George A. Tsihrintzis
Maria Virvou
Lakhmi C. Jain
Editors



SMART INNOVATION, SYSTEMS AND TECHNOLOGIES 25

Multimedia Services in Intelligent Environments

Recommendation Services

Springer

Smart Innovation, Systems and Technologies

Volume 25

For further volumes:
http://www.springer.com/series/8767

George A. Tsihrintzis · Maria Virvou
Lakhmi C. Jain
Editors

Multimedia Services in Intelligent Environments

Recommendation Services

Springer

Editors
George A. Tsihrintzis
Maria Virvou
Department of Informatics
University of Piraeus
Piraeus
Greece

Lakhmi C. Jain
School of Electrical and Information Engineering
University of South Australia
Adelaide, SA
Australia

ISSN 2190-3018 ISSN 2190-3026 (electronic)
ISBN 978-3-319-00374-0 ISBN 978-3-319-00375-7 (eBook)
DOI 10.1007/978-3-319-00375-7
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013939051

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The explosive growth of data and information over the past decades, along with enabling technology to store and deliver data, has led to a descriptive term such as ‘Big Data’. ‘Big Data’ refers to datasets so extensive and complex that they are characterized by four features: volume, velocity, variety and veracity (Belanger 2012). Data *aggregate* from multiple sources such as Tweets and sensors, *change in real time* such as stock market trades, *take multiple forms* such as unstructured images and video, and *include noise and false signals* so that they may be untrustworthy (IBM 2013). IBM estimates that the world creates 2.5 quintillion bytes of data a day, with 90 % of the world’s data created in the last two years alone (IBM 2013). The problem of locating data of interest to a specific user in the face of such overwhelming obstacles provided the impetus for this book on *Multimedia Services in Intelligent Environments—Recommendation Services* as part of a series of edited volumes on multimedia services in intelligent environments.

Recommendation Systems, also referred to as recommender systems, attempt to help a user by identifying relevant and valuable data *in the eyes of a specific user* and presenting them in appropriate and useable forms. In general, these systems approach the task of recommendation through either collaborative filtering or content-based filtering (Rajaraman and Ullman 2011). Collaborative filtering develops a prediction of user preference based on the past behavior of the user and similarity with decisions of others. Content-based filtering bases the prediction on properties of the item recommended, while hybrid systems combine the two approaches. Intelligent systems that integrate artificial intelligence and advanced processing capabilities into the mix are essential to achieving the goals of more accurate and personalized user recommendations. The editors of this book have contributed to the literature by focusing attention on research that brings these areas together—advances in intelligent systems, enhanced recommendation methods, and the challenging context of multimedia services. Their efforts, and those of the contributors, will provide the reader with a view of the current state of the field and serve as a roadmap for future research.

Gloria Phillips-Wren
Loyola University
Baltimore, MD
USA

References

1. Belanger, D.: Senior Research Fellow, Stevens Institute of Technology. It's About the Data!! Lessons from a Decade+ Experiment in Big Data. Keynote Presentation to the BI Congress III, SIGDSS/TUN Pre-ICIS Workshop, Orlando (2012)
2. IBM: http://www-01.ibm.com/software/data/bigdata/. Accessed on 29 March 2013
3. Rajaraman, A., Ullman, J.: Mining of Massive Datasets. Chapter 9: Recommendation Systems. Cambridge University Press: New York, pp. 305–339 (2011)

Preface

Multimedia services are now commonly used in various activities in the daily lives of humans. Related application areas include services that allow access to large depositories of information, digital libraries, e-learning and e-education, e-government and e-governance, e-commerce and e-auctions, e-entertainment, e-health and e-medicine, and e-legal services, as well as their mobile counterparts (i.e., *m-services*). Despite the tremendous growth of multimedia services over the recent years, there is an increasing demand for their further development. This demand is driven by the ever-increasing desire of society for easy accessibility to information in friendly, personalized, and adaptive environments. With this view in mind, we have been editing a series of books on *Multimedia Services in Intelligent Environments* [1–10].

Specifically, this book is a continuation of our previous books [1–4]. In this book, we examine recent *Recommendation Services*. Recommendation services appear in the mobile environment, medicine/biology, tourism, education, etc. The book includes 10 chapters, which present various recently developed recommendation services. Each chapter in the book was reviewed by two independent reviewers for novelty and clarity of the research presented in it. The reviewers' comments were incorporated in the final version of the chapters.

This research book is directed to professors, researchers, application engineers, and students of all disciplines. We hope that they all will find it useful in their works and researches.

We are grateful to the authors and the reviewers for their excellent contributions and visionary ideas. We are also thankful to Springer for agreeing to publish this book. Last, but not least, we are grateful to the Springer staff for their excellent work.

Piraeus, Greece — George A. Tsihrintzis
Piraeus, Greece — Maria Virvou
Adelaide, Australia — Lakhmi C. Jain

References

1. Tsihrintzis, G.A., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Advanced Tools and Methodologies. Studies in Computational Intelligence Book Series, vol. 120. Springer, Berlin (2008)
2. Tsihrintzis, G.A., Virvou, M., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Software Development Challenges and Solutions, Smart Innovation, Systems, and Technologies Book Series, vol. 2, Springer, (2010)
3. Tsihrintzis, G.A., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Integrated Systems, Smart Innovation, Systems, and Technologies Book Series, vol. 3, Springer, (2010)
4. Tsihrintzis, G.A., Virvou, M., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Recommendation Services. Studies in Computational Intelligence Book Series, Springer, (2013)
5. Tsihrintzis, G.A., Virvou, M., Howlett, R.J., Jain, L.C. (eds.): New Directions in Intelligent Interactive Multimedia Systems and Services, Studies in Computational Intelligence Book Series, vol. 142, Springer, (2008)
6. Tsihrintzis, G.A., Damiani, E., Virvou, M., Howlett, R.J., Jain, L.C. (eds.): Intelligent Interactive Multimedia Systems and Services, Smart Innovation, Systems, and Technologies Book Series, vol. 6, Springer, (2010)
7. Tsihrintzis, G.A., Virvou, M., Jain, L.C., Howlett, R.J. (eds.): Intelligent Interactive Multimedia Systems and Services, Smart Innovation, Systems, and Technologies Book Series, vol. 11, Springer, (2011)
8. Virvou, M., Jain, L.C. (eds.): Intelligent Interactive Systems in Knowledge-Based Environments, Studies in Computational Intelligence Book Series, vol 104, Springer, (2008)
9. Obaidat, M., Tsihrintzis, G.A., Filipe,J., (eds.): e-Business and Telecommunications: 7th International Joint Conference, ICETE, Athens, Greece, July 26–28, 2010, Revised Selected Papers, Communications in Computer and Information Science Book Series, vol. 222, Springer, (2012)
10. Tsihrintzis, G.A., Pan, J.-S., Huang, H.C., Virvou, M., and Jain, L.C. (eds.): Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2012). In: Proceedings of the Eighth International Conference, IEEE Conference Publishing Services, (2012)

Contents

Multimedia Services in Intelligent Environments: Recommendation Services 1
George A. Tsihrintzis, Maria Virvou and Lakhmi C. Jain

User Modeling in Mobile Learning Environments for Learners with Special Needs 7
Maria Virvou and Efthimios Alepis

Intelligent Mobile Recommendations for Exhibitions Using *Indoor* Location Services 19
Mersini Paschou, Evangelos Sakkopoulos, Athanasios Tsakalidis, Giannis Tzimas and Emmanouil Viennas

Smart Recommendation Services in Support of Patient Empowerment and Personalized Medicine 39
Haridimos Kondylakis, Lefteris Koumakis, Manolis Tsiknakis, Kostas Marias, Eirini Genitsaridi, Gabriella Pravettoni, Alessandra Gorini and Ketti Mazzocco

Ontologies and Cooperation of Distributed Heterogeneous Information Systems for Tracking Multiple Chronic Diseases 63
Amir Hajjam

Interpreting the Omics 'era' Data 79
Georgios A. Pavlopoulos, Ernesto Iacucci, Ioannis Iliopoulos and Pantelis Bagos

Personalisation Systems for Cultural Tourism 101
Katerina Kabassi

Educational Recommender Systems: A Pedagogical-Focused Perspective 113
Salvador Garcia-Martinez and Abdelwahab Hamou-Lhadj

Melody-Based Approaches in Music Retrieval and Recommendation Systems. 125
Bartłomiej Stasiak and Mateusz Papiernik

Composition Support of Presentation Slides Based on Transformation of Semantic Relationships into Layout Structure 155
Toyohide Watanabe and Koichi Hanaue

Multimedia Services in Intelligent Environments: Recommendation Services

George A. Tsihrintzis, Maria Virvou and Lakhmi C. Jain

Abstract The term Multimedia Services has been coined to refer to services which make use of coordinated and secure storage, processing, transmission, and retrieval of information which exists in various forms. As such, the term refers to several levels of data processing and includes such diverse application areas as digital libraries, e-learning, e-government, e-commerce, e-entertainment, e-health, and e-legal services, as well as their mobile counterparts (i.e., *m-services*). As new multimedia services appear constantly, new challenges for advanced processing arise daily. Thus, we have been attempting to follow relevant advances in a series of edited books on multimedia services in intelligent environments. This is the fifth volume on the topic. In our earlier books [1–4], we covered various aspects of processing in multimedia services, including *advanced tools and methodologies, software development challenges and solutions, integrated systems, and advances in recommender systems*. The current volume is devoted to *recommendation services*.

1 Introduction

The term *Multimedia Services* has been coined to refer to services which make use of coordinated and secure storage, processing, transmission, and retrieval of information which exists in various forms. As such, the term refers to several levels of data processing and includes such diverse application areas as digital libraries, e-learning, e-government, e-commerce, e-entertainment, e-health, and e-legal services, as well as their mobile counterparts (i.e., *m-services).* As new

G. A. Tsihrintzis (✉) · M. Virvou
University of Piraeus, Piraeus, Greece
e-mail: geoatsi@unipi.gr

L. C. Jain
University of South Australia, Adelaide, South Australia, Australia

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*,
Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_1,

multimedia services appear constantly, new challenges for advanced processing arise daily. Thus, we have been attempting to follow relevant advances in a series of edited books on multimedia services in intelligent environments. This is the fifth volume on the topic. In our earlier books [1–4], we covered various aspects of processing in multimedia services.

More specifically, in [1] we were concerned mostly with low level data processing in multimedia services in intelligent environments, including storage (Chap. 2), recognition and classification (Chaps. 3 and 4), transmission (Chaps. 5 and 6), retrieval (Chaps. 7 and 8), and securing (Chaps. 9 and 10) of information. Four additional chapters in [1] presented intermediate level multimedia services in noise and hearing monitoring and measuring (Chap. 11), augmented reality (Chap. 12), automated lecture rooms (Chap. 13) and rights management and licensing (Chap. 14). Finally, Chap. 15 was devoted to a high-level intelligent recommender service in scientific digital libraries.

In [2], we were concerned with various software development challenges and related solutions that are faced when attempting to accommodate multimedia services in intelligent environments. Specifically, [2] included an editorial introduction and ten additional chapters, as follows: Chap. 2 by Savvopoulos and Virvou was on "Evaluating the generality of a life-cycle framework for incorporating clustering algorithms in adaptive systems." Chapter 3 by Chatterjee, Sadjadi and Shu-Ching Chen dealt with "A Distributed Multimedia Data Management over the Grid." Chap. 4 was authored by Pirzadeh and Hamou-Lhadj and covered "A View of Monitoring and Tracing Techniques and their Application to Service-based Environments." Chapter 5 by Bucci, Sandrucci, and Vicario was devoted to "An ontological SW architecture supporting the contribution and retrieval of Service and Process models." Chapter 6 by D'Ambrogio dealt with "Model-driven Quality Engineering of Service-based Systems." Chap. 7 by Katsiri, Serrano, and Serat dealt with "Application of Logic Models for Pervasive Computing Environments and Context-Aware Services Support," while Chap. 8 by Patsakis and Alexandris covered "Intelligent Host Discovery from Malicious Software."

In [2], we also included three chapters on new theoretical results, development methodologies and tools which hold promise to be useful in the development of future systems supporting multimedia services in intelligent environments. Specifically, Chap. 9 by Fountas dealt with "Swarm Intelligence: The Ant Paradigm," while Chap. 10 by Artikis dealt with "Formulating Discrete Geometric Random Sums for Facilitating Intelligent Behaviour of a Complex System under a Condition of Major Risk." Finally, Chap. 11 by Artikis dealt with "Incorporating a Discrete Renewal Random Sum in Computational Intelligence and Cindynics."

In [3], we presented various integrated systems that were developed to accommodate multimedia services in intelligent environments. Besides the editorial introduction, [3] included thirteen additional chapters, as follows: Chaps. 2 and 3 were devoted to multimedia geographical information systems. Specifically, Chap. 2 by Gemizi, Tsihrintzis, and Petalas was on "Use of GIS and Multi-Criteria Evaluation Techniques in Environmental Problems," while Chap. 3 by Charou, Kabassi, Martinis, and Stefouli was on "Integrating Multimedia GIS Technologies in a Recommendation System for Geotourism." Chapters 4 and 5 covered aspects of e-entertainment systems. Specifically, Chap. 4 by El-Nasr and Zupko was on "Lighting Design Tools for Interactive Entertainment," while Chap. 5 by Szczuko and Kostek was on "Utilization of Fuzzy Rules in Computer Character Animation". Chapters 6 and 7 covered aspects of education and e-learning systems. Specifically, Chap. 6 by Nakatani, Tsumaki, and Tamai was on "Instructional Design of a Requirements Engineering Education for Professional Engineers," while Chap. 7 by Burdescu and Mihăescu was on "Building Intelligent e-Learning Systems by Activity Monitoring and Analysis." Chapters 8 and 9 were devoted to medical diagnosis systems. Specifically, Chap. 8 by Schmidt and Vorobieva was on "Supporting the Search for Explanations of Medical Exceptions," while Chap. 9 by Aupet, Garcia, Guyennet, Lapayre, and Martins was on "Security in Medical Telediagnosis."

Chapters 10 and 11 were devoted to telemonitoring systems. Specifically, Chap. 10 by Żwan, Sobala, Szczuko, and Czyzewski was on "Audio Content Analysis in the Urban Area Telemonitoring System," while Chap. 11 by Dalka, Szwoch, Szczuko, and Czyzewski was on "Video Content Analysis in the Urban Area Telemonitoring System." Chapter 12 by Karapiperis, Stojanovic, Anicic, Apostolou and Despotis, was on "Enterprise Attention Management." An additional chapter, namely Chap. 13 by Raij and Lehto, was on "e-Welfare as a Client-driven Service Concept." Finally, Chap. 14 by Panoulas, Hadjileondiadis, and Panas was on "Brain-Computer Interface (BCI): Types, Processing Perspectives and Applications."

