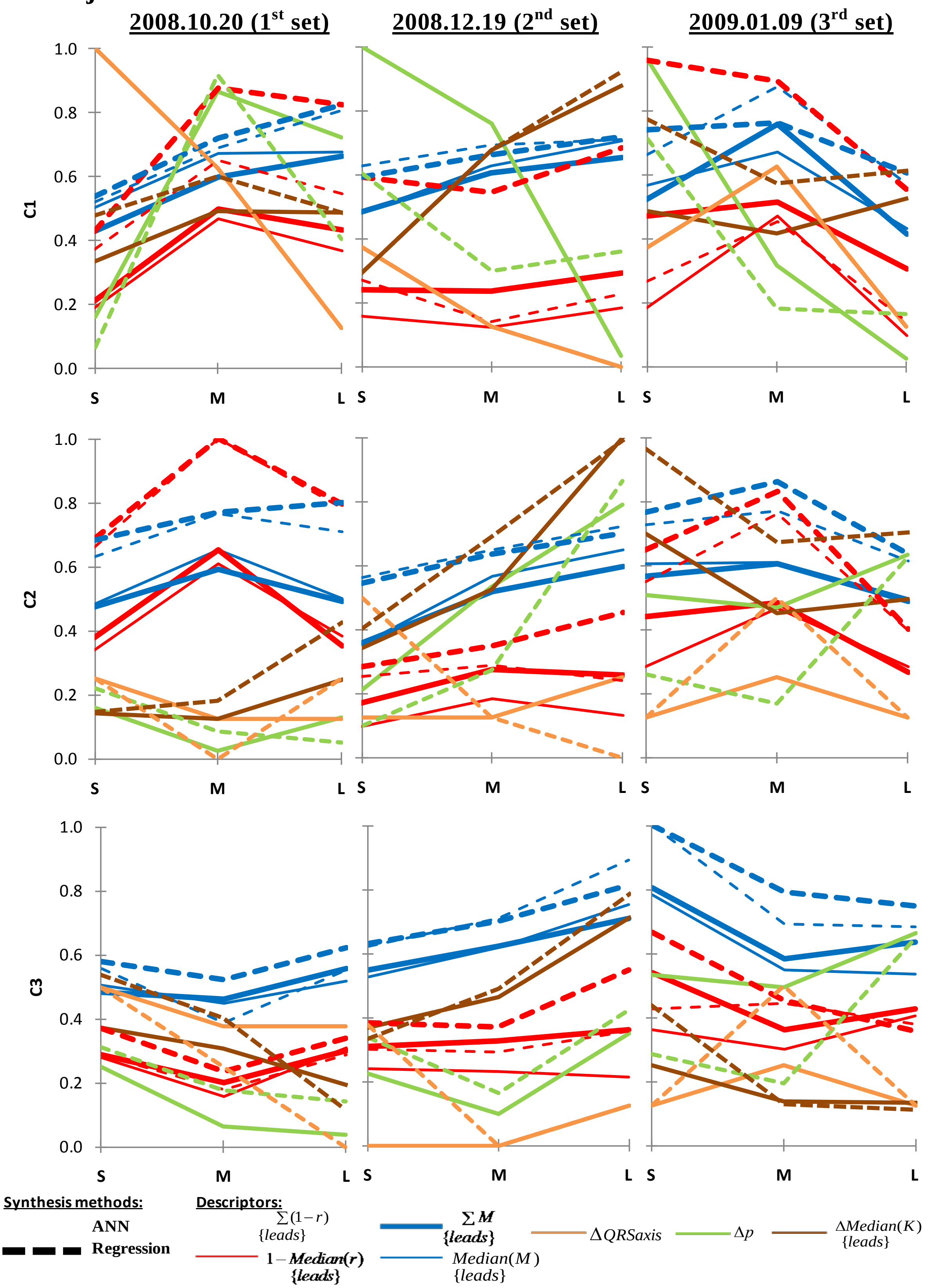


Fig.A.4.3_3 Normalised comparison results between a second standard 12-lead ECG (S2-ECG), not used for the synthesis models determination, and the 3 sets of 12-lead ECGs reconstructed from the 3 simplified ECGs recorded at 3 different dates, for the 4^{th} subject