In [4], we were concerned with a special class of software systems, known as *Recommender Systems*. Besides the editorial chapter, [4] included an additional eight chapters organized into two parts. The first part of the book consisted of Chaps. 2, 3, 4 and 5. In this part, we addressed various aspects of recommender systems which form the core of recommendation services. Specifically, Chap. 2 by Lampropoulos and Tsihrintzis was on "A Survey of Approaches to Designing Recommender Systems." Chapter 3 by Nguyen and Santos was on "Hybrid User Model For Capturing a User's Information Seeking Intent." Chapter 4 by Lamb, Randles and Al-Jumeily was on "Recommender Systems: Network Approaches." Chapter 5 by Felfernig, Jeran, Ninaus, Reinfrank, and Reiterer made contributions "Towards the Next Generation of Recommender Systems: Applications and Research Challenges."

The second part of the book consisted of four chapters and presented various new theoretical results and tools that were expected to be incorporated in and improve recommender systems and recommendation services. Chapter 6 by Toledo, Sookhanaphibarn, Thanwonmas, and Rinaldo was on "Content-based Recommendation for Stacked-Graph Navigation," while Chap. 7 by Sakamoto and Kuboyama was on "Pattern Extraction from Graphs and Beyond." Chapter 8 by Kinoshita was on "Dominant Adaptive Hierarchical Process as Measuring Method of Service Values," while, finally, Chap. 9 by Artikis and Artikis was on "Applications of a Stochastic Model in Supporting Intelligent Multimedia Systems and Educational Processes."

The book at hand is a continuation of our coverage of multimedia services in intelligent environments. In particular, this book is devoted to a specific class of multimedia services, called *Recommendation Services*, which are analyzed in the following section.

2 Recommendation Services

Recent advances in electronic media and computer networks have allowed the creation of large and distributed repositories of information. However, the immediate availability of extensive resources for use by broad classes of computer users gives rise to new challenges in everyday life. These challenges arise from the fact that users cannot exploit available resources effectively when the amount of information requires prohibitively long user time spent on acquaintance with and comprehension of the information content. Thus, the risk of information overload of users imposes new requirements on the software systems that handle the information. Such systems are called *Recommender Systems* and attempt to provide information in a way that will be most appropriate and valuable to its users and prevent them from being overwhelmed by huge amounts of information that, in the absence of recommender systems, they should browse or examine [5].

Besides the current editorial chapter, the book includes an additional nine (9) chapters. Specifically, Chaps 2 and 3 are devoted to *mobile* recommendation services. Chapter 2 by Virvou and Alepis is on “User Modeling in Mobile Learning Environments for Learners with Special Needs,” while Chap. 3 by Paschou, Sakkopoulos, Tsakalidis, Tzimas, and Viennas is on “Intelligent Mobile Recommendations using *Indoor* Location Services.”

Chapters 4, 5 and 6 are devoted to medical/biological recommendation services. Specifically, Chap. 4 by Kondylakis, Koumakis, Marias, Genitsaridi, Pravettoni, Gorini, Mazzocco, and Tsiknakis is on “Smart Recommendation Services in Support of Patient Empowerment and Personalized Medicine.” Chapter 5 by Hajjam is on “Ontologies and Cooperation of Distributed Information Systems for Tracking Multiple Chronic Diseases.” Chapter 6 by Pavlopoulos, Iacucci and Bagos is on “Interpreting the ‘omics’ Era Data.”

Chapter 7 by Kabassi is on “Personalization Systems for Cultural Tourism,” while Chap. 8 by Garcia-Martinez and Hamou-Lhadj is on “Educational Recommender Systems: A Pedagogical-Focused Literature Review.” Chapter 9 by Stasiak and Papiernik is on “Melody-based Approaches in Music retrieval and Recommendation Systems,” while Chap. 10 by Hanaue and Watanabe is on “Composition Support of Presentation Slides Based on Transformation of Semantic Relationships into Layout Structure.”

3 Conclusions

As multimedia services in intelligent environments become increasingly more demanding, new challenges appear which require even more sophisticated tools, methodologies, and integrated systems to solve them. Coincidently, the application areas of multimedia services continue to expand at a very high rate. As a result, the entire field of multimedia services in intelligent environments cannot be effectively covered in the five volumes published so far. Thus, it may be expected that additional volumes on other aspects of multimedia services in intelligent environments will continue to appear in the future.

References and Further Readings

1. Tsihrintzis, G.A., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Advanced Tools and Methodologies, Studies in Computational Intelligence Book Series, vol. 120, Springer-Verlag, Berlin (2008)
2. Tsihrintzis, G.A., Virvou, M., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Software Development Challenges and Solutions, Smart Innovation, Systems, and Technologies Book Series, vol. 2, Springer-Verlag, Berlin (2010)
3. Tsihrintzis, G.A., Jain, L.C. (eds.) Multimedia Services in Intelligent Environments—Integrated Systems, Smart Innovation, Systems, and Technologies Book Series, vol. 3, Springer-Verlag, Berlin (2010)
4. Tsihrintzis, G.A., Virvou, M., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments—Advances in Recommender Systems, Studies in Computational Intelligence Book Series. Springer-Verlag, Berlin (2013)
5. Lampropoulos, A.S.: Machine learning-based recommendation methods for multimedia data. Ph.D. Thesis, University of Piraeus (2010)

User Modeling in Mobile Learning Environments for Learners with Special Needs

Maria Virvou and Efthimios Alepis

Abstract This chapter presents how mobile computing can support the education of children with special needs in mainstream schools. Children with special needs may have problems in the physical and/or mental participation in classes. Moreover, they often need a higher level of supervision and coordination by other people related to them: school-teachers, their parents, therapists and therapists. However, despite their problems, ideally, such children should be included in mainstream education and develop their personalities and knowledge in the same way as other children of their age. Mobile computing offers great opportunities for remote learning, communication, participation and coordination of supervision. This chapter presents a mobile educational application that keeps history models of students and records common problems, weaknesses and progress so that it may be used effectively by all parties involved in the education of children with special needs.

1 Introduction

Social communities around the world have acknowledged the urging need to give children with special needs equal opportunities to education. Over the last two decades, several countries have led in the effort to implement policies which foster integration of children with special needs [1]. The rather daunting challenge for educators and researchers in the field of learning disabilities is to examine effective

M. Virvou (✉) · E. Alepis
Department of Informatics, University of Piraeus,
80 Karaoli
and Dimitriou Street 18534 Piraeus, Greece
e-mail: mvirvou@unipi.gr

E. Alepis
e-mail: talepis@unipi.gr

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*, Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_2,

instructional approaches relative to specific learning disabilities while at the same time recognizing the powerful influence of the social context [2].

However, such challenges often make it difficult for teachers of mainstream schools to have children with disabilities in their classes. For example, in a recent study conducted among teachers of mainstream schools in Hong Kong [3], it was found that most of the teachers (70 %) were in favour of the realisation of equal opportunities for students with disabilities. However, at the same time, many of the teachers who participated in the study, tended to agree with the statements "integration was a burden to the schools and teachers" (over 60 %) and a "painful struggle for special students" (over 48 %). This is probably the reason why the problems remain unsolved to a large degree in most countries. For example, [4] report that millions of students across the United States cannot benefit fully from a traditional educational program because they have a disability that impairs their ability to participate in a typical classroom environment. They continue to make the point that computer-based technologies can play an especially important role for such students.

Certainly, computers can provide the medium for more personalised learning adapted to the special needs of students with disabilities. Yet the benefits of computer-based learning for students with special needs can be increased significantly through web and mobile technology. There are several reasons for that. First, mobile phones are more widespread than computers, thus they provide wider access to computer programs without the extra cost of buying a computer. This is a significant asset for the educational facilities of students with disabilities. Indeed, as [4] point out, the use of computer technology can help considerably students with disabilities to keep up with their non-disabled peers; however, the cost of the technology is a serious consideration for all schools. Second, mobile phones provide the ultimate mobility which is so important for students with disabilities and the people who support them (parents, teachers, therapists, co-students). Students with special needs often find it more difficult to be physically present in a class or computer lab. In this respect, the mobile phone can give them access to all computer facilities from their home without the extra cost of buying a computer. Similarly, the people who support students with special needs can have access to educationally assisting computer facilities through their mobile phones. Thus they also do not need to have bought special computer equipment or have to be at a specific place to meet physically each other about the progress of the students with special needs. This is an extremely important asset for the efficient education of students with special needs. Indeed, as [5] highlights, coordination with agencies, health care providers and families is essential for the education of children with special care needs. Mobile phones can facilitate and even encourage such coordination to a greater extent than other technology equipment.

In view of the above, in this chapter, we describe a mobile tutor that is meant to help the education of learners with special needs. Thus, it takes into account what these needs are so that it may incorporate as many features as possible that will help the educational process of students with disabilities. The educational process involves both the children themselves and the social and medical environment that supports them.

2 Overview of a Mobile Educational System for Students with Special Needs

The mobile tutor that has been developed aims at accommodating the needs of three categories of students with special needs: students with moving difficulties, students with sight problems and dyslexic students. To achieve this, the mobile tutor incorporates the following main features: It incorporates a multi-modal interface so that it may facilitate the interaction with users having disabilities. Users may use a microphone as an alternative way of interaction, rather than the traditional use of keyboard and mouse. The whole interaction with the educational application is additionally supported through a mobile interface. Figure 1 illustrates a mobile connection with the educational application where a student may use his/her mobile device in order to complete a multiple choice question of an examination in a specific lesson. Secondly, there is a student modelling mechanism that reasons about users' actions and keeps long term information about them so that the interaction is dynamically adapted to their needs. This information can also be used by the supporting people (parents, teachers, therapists) to find out what the progress of the students is. The student modelling component reasons both about the cognitive and the affective state of students as this is revealed by the students' actions during their interaction with the system. All the above use cases are modeled, using user modelling techniques in two stages. User modelling occurs firstly by considering the special need that the patient has and secondly by the personal data stored for each user. Finally, the system facilitates the communication of the involved supporting people through SMS services that may either be through the educational application or through the usual telecommunications channels.

The system can interact with children that have sight problems or kinetic difficulties on their upper parts/ends and are not able to use the keyboard or even the mouse. In that case, the interaction can be done visually and acoustically. The user can see the screen and also listen to what the system has to say through the use of speech engines that produce synthesized voices. Users are also able to give input to the system through the use of special software for process and voice recognition. The commands that are included in the system as well as the insertion of text can be done orally, through the use of a microphone. The system has the potential of being trained and improved according to the user profile, that includes characteristics of the users own voice.

Dyslexic students may also benefit from the mobile tutor. Dyslexia is a brain malfunction that affects the way that the individual learns, which appears as a

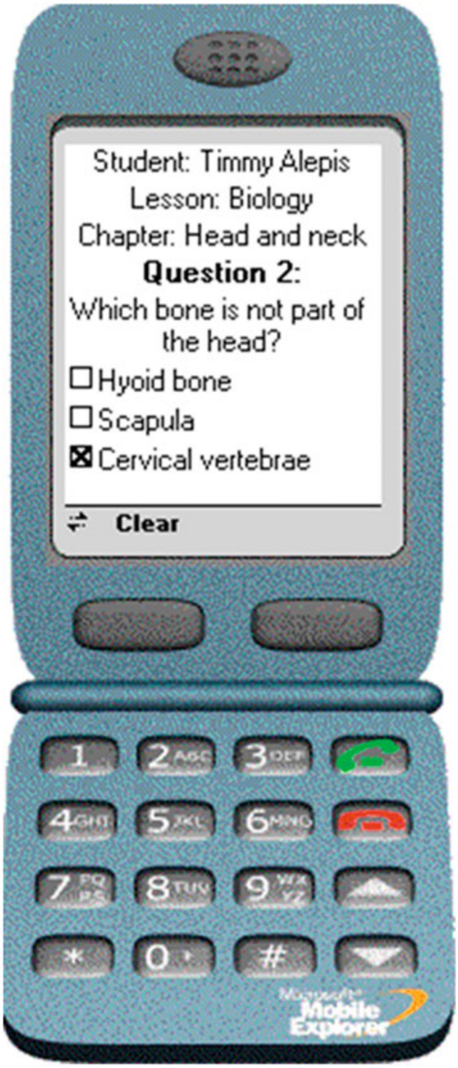


Fig. 1 Mobile interaction with the educational system

difficulty in reading and writing but also a difficulty in obtaining knowledge and skills in other areas of learning as well.

In the following subsections, the functionality of the system is described in terms of the targeted category of students with special needs.

2.1 Students with Moving Difficulties

The system can interact with people that have kinetic difficulties on their upper parts/ends and are not able to use the keyboard or/even the mouse. In that case, the interaction can be done visually and acoustically. The user can see through the use of the screen and also listen through the use of speakers. Correspondingly, the user is able to transfer data (users input) with help from the use of special software for process and voice recognition. The commands included in the system's educational process (such as the pressing of a specific button command), as well as the typing of plain text, can be achieved orally alternatively, through the use of a microphone. The system has the potential of being trained and improved according to the user profile, with the proper user model, that includes characteristics of the user's own voice. The percentage of success increases in accordance with the time the users interact with the system.

People with moving difficulties (ex. Lower ends) don't have great freedom of movement and transportation. Thus, we may suppose that the transportation of those individuals in special places equipped with personal computers equipped with the proper software would be difficult and in some cases even impossible. Under these circumstances, the use of a computer in a private place (such as the user's own house) is preferable. However, this presupposes that the person that will be supervising (the patient's doctor for example) the individual with special needs would be able to move in different places in short period of time and visit patients, so that he/she can check and evaluate their progress. Evidently, something like that (for example a doctor that would supervise more than 15 patients) would be particularly difficult, if we take into consideration the doctors' extremely heavy and stressful schedule. To accommodate this situation, the system alternatively supports access through mobile (cell) phones for user supervision purposes. With no obligation of buying extra equipment, doctors may use their mobile phone; connect cordlessly with the internet and gain access to the personal computers of their patients. They can accomplish that anytime, anywhere, thanks to the use of their mobile phone and through an existing web connection. Doctors can also check the progress of their patients, send and receive messages to or from them, and also exchange information and data with the medical application in reference of each user/patient. The main purpose of this educational application is to support the doctors and not to replace their personal contact with the patients.

2.2 Students with Sight Problems

In cases of children with sight difficulties, the interaction may be alternatively acoustical; meaning that the system may interact with the interacting user by using incorporated speech engines. Additionally, the user is being able to communicate orally with the system through the use of a microphone. Processing and recognition of oral speech take place as a next step. At the same time, the user can use a conventional keyboard or a keyboard specially adjusted for people with sight difficulties. Even the use of the mouse is partly evitable, since the movements of the mouse are accompanied with a corresponding vocal reply from the system. For example, the system can inform the user at any time for the mouses actual position, as well as which are the available user actions.

The communication with the system through the use of mobile devices could prove to be quite handy and useful for the doctors as well as for the persons with special needs. For example, even though the user interface of a personal computer is in many ways (audio-visual, speed) beneficial in comparison with the interface of a mobile device, it is rather expensive to buy (for someone who already hasn't it). On the other hand, a mobile phone is considerably cheaper and is something that almost every person, no matter his/her age, already possesses or will buy in the near future. Even more in countries were telecommunication is growing fast. By using the mobile facilities of the system, every user has the ability to get connected to the system's main server that has the medical software preinstalled and by using his/her mobile device is able interact with the medical application. The percentages of data that will be transferred from and to the mobile device and the main server, as well as the quality of the graphical interface are relative to the device the users own and/or the network they are connected to. The system incorporates a sophisticated software mechanism that adjusts the data to be transferred taking into consideration the limitations of each mobile device. No matter if the users have old or new models of mobile devices, at least the part of the submission of answers to test questions, either by entering simple text or by multiple choice, can be brought off. The only assumption to achieve that is an active network connection and a mobile device that supports network (wireless) connection.

2.3 Dyslexic Students

The term "dyslexia" means having difficulty with words and refers to the difficulty that a person has while reading, writing or spelling words that are either dictated to him or has to write on his own or copy them. Dyslexia is a brain malfunction that affects the way that the individual learns, which appears as a difficulty in reading and writing but also a difficulty in obtaining knowledge and skills in other areas of learning as well.

A more specified goal for the development of this software would be the support of reading and writing with the general use of the hearing sense:

1. Support during the reading: The dyslexic person has the facility of "hearing" his/her written material with the help of a synthetic voice through speech engines. In this way, word and meaning misinterpretations from the reading can actually be avoided. In addition, it is easier for the majority of dyslexic persons to hear something than to try reading it, since besides everything else dyslexic

people have been observed to have a weakness in the ability to focus while reading.

2. Support during the written expression and especially during tests: While using the computer's keyboard, dyslexic students are observed to muddle up or invert some letters and/or numbers. Help is provided in such cases in many levels.

In order to recognize a mistake, the system first compares the users' input with data from a dictionary of the language that the user uses. In this way, many orthographical mistakes may be located which are not necessarily due to dyslexia. The system can then suggest the replacement of the wrong word with word(s) that look alike and are orthographically correct.

Aiming mainly at dyslexic mistakes, the system incorporates a special facility of checking/comparing alphanumeric values, which is based on "sound comparison". Specifically, while the user submits string data on tests, there is a special comparison that takes place that compares the input and the correct answer using a sound comparing theory. Both the user's answer and the correct answer are transformed into phonemes using certain rules and then these phonemes are being compared (not the letters as usually). The idea that lies behind this specific "check" is that many times wrong words are written in a way that sounds the same as the correct ones. In addition, dyslexic students have been observed to have a tendency to exchange consonants. For the majority of such cases, the use of a wrong consonant does not change the sound and the vocalism of the word so much as it would if the wrong vowel was used.

In each case the system provides the quite simple yet important potential of reading through the application with the use of a speech engine for the text that the user types in real time. This means that the dyslexic user listens to what he/she types while typing through the keyboard on his/her personal computer. Thus, dyslexic students are helped not to be confused with visually similar words (and also semantically similar words) by hearing these words so that they may track their own mistakes more easily and more quickly. For example, they could write or read "Peter" instead of "Paul" as "similar" words, because firstly they are names, secondly because visually they have the same length and finally because both begin with the letter "P". However, if someone reads out loud these two words to a dyslexic, the difference becomes quite obvious to him/her.

Furthermore, the system offers the facility of the replacement of the keyboard as an input device by a microphone with the use of voice recognition. The user has the ability not only to input vocally but also to operate the program vocally. More specifically, the system teaches itself based on the user's voice and as a result the success ratio for correct input increases. Data input with the use of a user's own voice is something that can accommodate everyone, let alone dyslexics. All of the above facilities are based on incorporated theories and technologies in the program. The nature of dyslexia necessitates the creation of a knowledge database, which would draw from and incorporate data based on dyslexia. Moreover, it could incorporate the various mal-rules which we have assigned to the specific dyslexic and could also draw on the knowledge database on dyslexics. For example, although we cannot specify as a rule exploitable by the application that dyslexics always confuse or mess-up opposite concepts, we can however, "provide" the program with the information that a large percentage of dyslexics confuse "right" and "left" or "black" and "white".

Likewise, the system can create "dyslexic user models" so that better results may be achieved for each individual user. The interaction with the system is more individualized and adaptive to each dyslexic case. With the creation of dyslexic user models, the system will have the ability to track the particular user's weaknesses and offer the specific help required for this user's needs.

The supervision of dyslexics during their interaction with the system can also be achieved by a doctor with the use of wireless communication via a mobile phone. It is necessary for a professional to follow the progress of the dyslexics while they are interacting with the system in order to evaluate the results of the interaction (benefits, problems, etc.) and also to be in a position to adapt the system's parameters based on the users needs. Both instructors and doctors do not need to purchase a computer, desktop of laptop, but rather use their mobile device which they may already have. Additionally, the instructor is not required to have computer knowledge since he/she can the educational application consisting of mobile pages and controls, is quite easy and user-friendly. Access and user identification can be achieved at any moment and place and as many times required as one of the main advantages of mobile technology. Finally, with the use of this technology, the dyslexics are introduced to communication among themselves and the instructor(s) who can oversee them at any given moment from any given place.

3 Mobile Coordination of People Who Support Children with Special Needs

Children with special needs need help that should be provided not only from their family, but also from their school, specific organizations and foundations and the community indiscriminately. Comprehension and support for these children improve the more knowledge, we have in each particular situation and are critically associated with the quality of services we can afford to give. These services are presented, analyzed and reviewed in [6], where we can clearly see the fact that deficient support of people with special needs has to do in great percentage with the missing, quantitative and quantitative, of the corresponding mediums and services from the community. Moreover, the "provision" of help from a professional is particularly expensive and in most cases, either the government can't provide it, or the family of the patient can't afford it.

Since provision of help from a professional or doctor is expensive and taking into account the fact that people with special needs need continuous observation, the software we are proposing is developed mainly to complement the work of a doctor or a consultant and also to provide important and useful information to the parents.

All users, students with special needs, therapists, teachers, parents and co-students can use the application to cooperate in the educational process. All of them are not only able to have easy access to the data-bases of the application but they can also communicate with each other. Some usual cases of that kind of communication are the following:

1. A therapist sends remarks to a teacher about a student who possibly suffers from dyslexia and needs special care, tests and theory adapted to his/her problem.
2. A teacher informs a therapist that a specific student seems to have "unusual" problems with mathematics.
3. A teacher sends a message to a student's parent to inform him/her that their child's performance in a course is improving.
4. A teacher sends a message directly to a student to tell him to try to be more concentrated while spending time reading parts of the theory.
5. A student sends messages to a teacher about a problem s/he has while interacting with the mobile educational application.
6. Students send messages to each other in terms of collaboration.

The communication between them can be realized in many ways (Fig. 2). By using their mobile phone (and of course connecting to the application's mobile pages) therapists can send short messages via the short message service (SMS), either to the students or to their teachers and parents (If they also have mobile phones). The email service can be used as well in situations where something more than simple "short messages" is needed. Alternatively therapists can write the messages to the application's data-base. In this case, therapists have to declare the name of the receiver (whether s/he is a teacher, a parent or a student) and the application will use its audio-visual interface to inform him/her as soon as s/he opens the application.

In the first case (e-mail or SMS message) the message is written to the system's database and then is sent to an internet service that provides SMS and e-mail sending. Therapists are also able to send an e-mail directly through their mobile phones but it may be preferable for them to use the application to do that. The main reason for this is the fact that mobile networks are considerably slow and cost much and thus they may not be very convenient for the application's users. Thus, the application is expected to send an e-mail to the internet service. The body of this e-mail will be the body of the short message and the "subject" of the e-mail will specify the receiver by his/her mobile telephone number. Messages (if sent from the teacher to the student for example) can contain information about which test the student should visit next or about anything that the teacher thinks that the student should pay attention at. Messages can also be sent to parents (from therapists or teachers) who may find an easy way to stay informed about their child's progress.

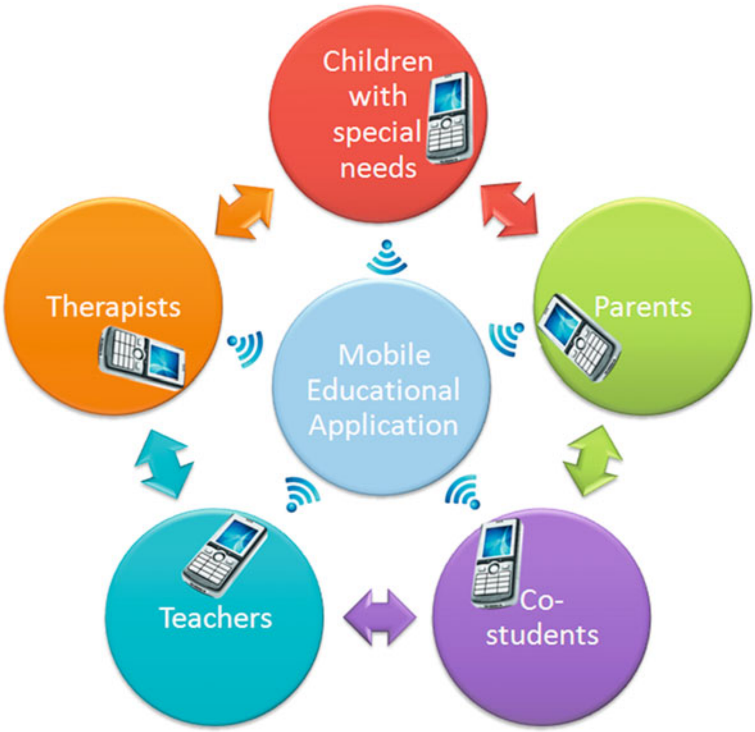


Fig. 2 Mobile communication through the educational application

Additionally, even more potential is given to the owners of the technologically latest mobile phones (Smartphones). They may use the quite recent but rapidly spreading technology of Bluetooth or infrared wireless communication. This means that they may alternatively use the above technologies in order to send and/or receive data to and from the main server of the educational application through their mobile device. Of course, there is a limitation in the distance between their device and the server (with Bluetooth approximately 10–20 m, while with infrared rays only a few centimeters), but considering the actually inexpensive way of communication, many users may find it quite "useful".

The discrimination between the users is done by the application (installed in the server) and for each different user a personal profile is created and stored in the data base. In the data base, other personal information such as e-mail accounts and mobile phone numbers are stored so that users won't need to recall them while trying to communicate with each other. Only the users name is actually needed. In order to accomplish that, user name and password is always required to gain access to the mobile web pages. On the first level of authentication, the application determines whether there is a valid username-password or not. On the second level of authentication, users are distinguished in simple users (students), teachers, parents and therapists.

Mentioning the security and privacy aspects of the mobile application it is also able to use the Secure Sockets Layer (SSL) in order to encrypt data during important data transactions. Both the IIS server and the NET Framework for mobile devices, the technologies that we have used, support this operation simply by inserting a valid certificate signed by a certificate verifying authority.

4 Conclusions

This chapter has presented a mobile educational application that supports the education of students with special needs. The system involves facilities that allow its use by the children with disabilities and also records their progress and common weaknesses so that it may become more personalized. Moreover, one important feature of the system is that it supports the communication and coordination of the people involved through the application, to reinforce efficient supervision and inclusion of the students with special needs in the social context of a classroom.

It is in our future plans to evaluate the resulting system in order to examine the degree of its usefulness as an educational tool for teachers, as well as the degree of usefulness and user-friendliness for the children with special needs as end users who are going to use the educational system.

References

1. Avramidis, E., Bayliss, P., Burden, R.: Student teacher’s attitudes towards the inclusion of children with special educational needs in the ordinary school. Teach. Teacher Educ. **16**, 277–293 (2000)
2. Kracker, M.: Classroom discourse: Teaching, learning, and learning disabilities. Teach. Teacher Educ. **16**, 295–313 (2000)
3. Pearson, V., Lo, E., Chui, E., Wong, D.: A heart to learn and care? Teachers’ responses toward special needs children in mainstream schools in Hong Kong. Disabil. Soc. **18**(4), 489–508 (2003)
4. Hasselbring, T.S., Glaser, C.H.W.: Use of computer technology to help students with special needs. Future Child. **10**(2), 102–122 (2000)
5. Barrett, J.C.: A school-based care management service for children with special needs. Fam. Community Health **23**(2), 36–42 (2000)
6. Chung, M.C., Vostanis, P., Cumella, S., Doran, J., Winchester, C., Wun, W.L.: Children with special needs: Use of health services, behaviour and ethnicity. Child. Youth Serv. Rev. **21**(5), 413–426 (1999)

Intelligent Mobile Recommendations for Exhibitions Using *Indoor* Location Services

Mersini Paschou, Evangelos Sakkopoulos, Athanasios Tsakalidis, Giannis Tzimas and Emmanouil Viennas

Abstract Information and Communication Technologies (ICTs) are utilized in an increasing number of museums and collection exhibitions world-wide. In this chapter, we present novel fully-automatic mobile assistant with indoor recommendation services. We will discuss novel efficient techniques in order to provide within a single software solution (a) typical, (b) semi automatic and (c) seamless-no human interaction tour guidance and recommendations during an exhibition tour. We have designed, developed, deployed and evaluated the solution at a real case exhibition to provide at a users' level an intelligent personalized virtual guide. The provided solution aims to assist visitors and to provide full automatic multimedia or audio guidance during exhibition visits using Wi-Fi based indoor and outdoor positioning, mobile messaging and wireless data provisioning. In the proposed chapter we will discuss the design, implementation, deployment and outcomes of the research and development endeavor for the integration of mobile multimedia, positioning and messaging services into a novel automatic personalized exhibition recommender-assistant. Deployment and evaluation issues of the proposed solution will be discussed for the case of the Museum "Digital

M. Paschou · E. Sakkopoulos (✉) · A. Tsakalidis · E. Viennas
Department of Computer Engineering and Informatics School of Engineering,
University of Patras, Rio Campus 26500 Patras, Greece
e-mail: sakkopul@ceid.upatras.gr

M. Paschou
e-mail: paschou@ceid.upatras.gr

A. Tsakalidis
e-mail: tsak@ceid.upatras.gr

E. Viennas
e-mail: biennas@ceid.upatras.gr

G. Tzimas
Department of Applied Informatics in Management and Finance, Technological Institution
of Messolonghi, Nea Ktiria 30200 Messolonghi, Greece
e-mail: tzimas@ceid.upatras.gr

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*,
Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_3,

Exhibition of History of Olympic Games in Antiquity" in Ancient Olympia (Greece, EU).