FOLIO ADMINISTRATIF

THESE SOUTENUE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

NOM : KRUPAVICIUTE DATE de SOUTENANCE : le 20 décembre 2011

Prénoms : Asta

TITRE : Conception de services personnalisés pour la capture ubiquitaire de signes vitaux en pSanté

NATURE : Doctorat Numéro d'ordre : -ISAL-

Ecole doctorale : Informatique et Mathématiques (InfoMaths)

Spécialité : Informatique

RESUME :

L'objectif de cette thèse est de concevoir une architecture de services Web pour la dé-termination automatique d'un système de capteurs personnalisé, embarqué sous forme de vêtement intelligent dédié au self-care, afin de permettre à tout utilisateur profane d'enregistrer lui-même son propre électrocardiogramme (ECG), à tout moment et n'importe où. Le défi principal réside dans l'orchestration intelligente et dynamique de services métiers en fonction du contexte, pour qu'ils fournissent à l'utilisateur une solution personnalisée optimale tout en maîtrisant la complexité inhérente à la dépendance au contexte des interactions homme-machine, à l'extraction des connaissances des signes vitaux spécifiques à un sujet, et à l'automatisation de la reconfiguration des services. Une solution à ce défi est de créer une intelligence ambiante qui étend la notion d'informatique ubiquitaire et est capable d'offrir à l'instar d'un expert du domaine, une assistance intelligente personnalisée à tout citoyen.

Nous proposons une méthodologie de construction d'une architecture (DM-SOA) orientée-services, dirigée à la fois par les données et par des modèles, pour la production de services métiers intelligents et tenant compte du contexte. Cette architecture permet l'automatisation d'activités sophistiquées et personnalisées qu'un expert mettrait en œuvre pour le traitement d'un cas individuel, à partir de ses connaissances professionnelles. La solution proposée est basée sur une nouvelle approche de modélisation dynamique de processus métiers, et l'implémentation de services reconfigurables automatiquement. Elle consiste à mettre en œuvre un environnement intelligent fondé sur une ontologie de processus métiers des concepts du domaine et de son contexte, et sur une base de règles pour l'orchestration contextuelle des services.

Pour valider le bien-fondé de notre approche, nous avons proposé une ontologie pour l'automatisation de processus d'aide à la décision et nous l'avons évaluée dans le domaine de la cardiologie, en l'appliquant au problème de la sélection la plus adéquate possible d'un système de positionnement d'électrodes, spécifique à chaque individu, et capable de fournir un signal ECG de contenu diagnostique similaire à celui d'un ECG standard à 12 dérivations. Pour répondre aux besoins en situation de self-care tout en préservant la qualité diagnostique des signaux enregistrés à la demande, nous proposons la réalisation d'un nouveau système prototype de capture ECG-3D à trois modalités. Ce prototype a été testé sur huit jeunes adultes volontaires sains (4 hommes et 4 femmes) présentant diverses caractéristiques morphologiques.

Dans ce contexte, l'intelligence ambiante est fournie par un ensemble de services de qualité professionnelle, prêts à l'emploi par des utilisateurs profanes. Ces services peuvent être accessibles au travail, n'importe où, via des moyens classiquement utilisés chaque jour, et fournissent une aide appropriée aux utilisateurs non-compétents. Une telle approche d'in-telligence ambiante s'inscrit dans la vision d'une société de l'information ambitionnant de faciliter l'accès à la connaissance et correspond aux objectifs à long terme du programme de recherche Information Society Technologies de l'Union Européenne. Cette approche devrait contribuer à l'amélioration de la santé des individus.

MOTS-CLES : Intelligence ambiante – SOA – ontologie – modélisation dynamique de processus métiers – orchestration des services Web – Aide à la décision – pSanté – ECG – électrocardiologie quantitative – capteurs biomédicaux – informatique ubiquitaire

Laboratoire (s) de recherche : SFR Santé Lyon Est UMS3453 CNRS / US7 INSERM – plateforme eTechSanté

Directeur de thèse : Prof. Paul RUBEL Co-directeur : Jocelyne FAYN, HDR

Président de jury : Prof. Régis BEUSCART

Composition du jury : Prof. Régis BEAUSCART, Prof. Guy CARRAULT, Prof. Corine CAUVET, Dr. Jocelyne FAYN, Prof. Eric McADAMS, Prof. Paul RUBEL, Dr. Ana SIMONET, et Prof. Christine VERDIER