1 Introduction and Motivation

Multimedia application use electronic information that may integrate different types of objects including text, images, video and audio. Information and Communication Technologies (ICTs) used in appropriate ways in a museum or an exhibition can result in a functional upgrade of the visitor's experience [1]. The added value includes promotion and enhancement of educational, research or entertainment purposes. In the past few years personalized services [2] of Information technology have begun to be adopted in a rapidly increasing number of cultural and touristic exhibitions around the world [3–6]. The aim of these services is to provide an additional and perhaps more interesting dimension to the presentation of the exhibits and ultimately attract more visitors [7, 8].

In fact Smartphone application marketplaces and stores include a number of apps that serve as a substitute of a typical travel guide. Unfortunately, such applications are not possible to support automatic recommendation and guidance whilst indoors due to lack of GPS signal. There are semi-automatic attempts to provide guidance using external aids such as numbered signs, QR codes or RFID signs. However, all the former approaches involve either architectural interventions at the exhibition halls (e.g. adding labels and posts which may spoil exhibits presentation etc.) or human interaction and explicit user choices in order to provide recommendations and further assistance. In our approach we propose a novel smartphone app solution to serve as a recommender personalized museum guide. To deliver the recommendation a novel indoor position solution based on Wi-Fi technology has been employed. The automatic recommendation system delivers accurately exhibits presentations and multimedia details detecting the position of the visitor in the museum with less than 1 m accuracy. The proposed solution has been deployed and working in the Museum "Digital Exhibition of History of Olympic Games in Antiquity" in Ancient Olympia (Greece, EU).

In our case study and implementation environment of the proposed mobile recommendation museum assistant, the guests of the Digital Exhibition of History of Olympic Games are invited to learn about the Games and the Greek culture through interactive media. These interactive modules include interactive discussion tables and applications that visitors are able to interact with, by touching the monitor. Additionally, a timeline system is available, which consists of a vertically mounted display, and can be moved along the horizontal axis by the visitor. Various dates related to the celebration of the ancient Olympic Games are carved on the wall behind the screen. Dynamic multimedia information is presented, depending on theme chosen and the date corresponding to a specific screen position. A researcher information station is provided, whose purpose is to enable specialized visitors (i.e. archaeologists, researchers and students) to explore in depth the history of the area and the exhibits and allow them to print anything they consider interesting.

The main objective is to render a visit to the Museum an experience that visitors will remember with pleasure, by providing high quality services. Therefore, a set of special services and facilities are provided that are divided chronologically into three categories: "Before the visit", "During the visit" and "After the visit". As a result visitor management is facilitated and the overall experience of the visitor is improved. The visitor has a number of digital media available, which is the key to a smooth transition from the Archaeological Exhibition itself to the Digital Gallery and aims at attracting visitors to the main Digital Gallery. The proposed mobile museum guide assists the visitor with automatic guidance or typical step by step semi-automatic assistance and recommendations throughout his/her visit.

Details on the mobile assistant are discussed as follows: Sect. 2 provides an overview of related work. Section 3 lays down the outline of the storyboard designed to depict the application logic. In Sect. 4 the architecture of the system is discussed. Section 5 presents indoor positioning and technologies. Finally, the case study functionality and evaluation of results are discussed in Sect. 6. Section 7 concludes the chapter and provides future ideas for further investigation.

2 Related Experience

Multimedia tools used in museums facilitate communication of large amounts of information in a user-friendly and interesting manner. In parallel, multimedia tools allow visitors to access the information they require at their own pace [3]. Recently, many researches and approaches to analyze personal behavioural data and applications which use these logs are reported [9, 10]. Thanks to downsized and high-performance device such as smart phone, recording people's life becomes easier. Besides, by mining these data, we can receive useful feedbacks. Wi-Fi based indoor location systems have been shown to be both cost-effective and accurate, since they can attain meter-level positioning accuracy by using existing Wi-Fi infrastructure. However, those systems still introduce technical challenges that have to be faced and it turns out that creating a useful indoor navigation app requires more than navigation [5].

A sensor-assisted adaptation method is proposed in [4], which employs RFID sensors and environment sensors to adapt the location systems automatically to the changing environmental dynamics. The proposed adaptation method performs online calibration to build multiple context aware radio maps under various environmental conditions. An image-based indoor positioning system for digital museum application is presented in [11], using images to which location information is embedded. By using robust image matching, fast nearest neighbour search algorithm based Locality Sensitive Hashing (LSH), and the confidence measure which takes into consideration of the location of the objects in the query image, the presented system aims at estimating users' locations with high accuracy and in a short time, which is supported by conducted experiments.

3 Design and Storyboard

The key aim of technology in the premises of a museum is multiple, offering new dimensions in the experience of the visitors. In this chapter research and development results are presented as the outcome of the endeavor for the delivery of the Museum "Digital Exhibition of History of Olympic Games in Antiquity" in Ancient Olympia (Greece, EU). The novel mobile solutions implemented in this case have a triple target:

- Visitor participation in collective activities, through the implementation of digital assistants that encourage collective actions (e.g. mobile messaging, family member/friend position detection).
- Election of the "games" as a social commodity and a basic ingredient of the ancient Greek life-style, the perpetual contrast of the opposites, the personal freedom and the competitive spirit.
- Unobtrusive integration of all digital means with the available ancient exhibits within the museum.

Among the numerous different digital solutions delivered in the museum, the most important role plays the Automatic Personalized Tour with mobile devices (PDA, smartphones), which every visitor receives upon his/her entrance in the museum. It is based upon a modular architecture that ensures the separation of concerns maximizes re-usability of sub-systems and delivers fine-tuned positioning, messaging and personalized multimedia presentation services.

In this chapter the architecture of the Automatic Personalized Tour with mobile devices solution is discussed and its high level functionality in the area of the museum. A discussion of the necessary subsystems, the storyboard, technological hints and the final outcome will be outlined to show the feasibility and effectiveness of the proposed approach. The overview of the functional specifications is outlined below.

(a) A visitor receives essentially a portable device (smartphones) properly customized. The handling should be simple and done by touch screen. He is given the possibility for presentation of multimedia data in one of at least four languages available. It is possible to choose among different tours such as a short content tour or a fully detailed tour.

(b) The automatic personalized assistant has a novel system for identifying and presenting the position/location of each visitor in the wide area of the museum. The positioning is based on two interconnected wireless Wi-Fi 802.11 g network infrastructures; (1) one within the 14 exhibition halls and (2) a second outside and around the museum. Contrary to previous approaches, the solution is designed and implemented from scratch in order to achieve minimum error location detections, minimum network bandwidth consumption and number of handshakes as well as minimizing PDA performance cost. This way, the visitor can wander on the premises of the museum freely wearing the mobile device and the headphones and depending on the visitor's current position, the equivalents audiovisual multimedia data is presented to him/her. Moreover, the visitor has information on the position of his family and/or his friends in the museum as the tour assistant has the possibility of establishing a group of tour members move independently if they wish.

(c) Alternatively, the assistant smartphone tour app can operate non-automatically allowing the user to choose directly the multimedia presentation that wants to watch and listen to independently of his/her position in the museum.

(d) In addition, the tour assistant supports mobiles instant messages between visitors in the same group of users. Also allow each user to receive information messages from the museum information desk through a specially designed Monitoring and Management Visitors Administration Console.

Another key advantage of the proposed approach is that each subsystem is supported by an XML based layer that allows (a) clear separation of concerns and (b) maximizes the chances of re-usability of the sub-modules (e.g. mobile positioning, multimedia data provisioning, messaging, user profile and logging). The fundamental non functional specification is the XML based data interchange at all layers of the solution, based on the idea of multi-layered profiles of [6]. The content and customization of the mobile application and services are all XML based in order to achieve maximum scalability and interoperability. Personalization is also achieved through XML based user profiles and visitor's choices XML storage. XML support is also the basis of the messaging approach in order to enable extensions and allow interconnection with messaging servers outside of the museum if needed.

The application is designed to help navigate each visitor within the museum. The museum is physically divided in 14 exhibition halls and thematically divided into 12 thematic sections. Each section has from 1 up to 4 thematic keys. The solution is possible to handle multimedia data for any number of halls and/or thematic rooms and keys of the data presented. The term "keys" includes extra information (it can be regarded as a sub-unit) that are available to the visitor after watching the content on a specific section chosen. So if there are 2, he/she keys will be able to select and view two further sub-units related to this topic.

There are two main different types of operation for the application:

- Automatic operation (automatic positioning of the user current position inside the museum using the Wi-Fi network).
- Manual operation (the visitor choose manual the presentation of a thematic section which would like to attend).

There are three different ways of tour:

- Audio tour (contains only an audio presentation on the thematic sections).
- Detailed audio tour (contains only an audio presentation on the thematic sections and the keys).
- Audiovisual tour (contains an audiovisual presentation on the thematic sections and the keys).

Another observation is that our application is multi-language and more specifically supports the following languages: Greek (Gr), English (En), German (De), French (Fr). The XML underlying customization layer allows the integration of any number of additional languages. For the latest Olympic Games a prototype for the Chinese language has also been developed. A number of sub-cases are also included in the scenario to deal with exceptions and handle possible errors such as failure of Wi-Fi due to electrical shortcuts, malfunction of the PDA device that leads to automatic locking of the PDA and more.

4 Architectural Issues

Next the overall architecture of the portable devices system is presented and the building modules are analyzed below. The personalized assistant consists of and cooperates with the following components:

- Mobile Devices Subsystem at a Visitor level (MDSC): it constitutes the smartphone application that can be given installed at the mobile devices offered by the Museum or maybe downloaded through the local area wireless network.
- Mobile Devices Management and Disposal Subsystem (MDMDS): it includes management software for managing visitors, mobile devices docking station-carousel, preparation of MDSC during the charging and the updating of data in MDSC, as these are taken from the DMS subsystem and the operation/subsystem issuing data for portable devices.
- Networking Services Subsystem (NSS): it includes services and management functions of the MDSC subsystem such as location positioning, monitoring the location of visitors inside the museum, where WNIS and Communication Server is available.
- Data Management System (DMS): it includes the special operation of publication of multimedia data that is presented in MDSC subsystem.
- Wireless Network Infrastructure Subsystem (WNIS): it includes the necessary infrastructure for the availability of wireless interconnection of MDSC.

A first-level representation of the overall architecture of the subsystem together with the associated systems and subsystems with which it interacts shown in the figure below (Fig. 1), and then presented in more detail each module.

4.1 Mobile Devices Subsystem at Visitor Level

This subsystem is the basic application that will present the solution that will be installed on mobile devices. It consists of a Cell Application which includes the following three applications:

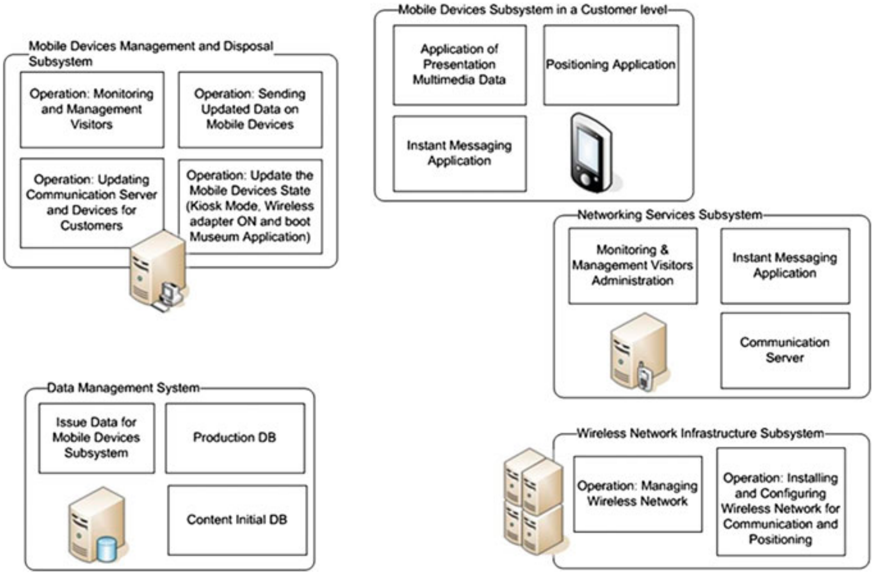


Fig. 1 Overall architecture of the system together with the associated systems and subsystems

- Application of Presentation Multimedia Data
- Instant Messaging Application
- Positioning Application.

The following are the general operating requirements of the subsystem, the way interfaces with other subsystems and analyzed in greater depth the applications that compose the particular subsystem.

4.1.1 Operational Specifications of Subsystem of Mobile Devices at Visitor Level

This subsystem (MDSC) is the backbone of Mobile Device Subsystem. The basic functional specifications are as follows:

- Presentation of multimedia data in four languages for content applications.
- Identifying and presenting the position of the mobile devices at the site of the museum where there is a wireless network infrastructure on the basis of information from the subsystem NSS. This feature has three objectives.
  - Present the location of the device/visitor within the smartphone application museum map.
  - Updating the position of the device in NSS subsystem.

  – Show the position of the device/visitor within the smartphone application museum map of other devices/visitors usually of the same tourist group with the visitor (e.g. fellow student, family member etc.).

• Network management and instant messaging.

4.1.2 Underlying Infrastructure for the Smartphone App: Cell Application

Overview/Operations: The Cell Application implements communication among the Application of Presentation Multimedia Data, Instant Messaging Application and Positioning Application, as well as communicating with the NSS. It is somehow a cell, which consolidates these three applications. For example, when the Positioning Application detect the position of the visitor in the area of the museum notify the cell application and it undertakes to inform the NSS subsystem and the Application of Presentation Multimedia Data.

Briefly, the operations provided by the Cell application layer are as follows:

• It constitutes the administrator of communication between the remainder three applications.
• It allows the communication with the other subsystems.
• It manages the situations of errors.
• It manages the messages that are presented in the user.
• It manages the situations of network loss.

4.1.3 Presentation Multimedia Data

Overview/Operations: The main role of this application is the presentation of multimedia data in multiple languages. More specifically, the application has the following features:

• Provides a multimedia presentation of data in different languages and for the case study present there are four (4) languages available. It is possible to extend the localization and provide any number of different languages using the XML based on the dynamic configuration designed for the proposed personalized museum assistant.
• It provides two types of operation: Auto (changing the content presentation automatically depending on the position of the user in the museum) and Manual.
• It provides three types of presentations: Passive Acoustics (the visitor may listen to provided guided tours using the automatic recommendations based on his/her location in the museum—as a consequence the visitor just hangs the museum guide from his shoulder or neck and walks around the museum freely), Interactive Acoustics (the visitor may choose among the presentations available for

the exhibits) and Audio-Visual (the visitor may choose among the multimedia presentations available for the exhibits at his/her own pace).

- Interacts with the Core Application in order:
  - to receive data concerning the location/position of visitors in the museum (indoors and outdoors) and behaves accordingly
  - to receive the name of the visitor and display it in the welcome message
  - to deliver messages which should be displayed at the user
  - to manage the exception cases in lack of network or positioning into an exhibition hall
  - to set the language of the application upon registration.
- It manages with appropriate updates to the user situations where there is no network and cannot locate the exhibition hall in which the visitor is the given moment.
- It provides the possibility of depiction of place of user in the museum map. This map functions also as a means of navigation in the application.
- It dynamically presents application data using external image files (.jpg format), video (.swf Flash format) and audio (.mp3 format). Thus it is very easy to change the content of the application. The same is true with all the text that exists in implementation. The place of files is given dynamically from Core Application.
- It provides the possibility of on-line help in the user.

4.1.4 Instant Messaging Application

Overview/Operations: The Instant Messaging application gives the possibility in each visitor of directly communicating with text messages with the users that belong in the same team of users thus as they were declared at the process in the MDMDS. Also allow each user to receive alert messages from the system administrator through the Monitoring and Management Visitors Administration Console. Analytically, the operations that provide the particular application are following:

- Sending and receiving messages via direct keyboard users who have been declared in the same group of users.
- Display the state of Presence (Online/Offline) for specific users.
- Get information messages from the administrator of system.
- Display of position (at the hall) of users reported in the same group of users.
- Display of language of users that has been declared in the same team of users.

4.1.5 Positioning Application

Overview/Operations: Using the available Wi-Fi network, the topology of exhibition halls and the wireless network card that has a PDA, the Positioning Application undertaken to determine the position of each visitor to the site and then notify the Cell Application to manage this information.

4.2 Mobile Devices Management and Disposal Subsystem

This subsystem includes the management of the carousel, the management of incoming and outgoing visitors, the preparation of MDSC during the charging and the updating of data in MDSC, as these are taken from the DMS subsystem and the operation/subsystem issuing data for portable devices.

4.2.1 Operational Specifications of Mobile Devices Management and Disposal Subsystem

The basic functional specifications of the subsystem in relation to other subsystems operating PDA applications are as follows:

- Management and monitoring of visitors at dispense and return of equipment process and setting groups of visitors.
- Update subsystem MDSC and NSS (user and group information at the Communication server) for instant messaging (groups).
- Update of multimedia data with new versions in MDSC as taken by the NSS subsystem and other necessary files for the functioning of the MDSC per PDA (executables, file XML, etc.) are available in pre-spatial point. This procedure is performed when the PDAs is in the carousel.
- Turns in running the MDSC application (kiosk mode), shortly before the PDA charged to a user, with available network connection (it checks the availability from the MDSC), with turned on and regulated the backlight, turbo CPU mode, the sound as expense in headset exclusively, network regulations (IP, DNS) and with checked the state of battery.

4.2.2 Interaction with Other Subsystems

This subsystem communicates with the MDSC subsystem to carry out the functions 2, 4 as described above. If it does not make this process the PDAs cannot be used. The MDSC interact with NSS to be informed about any available new versions of multimedia data needed to be transferred in accordance with the operation 3.

4.2.3 Correct Operation Issues

In each appliance in MDSC subsystem exist various files of application in a specific point thus as they are created afterwards the process of installation. During the upgrade process may need some files to be replaced later with files from MDMDS. It is also worth noting that each device in MDSC subsystem linked to a user of the NSS and more specifically with a user of Communication Server. The MDMDS subsystem should write to each mobile device a file of user information.

4.3 Networking Services Subsystem

This subsystem includes services and management functions of the MDSC subsystem such as location positioning, monitoring the location of visitors inside the museum, where WNIS and Communication Server is available.

It consists of the following two applications:

- Communication Server
- Monitoring and Management Visitors Administration Console.

The following are the general operating requirements of the subsystem, the way interfaces with other subsystems and analyzed in greater depth the applications that compose the particular subsystem.

4.3.1 Operational Specifications of Networking Services Subsystem

The basic functional specifications are as follows:

- Monitoring the position of visitors inside the museum in real time according to information received from the MDSC subsystem where WNIS and Communication Server are available.
- Send information messages to visitors as a unit, group or exhibition hall.
- Communication Server.

4.3.2 Interaction with Other Subsystems

The NSS ties with MDSC for sending and receiving messages. This happens through the communication server, which will be analyzed afterwards. The NSS also informed by the MDMDS with incoming and outgoing groups of visitors. The NSS makes dependent the operation completely from his availability of access on the WNIS.

4.3.3 Communication Server

Overview/Operations: This server other than the operation in support of Instant Messaging operates as a Communication Server to link subsystems. More specifically throughout the following:

- Any device (PDA) that runs the MDSC subsystem sends information about its position and the language of choice of the visitor on the Communication Server so that it manage the information encouraging them to the Monitoring and Management Visitors Administration Console Application and other users in the same group of users.
- Creation of users and user groups. NSS subsystem manage so that users belonging to the same group to share their presence. The concept of presence is the meaning of the situation of each user. For example, if a family consisting of father, mother and child come to the museum will have during the check-into join the same group with the result that everyone in this group can see at any time if any of the members of the same group is connected, in which exhibition hall he/she is found the given moment and may send a text message.

4.3.4 Monitoring and Management Visitors Administration Console

Overview/Operations: The functionality of the application divided into the following two sections:

- The first part deals with the implementation of sending and receiving text messaging with all users of the system inside the museum.
- And the second section aims to give to the system administrator an overview of what happens inside the museum in real time.
- More specifically, through this application achieved the following:
- Monitoring the position of visitors in real time according to information received from the MDSC subsystem and operation of positioning where is available and properly configured the WNIS.
- Send information messages to visitors at:
  - single visitor,
  - specific language group,
  - specific exhibition hall group.
- Management of predefined messages.
- Group contact by language and location (exhibition hall).
- Presentation of number of visitors and language of these on a map.
- Management of presentation of visitors in the map (definition top-level visitors per room with a color code).
- Instant Messaging Logging (history).

- Logging the position of the visitors inside the museum in order to come out traffic statistics for each exhibition hall.
- Recording the presentation language of visitors in the MDSC in order to come out statistics.

4.4 Data Management System

The data management system (DMS) includes the operation of publication of multimedia data and additional files with settings (new versions). Once a new update data from the DMS, these data pass into MDMDS, to transfer data to devices. This subsystem is linked solely to MDMDS. Once a new updated data from the DMS, these data pass into MDMDS, to transfer data to devices. The particular sub system is interlinked exclusively with the MDMDS.

4.5 Wireless Network Infrastructure Subsystem

The wireless network infrastructure subsystem (WNIS) contains the necessary infrastructure for the availability of wireless interconnection of MDSC, and necessary facilities for the successful positioning operation of a device/visitor, wherever available. To achieve the positioning operation of a device/visitor ought to be regulated according WNIS and as this will be determined during the process of calibration of PDAs in the area of the museum.

5 Indoor Positioning and Technologies

This section describes the technologies involved in the development of the proposed applications and the development tools used to design and implement them.

5.1 Microsoft .NET Compact Framework

The Microsoft .NET Compact Framework (.NET CF) [12] is a version of the .NET Framework that is designed to run on Windows CE based mobile/embedded devices such as PDAs, mobile phones, factory controllers, set-top boxes, etc. The .NET Compact Framework uses some of the same class libraries as the full .NET Framework and also a few libraries designed specifically for mobile devices.

In addition, .NET Compact Framework makes possible to host ActiveX Controls (Flash Player controls in our case) in an application that uses the .NET Compact Framework.

5.2 XMPP Protocol

Extensible Messaging and Presence Protocol (XMPP) [13, 14] is an open, XML-inspired protocol for near-real-time, extensible instant messaging (IM) and presence information. The protocol is built to be extensible and other features such as file transfer and Voice over IP have been added.

5.3 Openfire

Openfire [15] (previously known as Wildfire Server) is a Jabber/XMPP server written in Java. It allows the implementation of XML based messaging services to inform and allow inter communication of the visitors chat clients.

5.4 Windows Media Player Mobile

Media Player Mobile [16] closely resembles the capabilities of the Windows version of WMP. Windows Media Player Mobile software includes a default skin which is a customized user interface to Windows Media Player that provides a unique appearance and functionality for Windows Media Player Mobile. You can create your own buttons in order to seamlessly integrate your customized skin with your application.

5.5 Macromedia Flash Player

The development platform used for the creation of this application is Macromedia Flash Professional [17]. The software used on customer to run the application is Macromedia Flash Player [18]. The Flash is a tool used by designers and developers responsible for creating presentations, applications and in general content that enables interaction with the user.

5.6 Microsoft Visual Studio

Microsoft Visual Studio [19] is the main Integrated Development Environment (IDE) from Microsoft. It can be used to develop console and graphical user interface applications along with Windows Forms applications and web applications in both native code and managed code for all platforms supported by Microsoft Windows, .NET Framework and .NET Compact Framework. Visual Studio Team System provides a set of software development, collaboration, metrics, and reporting tools in addition to the features provided by Visual Studio Professional.

6 Case Study: The Digital Exhibition of the History of Ancient Olympic Games and Application Visual Evaluation

This section presents the architectural design, implementation aspects and functional details that were taken into consideration in the implementation of a real-world, large-scale system for the Digital Exhibition of the History of Ancient Olympic Games.

Language Selection: Initially, the visitor has to choose the language of the presentation as shown in Fig. 2a. This screen serves as a boot option for the user, which means that unless the user chooses the language, the service will not proceed (whether or not moving in the museum). The design of the systems allows the integration of new languages with no restrictions.

Operation Type Selection: The visitor is possible to choose among the different operation types in the respective selection screen (Fig. 2b). The user is required to select the type of operation that is automatic operation (automatic positioning of the user current position inside the museum using the Wi-Fi network) or manual (the visitor choose manual the presentation of a thematic section which would like to attend).

Tour Type Selection: Then the user will see the tour type selection screen (Fig. 2c). Options include audio tour where the visitor listens for each thematic section a brief tour, detailed audio tour where outside short acoustic tour on the thematic sections, the visitor can hear also the keys tours and finally audio-visual tour where the visitor listens but also watches video and images.

Then, depending on the type of operation the user has chosen there are the following options:

1. In case of choosing an automatic type of operation, if it is detected within an exhibition hall: (a) If yes, the application loads the presentation for the corresponding thematic section. (b) If no (for example the visitor may be at the cafeteria), the system will inform the user that is standing outside of any exhibition hall and goes into standby mode until the user to enter an exhibition hall.

Fig. 2 Mobile application screenshots starting from *upper left corner*. **a** Language selection. **b** Operation type selection. **c** Tour type selection. **d** Exhibition halls—thematic sections selection screen

2. In case of choosing a manual type of operation the user will redirect to the Thematic Section Selection Screen (Fig. 2d) and the user will manually select the presentation of a thematic section which would like to attend.

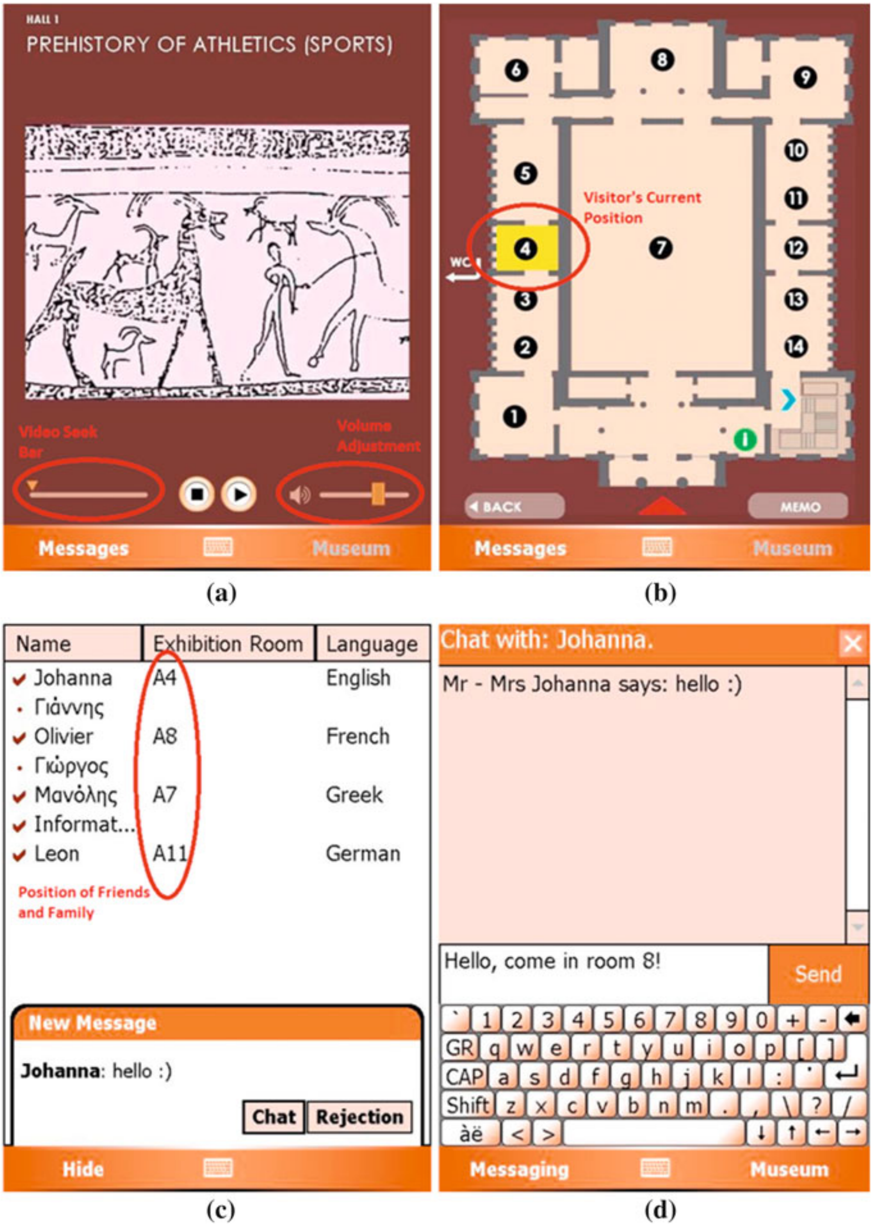


Fig. 3 **a** Audiovisual tour playback. **b** Museum map. **c** Contacts, positioning information and incoming notification bubble. **d** Mobile online chatting

Exhibition Halls—Thematic Sections Selection: After the user selects the manual type of operation will then receive a menu that contains all the exhibition halls—thematic sections (Fig. 2d).

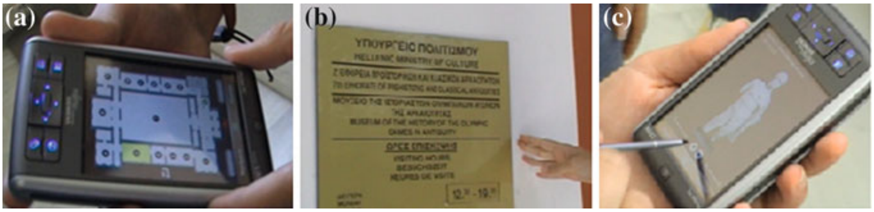


Fig. 4 Mobile application real life snapshots in the museum. **a** The visitor is automatically located automatically by the positioning system at the *yellow highlighted* exhibition room no. 5. **b** The entrance of the museum. **c** Watching and listening to a video

Instant Messaging Service: The Instant Messaging Service gives the possibility in each visitor of directly communicating with text messages (Fig. 3d) with the users that belong in the same team of users (Fig. 3c).

Positioning Service: The Positioning Service undertaken to determine the position of each visitor to the site (Figs. 3b, 4a).

Multimedia Data Provisioning: The main role of this service is the presentation of multimedia data in four languages (Figs. 3a, 4c).

7 Conclusions and Future Steps

In conclusion, the case study of the Digital Exhibition of History of Ancient Olympic Games in Ancient Olympia smartphone application for personalized recommendation museum guide has been presented. It allows the automatic indoor (and outdoor) location based recommendation for multimedia presentations about exhibits using Wi-Fi i.e. with no intervention in the museum. The proposed system is accompanied by all the supporting services and applications to provide excellent quality, innovative (state-of-the-art) technology results and efficient operational framework in real life museum. The key aim is to serve as a model for integrating ≪digital services≫ in museums and exhibitions in general with no limitations. The proposed recommendation museum guide solution is designed and provided as an integrated platform. As such it allows museum management and administration to extend or alter the exhibition adding or changing the exhibits in the future. Moreover, it is social web enabled as it allows to ≪link≫ with other museums around the world.

Future steps include the integration of image recognition features using the smartphone embedded camera. This will allow the user to watch or listen multimedia presentation in the paced—manual tour type after taking a photo of an exhibit. In this way the manual choices will turn to semi-automatic recommendations with user intervention. Furthermore, study towards the use of the mobile's camera in order to deliver augmented reality services upon exhibits indoors is also a possible direction. In fact, *indoor* augmented reality is currently implemented using either explicit user intervention or using standing (not mobile) cameras and completely differs from outdoor augmented reality approaches.

References

1 Chou, L.D., Lee, C.C., Lee, M.Y., Chang, C.Y.: A tour guide system for mobile learning in museums. In: Proceedings of the 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04), p. 195 (2004)
2. Fantoni, S.F.: Personalization through it in museums. Does it really work? The case of the marble museum website, Carrara Marble Museum, Cultural Institutions and digital technology, Ecole du Louvre, Paris, France (2003)
3. Allison, D., Gwaltney, T.: How people use electronic interactives: Information age—people, information and technology. In: Bearman, B. (ed.) Proceedings of the First International Conference on Hypermedia and Interactivity in Museums (ICHIM'91). Archives and Museum Informatics, PA, pp. 62–73. Pittsburgh, 14–16 Oct 1991
4. Chen, Y.C., Chiang, J.R., Chu, H., Huang, P., Tsui, A.W.: sensor-assisted Wi-Fi indoor location system for adapting to environmental dynamics. In: Proceedings of MSWiM '05. Montreal, Quebec, Canada, 10–13 Oct 2005
5. Farina, N.: Why indoor navigation is so hard—your phone can get you to the museum, but it can't guide you to the T-Rex. http://radar.oreilly.com/2011/10/indoor-navigation.html (Cited 15 June 2012)
6. Garofalakis, J., Sakkopoulos, E., Sirmakessis, S., Tsakalidis, A.: Integrating adaptive techniques into virtual university learning environment. In: Proceedings of the IEEE International Conference on Advanced Learning Technologies, pp. 28–33 (2002)
7. Koshizuka, N., Sakamura, K.: Tokyo university digital museum. In: Proceedings of the Kyoto International Conference on Digital Libraries (2000)
8. Lonsdale, P., Baber, H., Sharples, M., Byrne, W., Arvanitis, T. N., Brundell P., Beale, R.: Context awareness for MOBIlearn: Creating an engaging learning experience in an art museum. In: Proceedings of the 3rd MLEARN Conference, Rome, Italy, pp. 115–118 (2004)
9. Gemmell, J., Bell, G., Lueder, R., Drucker, S., Wong, C.: My—LifeBits: ful_lling the memex vision. In: Proceedings of the ACM Multimedia, pp. 235–238 (2002)
10. Healey, J., Picard, R.W.: StartleCam: A cybernetic wearable camera. In: Proceedings of the 2nd IEEE International Symposium on Wearable Computers, p. 42, 19–20 Oct 1998
11. Kawaji, H., Hatada, K., Yamasaki, T., Aizawa, K.: An image-based indoor positioning for digital museum applications. In: Proceedings of the 16th International Conference on Virtual Systems and Multimedia (VSMM), pp. 105–11 (2010)
12. Microsoft .NET Compact Framework 3.5—http://msdn.microsoft.com/en-us/netframework/aa497273.aspx (Cited 15 June 2012)
13. agsXMPP SDK—http://www.ag-software.de/index.php?page=agsxmpp-sdk (Cited 15 June 2012)
14. XMPP Protocol—http://www.xmpp.org/ (Cited 15 June 2012)
15. Openfire—http://www.igniterealtime.org/projects/openfire/index.jsp (Cited 15 June 2012)
16. Windows Media Player Mobile—http://www.microsoft.com/windows/windowsmedia/default.mspx (Cited 15 June 2012)
17. Macromedia Flash Professional 8—http://www.adobe.com/products/flash/ (Cited 15 June 2012)
18. Macromedia Flash Player 7 for Pocket PC—http://www.adobe.com/products/flashplayer_pocketpc/ (Cited 15 June 2012)
19. Visual Studio 2008—http://msdn.microsoft.com/en-us/vstudio/products/default.aspx (Cited 15 June 2012)

Author Biographies

Mersini Paschou has graduated from the Department of Computer Engineering and Informatics of the University of Patras in 2007 and obtained M.Sc. degree in "Computer Science and Engineering" in 2009. Currently, she is a PhD candidate of the same department. She has been a member of Graphics, Multimedia and GIS Laboratory since 2006 and has participated in numerous National and European projects as a software engineer. Her research interests focus on the areas of Web Services, Information Systems, Web Technologies and Mobile Web Applications.

Evangelos Sakkopoulos was born in Greece, in 1977. He holds a PhD from the Computer Engineering and Informatics Department, University of Patras, Greece since 2005. He has received the M.Sc. degree with honors and the diploma of Computer Engineering and Informatics at the same institution. His research interests include Software Engineering, Mobile Apps, Web Services, Web Engineering, Web based Education and Intranets. He has more than 40 publications in international journals and conferences at these areas.

Athanasios K. Tsakalidis was born in Katerini (Greece) in 1950. In 1973, he obtained his Diploma of Mathematics from the Mathematics Department of Aristotle University of Thessaloniki (Greece). In 1981 he obtained his Diploma of Informatics from the University of Saarland, Saarbuecken (Germany) and in 1983 he obtained his PhD in Computer Science from the same University. For 12 years, he was the student and colleague of Professor Kurt Mehlhorn (Director of Max-Planck Institute of Informatics in Germany). Today, he is a Professor in Computer Engineering and Informatics Department of University of Patras, in which he was chairman for 8 years. He was member of the Board of Directors of R.A. Computer Technology Institute (R.A.CTI), Research and Development Coordinator of the same Institute, for 10 years. His research interests include: Data Structures, Graph Algorithms, Computational Geometry, Expert Systems, GIS, Medicine Informatics, Database Systems, Multimedia, Information Retrieval, and Bioinformatics. He co-authored chapter "Data Structures" of "Handbook of Theoretical Computer Science" (Elsevier Publications and MIT-Press 1990, work of 48 authors from 12 states). He has also authored more than 200 papers in international journals and international conferences proceedings with special contribution in resolution of basic problems in the area of data structures.

Giannis Tzimas is currently an Assistant Professor in the Department of Applied Informatics in Management and Finance of the Technological Educational Institute of Mesolonghi. Since 1995, he is also an adjoint researcher in the Graphics, Multimedia and GIS Lab, Department of Computer Engineering and Informatics of the University of Patras. He graduated from the Computer Engineering and Informatics Department in 1995 and has participated in the management and development of many Research and Development projects funded by national and EU resources, as well as the private sector. He also works in the Internet Technologies and Multimedia Research Unit of the Research Academic Computer Technology Institute (RACTI), as a technical co-ordinator, from 1997 till now. His research activity lies in the areas of Web Engineering, Web Modelling, Bioinformatics, Web Based Education and Intranets/Extranets. He has published a considerable number of articles in prestigious national and international conferences and journals.

Emmanouil Viennas was born in Athens, Greece in 1985. He received the Diploma of Computer Engineering and Informatics in 2008 and the master degree in Computer Science and Engineering in 2010, both from University of Patras, Greece. His research interests lie in the areas of Web Engineering, Software Engineering, Mobile Applications and Information Systems. He is the author of more than 10 technical papers published in international journals and conferences. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Association for Computing Machinery (ACM), the Technical Chamber of Greece and the Hellenic Association of Computer Engineers and Communications.

Smart Recommendation Services in Support of Patient Empowerment and Personalized Medicine

Haridimos Kondylakis, Lefteris Koumakis, Manolis Tsiknakis, Kostas Marias, Eirini Genitsaridi, Gabriella Pravettoni, Alessandra Gorini and Ketti Mazzocco

Abstract Medicine is undergoing a revolution that is transforming the nature of healthcare from reactive to preventive. The changes are catalyzed by a new systems approach to disease which focuses on integrated diagnosis, treatment and prevention of disease in individuals. This will replace our current mode of medicine over the coming years with a personalized predictive treatment. While the goal is clear, the path is fraught with challenges. The p-medicine EU project aspires to create an infrastructure that will facilitate this translation from current medical practice to personalized medicine. This Chapter focus on current research activities related to the design and implementation of an intelligent patient empowerment platform and its services. The focus of our work concerns the nature of the interaction between health institutions and individuals, particularly the communicative relation between physicians and patients, the ways of exchanging information, the nature of the information itself and the information assimilation capabilities of the patients. Our practical focus is the domain of cancer patients, whether in normal treatment or participating in clinical trials. The ultimate objective is to implement a smart environment (recommender system) able to act as a decision support infrastructure to support the communication, interaction and information delivery process form the doctor to the patient. A prerequisite of personalized delivery of information and intelligent guidance of the patient into

H. Kondylakis · L. Koumakis · M. Tsiknakis (✉) · K. Marias · E. Genitsaridi
Computational Medicine Laboratory, FORTH-ICS, Heraklio, Crete, Greece
e-mail: tsiknaki@ics.forth.gr

M. Tsiknakis
Department of Applied Informatics and Multimedia, Technological Educational Institute, Heraklion, Crete, Greece

G. Pravettoni · A. Gorini
Centro Interdipartimentale di Ricerca e Intervento sui Processi Decisionali (IRIDe), Università degli Studi di Milano, Milan, Italy

K. Mazzocco
Department of Decision Sciences, Milan and eCancer—Cancer Intelligence Limited, Bocconi University, Milan, Italy

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*, Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_4,

his/her treatment plans is our ability to develop an appropriate and accurate profile of the user. In the p-medicine project we focus on modeling and profiling the psycho-cognitive capabilities of the patient based on questionnaires and other information features and behaviors extracted from a personal health record of the patient. In this chapter we will provide a systematic review of user profiling techniques and approaches and present our results in developing a psycho-cognitive profile of the user/patient. Subsequently we will describe the details and challenges of implementing the recommendation system and services using a combination of methods to counter-balance the intrinsic weaknesses in various algorithmic approaches. We will review solutions that have combined demographic user classes and content-based filters using implicit behavior and explicit preferences, collaborative filtering and demographic or collaborative filtering and knowledge-based filters. Finally, our approach will be fully described, which uses an adaptive user interface for the presentation of the e-consent, an ontology and a semantic web rule language to formally describe patient choices, and a reasoning engine to handle access and personalized delivery of pertinent disease related information.

1 Introduction

Personalized medicine refers to the tailoring of treatment to the individual characteristics of a patient [1]. It is mainly based on advances in the field of 'Pharmacogenomics', the study of genetic variation that uses genomic technology and the study of global properties of the human genome to gain insights about the mutual effects of relevant genes and a certain drug treatment. Advances in fundamentals of genomic medicine have facilitated the development of diagnostic tools and chemical agents targeting diseases such as cancer. By including genetic tests, one is nowadays able to predict the effectiveness of a certain drug treatment in an unique patient, thereby supporting the physician in the medical treatment decision-making process. The implementation of personalized medicine into clinical practice would lead to more cost-effective health care that is able to predict sooner, diagnose more accurately and treat more effectively, leading to a large number of lives saved.

The p-medicine EU project[1] aspires to create an infrastructure that will facilitate the translation from current medical practice to personalized medicine. Essential to the realization of personal medicine is the development of information systems capable of providing accurate and timely information about potentially complex relationships between individual patients, drugs, and tailored therapeutic options [2].

This Chapter focuses on current research activities related to the design and implementation of an intelligent patient empowerment platform and its services. The entry point for the patients is the p-medicine portal. After registering to the

[1] http://www.p-medicine.eu/

portal the patients can access all patient empowerment services provided from the platform. The environment must not only represent data in a convenient format, but data must also be translated into language that is understandable to the patient. This is because the empowerment process implies that patients are able to understand medical statements, as well as legal and ethical considerations.

To achieve this, a collection of intelligent techniques is used to construct a patient profile. A profiling server collects information from different techniques and combines them to construct patient profiles. Central technique to our approach is the ALGA-C questionnaire which is used to collect psycho-cognitive information about patients.

The benefits of constructing patient profiles are many:

- Optimize information delivery from doctors to patients: Doctors, having a graphical summary of the patient profile, can rapidly adjust the content and the level of verbal information to the patient's perceived needs and their level of understanding.
- Optimize information delivery to patients in the patient empowerment environment: Information delivery is optimized according to each specific profile. Predefined rules use patient profile to personalize the contents of the information presented and to customize ways by which users complete their tasks in the patient empowerment environment. This makes it easier for the patient to decide what interest to him/her is at the moment.
- Identify relevant trials for enrollment: Besides using patient profiling to adapt the information provided, it can also enable the automatic identification of possible clinical trials that the patient could be enrolled. His/her clinical information is matched against the eligibility criteria of several trials and possible matches are identified. This matching uses advanced algorithms over a semantically-enriched clinical trial repository.

However, before any sharing of health information can take place, patients must give their consent. Our approach uses an adaptive user interface for the presentation of the e-consent, an ontology and a semantic web rule language to formally describe patient choices, and a reasoning engine to handle the access and the management of the personal health information.

The rest of this Chapter is structured as follows: Sect. 2 provides a short review on the profiling techniques used for similar purposes. Then, Sect. 3 describes the challenges in the area and the approach we selected to respond to these challenges using patient profiling. Section 4, presents shortly the architecture of our platform and finally Sect. 5 concludes this chapter and discusses future directions.

2 Review and Methods

The uniqueness of every individual patient is not only determined by his/her unique genetic material but also by his/her unique personal psychological profile including psychological, psychosocial, cognitive and behavioral aspects. These dimensions mainly influence a patient's quality of life as well as patient satisfaction. Psychological aspects such as levels of anxiety, pain and depression as an example, substantially impact on a patient's quality of life and, in turn, influence the medical process and healing to a considerable degree.

Contemporary enhanced awareness of the strong correlation between quality of life aspects and the process of cure have led to recent attempts of finding reasonable and successful ways of including quality of life measures into the clinical process. In this section we examine in a more general sense what information constitutes a user profile, how the profile is represented, how it is acquired and built and finally how the profile information is used. Then, we focus on patient profiles and we show how these techniques are usually employed to assess the quality of life of a patient.

2.1 What Information Constitutes a User Profile?

The most common user profile contents [3] are the following:

- **User interests**: They describe the concerns of the users. They can be short-term or long-term.
- **Knowledge, background and skills**: In intelligent tutoring systems and in adaptive environments it is essential to provide proper assistance to the user and to adapt the content of the information according to user's, knowledge and skill. Some systems try to categorize users as expert, intermediate or novice depending on how well they know the application domain. Finally the user's background refers to the user's characteristics that are not directly related to the application domain.
- **Goals**: Goals represent the user's objective or purpose with respect to the application he is working with, that is what the user wants to achieve [4]. Determining what a user wants to do is not a trivial task. Plan recognition is a well-known techniques used for this purpose [5].
- **Behavior**: User's behavior with a software application is an important part of the user profile. If a given user behavior is repetitive, then it represents a pattern that can be used by an adaptive system or an intelligent agent to adapt a web site or to assist the user according to the behavior learnt. The type of behavior modeled depends on the application domain.
- **Interaction preferences**: A relatively recent component of a user profile is interaction preferences, that is, information about the user's interaction habits and preferences when he interacts with an interface agent. For example, a user may prefer warnings, suggestions, or actions on the user's behalf. In addition, the agent can provide assistance by interrupting or not interrupting the user's work.
- **Individual characteristic**: This item includes mainly demographic information such as gender, age, marital status, city, country, number of children etc. Another characteristic that is included there is the user's learning style mainly

for tutoring systems. Finally, personality traits are also important features in a user profile. A trait is a temporally stable, cross-situational individual difference. One of the most famous personality models is OCEAN [6]. This model comprises five personality dimensions: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Personality models and the methods to determine personality are subjects widely studied in psychology [7, 8]. In the area of user profiling, various methods are used to detect user's personality. For example, in [9] facial actions are used as visual cues for detecting personality.

- **Contextual Information**: The user's context is also relatively new feature in user profiling however, it is of great importance in medical area. According to Abowd et al. [10] context is any information that can be used to characterize the situation of an entity (person, place etc.). The environmental context captures the entities that surround the user whereas the personal context includes the physiological context (blood pressure, retinal pattern) and the mental context (mood, expertise, angriness, stress). The social context describes the social aspects of the current user context (friends, relatives). The spacio-temporal context describes aspects of the user context relating to the time and spatial extent for the user context.

2.2 How Profiling Information is Represented?

The most common representation of user interest is keyword-based models where interests are represented by weighted vectors or keywords [11]. Weights usually represent the relevance of the word for the user or within the topic. Another representation of user interest is through topic hierarchies [12], where each node in the hierarchy represents a topic of interest for a user, which is defined by a set of representative words. The most common representation of knowledge, background and skill is through models that keep a score on certain knowledge topics that a user might know [13] or does not know [14].

Goals or intentions can be represented in different ways such as Bayesian networks [15]. In this representation, nodes represent user tasks and arcs represent probabilistic dependencies between tasks. Given evidence of a task performed by the user, the system can infer the next most probable task and hence the user's goal.

2.3 How Profiling Information is Obtained?

In order to build a user profile, the information can be provided explicitly by the user through implicitly observation of the user's actions.

The simplest way of obtaining information about users is through the data they submit via forms or other user interfaces provided for this purpose. Especially for patients their profile information commonly assessed by patient-reported outcome measures (PROMs) including Health-Related Quality of Life (HRQL) information. PROMs can be defined as “reports coming directly from patients about how they function or feel in relation to a health condition and its therapy, without interpretation of the patients’ responses by physician or anyone else” [16]. Such measures can thus be described as instruments which provide patient-based information about health, illness and the effects of treatment. A large number of measures providing HRQL information are currently available. These instruments embrace a broad range of health dimensions such as physical, psychological and social functioning [17]. In addition to these aspects which are directly related to the quality of life of a patient, PROMs sometimes investigate broader constructs such as impairment, disability and handicap, also influencing quality of life to a substantial degree. The disadvantages of this method is that users most of the times are not willing to fill in long forms providing information about them. Moreover, users do not always write the truth when completing forms about themselves or they might not know how to express their interests or what they really want.

To cope with these disadvantages, long questionnaires are often replaced by rating specific pages, topics or sections (Facebook “Like” for example, or Google “+1”) using simple questions. Another way of providing explicit information is through the “Programming by Example/Demonstration” paradigm [11]. In this approach, the user demonstrates examples to the computer and a software agent records the interactions and writes a program that corresponds to the user’s actions.

On the other hand, in order to implicitly collect information about user’s actions, their actions should be logged and patterns should be discovered using *Data Mining*, *Information Retrieval* or *Machine Learning* techniques [18–20]. However, to be able to discover patterns the user behavior should be repetitive, and the behavior observed should be different on different users.

When no information about a user is available, Key [21] introduced another methodology which uses a stereotype as default information. Stereotypes enable the classification of users as belonging to one or more of a set of subgroups, and also the integration of the typical characteristics of these sub-groups into the individual user profile.

3 Profiling Mechanisms for Patient Empowerment

3.1 Challenges and Opportunities

In an epoch where shared decision making is gaining importance, patient’s commitment and knowledge about his/her health condition becomes more and more relevant. Nowadays patients spontaneously search for information on internet, as if they were thirsty of knowing what their future might prospect to them. Navigating around the web might provide valid information, but it can also make the patients get lost in the mess of websites, unknown sources, and inaccurate information.

With this background of possible biased information, patients interact with physicians to choose the best plan of action in order to reach the best possible outcome. As Geissler [22], Director of the European Cancer Patient Coalition and founder of Chronic Myeloid Leukemia Advocates Network, says “empowerment of patients is a pre-requisite for health, and access to high quality health information is fundamental to achieve this”.

In order to have “high quality” health information it is not sufficient that its content is accurate. It is also necessary that accurate content is comprehensible to the person who has to use that information. Health literacy has been defined as “the degree to which individuals have the capacity to obtain, process and understand basic health information and services needed to make basic health decisions” [23]. The capability to obtain, process, and understand basic health information, make patients able to look after themselves well or make good decisions on health [24]. Because health literacy is central to enhancing involvement of patients in their care, all strategies to strengthen patient’s engagement should aim to improve it. In view of that, several factors other than the mere medical jargon can impair patient processing and understanding of health information. Psychological aspects and cognitive style, for example, are two more elements that affect the way patient approach, select and retain information.

Assessment of the patient profile in the clinical context can benefit patient care in several occasions:

1. Increasing physicians’ and nurses’ awareness of patients.
2. Identifying and prioritizing problems such as physical or psychosocial problems that otherwise might have been overlooked and remain unrecognized.
3. Identifying patients’ preferences among outcome goals (that often differ substantially from the physicians’ preferences).
4. Enabling an anticipation of benefits regarding patients’ adherence to treatment.
5. Allowing for the monitoring of disease progression and treatment that may not be revealed via clinical testing.

As a result of including patient profiling into the treatment process, the patient might feel better cared for which positively influence his/her emotional functioning and the communication between the physician and the patient is facilitated and improved.

As a result, shared decision-making will be promoted. It is essential for the patient to be actively involved and to participate in the decision making process concerning treatment in order to assure that decisions are consistent with the patient’s values, preferences and general considerations. Shared decision making based on patient centered medicine involves a bi-directional information exchange between the physician and the patient [25]. For shared decision making to be effective, the content of communication in a medical consultation should include both factual data and patients’ considerations. While the former is derived from clinical tools providing information about genetics and supposed treatments, the latter is provided by patient profiling techniques providing HRQL information. Barnato et al. [26] noted that

> in an ideal world […] patients would come to a cancer consultation armed with sufficient knowledge, clarity about their personal value, and the ability to engage in a thoughtful discussion about the pros and cons of treatment options. Providers, in turn, would be prepared to support their patients, armed with an understanding of the patient's knowledge gaps, personal values about possible outcomes and treatment preferences (p. 627).

However, clinical consultations take place under conditions of limited time where physician talk sometimes overwhelms patient's preferences and considerations. This gap between an optimal and many actual encounters could be reduced by implementing smart patient profiling techniques that raise awareness of patient considerations, facilitate the discussion of these aspects and thereby actively involve the patient into the medical decision process. This shared decision making process empowers the patient because it provides him/her with the chance of making his/her own, well-discussed and well-informed, choice concerning the treatment.

Some governmental and professional organizations have advised routine screening for the presence of heightened psychological distress in cancer patients (NICE, Rebalance Action Focus Group). However, there are several barriers preventing the routine use of screening or PROMs in clinical practice. These barriers can be classified in practical, attitudinal and methodological barriers [17]. The practical barriers include a lack of IT support concerning storing and retrieval of data, a lack of time and money needed to collect, analyze and appropriately use data as well as physicians' lack of familiarity and knowledge in the field of HRQL measures. Additionally, the format and length of most existing tools is a barrier to the adoption in clinical practice because a considerable amount of time is required for administering, scoring and interpreting the self-report measures [27].

Addressing these challenges and trying to merge the two aspects of personalized medicine (genetic and psychological dimensions) we will develop a patient profiling server to capture the personal profile of the patient.

3.2 Patient Profiling Server

Patient profiling server is the central point for the patient's data analysis, and is currently in the analysis phase. Profiling questionnaire, PHR and EHR systems are the main data sources of the profiling server. These sources with specific application programming interfaces (APIs) will communicate and export the data for further analysis.

The patient profiling server will also provide the necessary services for combining the different sources. Information collected from the sources will be exploited in conjunction with the provided knowledge discovery tools, in order to form a platform for the patient empowerment.

In essence, the aim of the server is to provide the necessary methods and algorithms to collect, merge and analyze the patient's data. The server will be able to develop a patient profile and to operate as an integrated analysis environment for patient data analysis and knowledge discovery tools (Fig. 1).

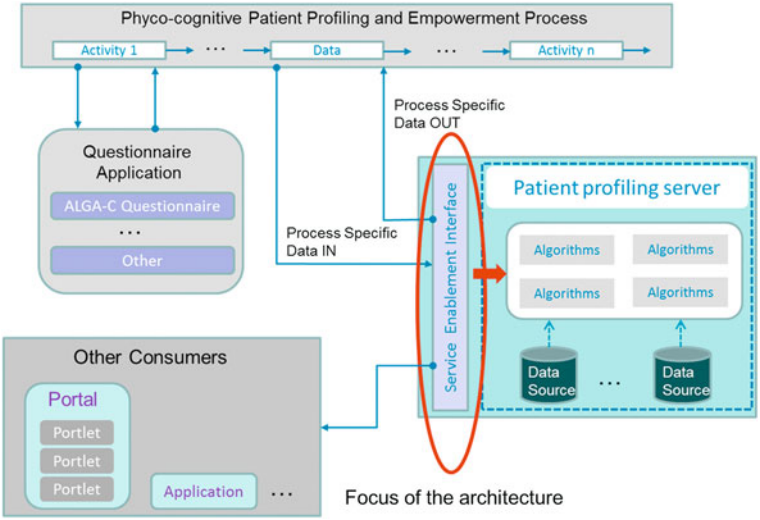


Fig. 1 Patient profiling server architecture

A variety of knowledge discovery tools exists in the public-domain (e.g., weka,[2] R-package/Bioconductor,[3] BioMoby[4]). We focus on a specific domain of data mining in order to discover patterns using Data Mining, Information Retrieval or Machine Learning techniques. Bayesian network, association rules mining and case-based reasoning are three well known techniques for solving such problems. These techniques will be evaluated and the best or all of them will be implemented within the patient profiling server.

1. A **Bayesian network** is a directed acyclic graph where nodes represent random variables and arcs represent probabilistic correlations between variables [18]. An important characteristic of BN is that Bayesian inference mechanisms can be easily applied to them. The goal of inference is typically to find the conditional distribution of a subset of the variables, conditioned on known values for some other subset (the evidence).
2. **Association Rules Mining** is a data mining technique widely used to discover patterns from data. An association rule is a rule which implies certain association relationships among a set of objects in a given domain, such as they occur together or one implies the other. Association rule mining algorithms tend to produce huge amounts of rules, most of them irrelevant or uninteresting. Therefore, post-processing steps, such as rule pruning [19] and grouping are needed to obtain valuable information to build a user profile.

[2] http://www.cs.waikato.ac.nz/ml/weka/

[3] http://www.bioconductor.org/

[4] http://biomoby.open-bio.org/

3. **Case-Based Reasoning** is a technique that solves new problems by remembering previous similar experiences [20]. A case-based reasoner represents problem-solving situations as cases. Given a new situation, it retrieves relevant cases (the ones matching the current problem) and it adapts their solutions to solve the problem.

Besides there three mainstream approaches, other are used as well such as genetic algorithms [28], neural networks [29], kNN-algorithms [30], clustering and classification techniques, fuzzy logic or combinations among them [31]. Thus the patient profiling server can be viewed as a data mining repository which will operate over patient's data sources. Algorithms and tools used within the system can evolve according to the end user's needs and the nature of the data sources.

Main instrument of collecting personal information is ALGA-C, a short, easy-to-use, and acceptable questionnaire that generates a personal profile of the patient. This personal profile questionnaire for cancer patients is aimed to produce data which support the physicians' decision making process by being merged with other data such as molecular information and imaging data, thanks to advances in IT technology. Furthermore, data generated by the instrument can also be used to monitor the patient's quality of life, thereby facilitating the patients' involvement in the clinical decision process, finally leading to patient empowerment.

3.3 ALGA-C

As previously mentioned, many approaches to personalized medicine include medical relevant data such as genetic and biomolecular information that are used to predict the best treatment choice including predictions about adverse effects. However, we claim that these approaches neglect the unique influence of psychological dimensions and cognitive style. Although a variety of biological and medical information have the power to predict the best treatment choice and thus lead to an increased life expectancy on the side of the patient, approaches that exclusively rely on biological and medical information miss a substantial part of the human being: psychology. By neglecting unique psychological information, these approaches limit their own effectiveness since psychological dimensions impact on treatment effectiveness to a considerable extent. Psychological influences limit the predictability of treatment effectiveness and thereby undermine the potential benefit of an approach that solely relies on biological and medical information. To illustrate the power of psychology regarding treatment effectiveness, we refer to the hypothesized example of a patient who is very anxious and possesses a moderate through high level of depression. A treatment that was predicted to yield good results because it was perfectly determined by analyzing biomolecular, genetic and imaging data, may turn out to be ineffective for the anxious and depressed patient. This patient might be unresponsive to the treatment due to noncompliance to the treatment caused by the patient's heightened psychological distress. Given these important psychological influences, we suppose that current approaches to personalized medicine relying exclusively on biological and medical information, lack to consider the psychological component that contributes to human uniqueness in the same way genetic information does. A human being cannot be considered as unique by only referring to him/her as a biological and genetic entity. Instead, what makes a human being unique is also his/her specific needs and value, habits and behaviors, hopes and fears, beliefs and cognitive dispositions [32].

The assessment of psychological and personal variables serves the additional purpose of monitoring the patient's perceived health status and general quality of life. Different measures for the assessment of health-related quality of life of patients have been developed. However, most of them measure only one broad area, for example, psychological distress including depression and anxiety, psycho-social problems including dimensions such as social abilities, body image and financial problems, or health status in general including symptoms resembling pain, fatigue and physical problems. We claim that there is an urgent need for an instrument that measures all these different areas in order to represent a broad picture of the patient without missing essential information. Additionally, to our knowledge, there is no single measure including cognitive aspects of the patient such as memory and attention, rumination, or self-efficacy. One of the biggest problems in developing an instrument that measures all these broad areas is the fact that such an instrument needs to be very short and easy to fill out otherwise patients are not willing to comply. It is thus a major challenge to construct an instrument that measures all these areas and that is short and easy to fill out at the same time. The ALGA-C wants to be a short and easy to fill out instrument to measure four broad areas: perceived health state, psychological aspects, psycho-social as well as cognitive aspects, all investigated by different sub-dimensions. Bellow we describe in detail the areas that the aforementioned ALGA-C questionnaire focuses, shown also on Fig. 2.

3.3.1 Quality of Life and Perceived Health State

The way a patient perceives his/her own health state is determining his/her quality of life to a large degree. Symptoms such as pain, sleep problems, fatigue, general physical problems and diminished appetite limit an individual's functioning, its perceived well-being and therefore its quality of life. Investigating these symptoms is essential as many cancer patients experience these problems to a heightened degree compared to healthy individuals.

Another symptom often related to impaired physical functioning is the loss of appetite in cancer patients as a result of chemotherapy or other treatments. In general, these illness- and treatment-related symptoms often interfere with the daily activities of a patient, lowering his/her general quality of life. Therefore, the assessment of these symptoms is of great importance in order to discover them, treat them where necessary and thereby improve patients' quality of life.

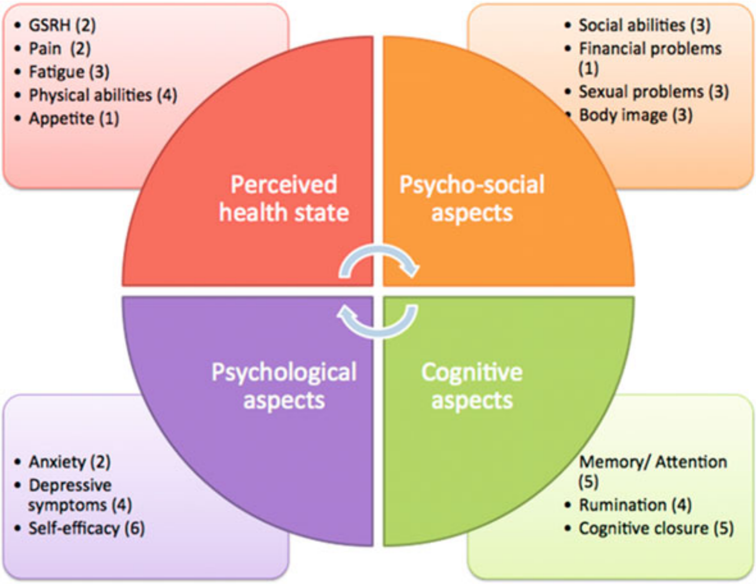


Fig. 2 A schematic representation of ALGA-C. The *circle* represents the four main macro-areas, and the *four boxes* contain the different sub-dimensions with the relative number of questions in the brackets

3.3.2 Quality of Life and Psychological Aspects

The most important psychological dimension related to quality of life is psychological distress, which includes anxiety as well as depressive symptoms. A failure to detect and treat heightened levels of psychological distress limits the effectiveness of cancer treatments, decreases patients' quality of life, and negatively influences health care costs.

Arguments for including a measure of psychological distress in cancer patients are that it has been found to be a very common phenomenon in patients [33] and that heightened levels of anxiety and depression are associated with several negative outcomes containing patient noncompliance and non-adherence to treatment recommendations [34] and decreased patient satisfaction concerning care [35]. These negative outcomes, in turn, decrease a patient's quality of life.

Given the high prevalence rates of depression, anxiety, or both in cancer patients, and the increased chance for these symptoms to persist for years after diagnosis, their detection and treatment is essential. In order to optimize treatment, it is necessary to further investigate the main risk factors for continued psychological distress and clarify differences in results yielded by previously mentioned studies.

3.3.3 Quality of Life and Psychosocial Aspects

Cancer treatment often leads in major changes in patients' physical appearance. It may, for example, result in alterations of body image through the loss of a body part, disfigurement, scars and skin changes. Whereas the effects of chemotherapy, such as hair loss, weight loss or weight gain, are rather transient and reversible to a certain degree, effects of surgery are often permanent. Also radiotherapy may have considerable effects since it may cause tissue damage with insidious changes over many years. Although differences in methodology and measurement have led to a wide range of severity and frequency of body image outcomes reported by different studies, the vast majority of research indicates that body image is a key determinant of differences in quality of life.

In general, an extensive body of literature and studies regarding the effects of treatment on body image and other psycho-social aspects has mainly focused on breast cancer patients. However, a large number of patients across many disease groups and treatment types can be affected. Since body image and sexual problems were demonstrated to result from cancer treatment, especially chemotherapy, it is important to investigate these possible effects given their potential impact on mental health, partner relationships, social support and quality of life in general.

Social and interpersonal support was also found to be critical for cancer patients after completing treatment. A lack of social support is reasoned to be a risk factor for the development of depressive symptoms. However, patients often experience that their physical conditions or other effects resulting from treatment interfere with the family life, their social activities and their working abilities. Such impairments might result in social problems leading to a limited social and interpersonal support, which, in turn, negatively influence the patient's conditions and quality of life.

Next to social and interpersonal problems, financial problems might further impact on a patient's quality of life in a negative way. A study that examined differences in the quality of life of breast cancer survivors one year after diagnosis in comparison with women from the general population suggests that almost 40.0 % of breast cancer survivors reported to have financial problems [36].

3.3.4 Quality of Life and Cognitive Aspects

Potential health, psychological and psycho-social problems that result as side effects from cancer treatment might be most obvious and accordingly cared for in the first place. However, more subtle side effects, such as possible changes in cognitive functions, have received little attention, even though many patients report to suffer from subjective changes in memory and clear thinking during and after treatment [37]. Even small impairments in cognitive functioning are expected to influence a patient's ability to function, thereby negatively impacting his/her quality of life. The fact that subtle changes in cognitive abilities are sometimes difficult to detect might be a possible explanation for the lack of appropriate measures investigating cognitive dimensions in cancer patients. However, we claim that it is crucial to measure these dimensions in order to improve patients' quality of life and additionally support the physician in the medical decision-making process by taking into account patient's cognitive characteristics.

Difficulties in the ability to remember, think and concentrate are often reported by breast cancer patients receiving chemotherapy. Evidence supporting this fact comes from a study that examined differences in cognitive functioning between breast cancer patients currently receiving chemotherapy and healthy control subjects [37]. In comparison to healthy women, breast cancer patients were found to have decreased cognitive functions, especially with regard to the memory domain. Furthermore, patients' cognitive functions were not only found to be impaired during the time of treatment but residual effects were detected even after completion of chemotherapy.

Next to the immediate impact of impairments in memory on the quality of life in patients, memory and concentration problems might also result in a patient's lowered chance of actively participating in the medical and treatment decision process when cognitive problems are not appropriately cared for. Impairments in memory and attention might account for the well-observed fact that patients do not fully understand and remember what was discussed in the encounter with the physician. By previously informing the physician about memory and attention problems, he/she might be able to adjust his/her way of communicating with the patient, thus ensuring shared decision making and patient empowerment.

Another cognitive dimension crucial of investigation is rumination which can be defined as the tendency respond to distressing situations in a maladaptive manner including repetitively and passively focusing on symptoms of distress [38]. The reason why the investigation of rumination is important in the context of cancer is the fact that it enhances the effects of depressed mood on thinking, it interferes with effective problem solving and it increases the chance of losing social support [38]. Furthermore, through these possible consequences, rumination possesses the potential to maintain and exacerbate initial depressive symptoms, possibly turning them into episodes of major depression. There is thus an increased chance for people who engage in rumination to experience prolonged periods of depression and develop clinically significant depressive disorders. Next to the correlation between rumination and depression, studies found rumination to be positively related to coping styles that include suppression or avoidance as reaction to distressing feelings and thoughts. It seems likely that the application of these coping styles as well as further impairment in concentration and memory, that is likely to result from rumination, influence treatment and cure of cancer in an undesired way. Given the fact that several treatments such as behavioral activation interventions, mindfulness and acceptance based approaches as well as interpersonal therapies were found to be effective in treating rumination, it is reasonable to investigate this dimension in cancer patients. Detecting and, where necessary, treating, patients who have the tendency to ruminate, is reasoned to positively influence depressive symptoms, treatment, cure and patients' quality of life.

Self-efficacy which is one of the core aspects of Bandura's social-cognitive theory [39] refers to an individual's belief of being able to control and adapt to a wide range of challenging and demanding situations. It is referred to as "a broad and stable sense of personal competence to deal effectively with a variety of stressful situations" [40]. Individuals possessing high levels of self-efficacy are optimistic and self-confident of their own coping abilities when confronted with life stressors such as a diagnosis of cancer. A strong sense of self-efficacy was found to be generally correlated with better health [39, 40]. The additional fact that a rather low sense of self-efficacy is associated with depression and anxiety turn the dimension self-efficacy to be of crucial in the context of cancer.

Cognitive closure is a dispositional construct that is referred to as a latent variable manifested through several different aspects including the desire for predictability, the discomfort with ambiguity and close-mindedness. We reason that cognitive closure is related to a patient's preferences for the amount of information he/she wishes to receive from the physician and might therefore impact on the patient's level of involvement in the clinical decision making process. The assessment of cognitive closure might thus assist the physician in his/her decision how much information should be best provided to the patient in order for the patient to make an informed choice. If greater levels of informed choice are associated with greater levels of patient satisfaction and compliance to treatment, we should seek to better understand and measure patients' preferences. When talking about patients' preferences, it is critical to distinguish between preferences for specific health states and utilities and preferences for specific treatments. The investigation of the latter involves the assessment of risk.

In general, better understanding of patients' preferences concerning different treatment choices is central to the concept of informed choice and shared decision making of the physician and the patient [41]. Preference based decision making involves the assessment of risks and benefits of the alternative options for action. Therefore, the investigation of patients' perceptions of risk is essential given that informed treatment decisions must be made at the individual level after consideration of perceived risks as well as benefits. Another reason for the assessment of risk perception in patients considers that the attainment of improved health care relies on patients participating in clinical trials. Research evidence suggests that educating patients about clinical trials has a low impact on their willingness to participate. Instead, patients' choice of participating in clinical trials was rather influences by their perception of risk [42].

In conclusion, considering that problems related to these four broad areas will limit the daily activities of cancer patients and thus influence their overall quality of life, an instrument that allows an early detection of these variables is crucial because timely interventions are likely to improve treatment outcomes. Furthermore, with respect to the aim of integrating psychological data with medical information, assessment of these dimensions at an early point in the health care process is necessary.

4 p-Medicine Interactive Empowerment Services

Addressing these challenges and trying to merge the two aspects of personalized medicine (genetic and psychological dimensions) we need an environment able to produce data which support the physicians' decision making process by being merged with other data such as molecular information and imaging data.

Obviously we are talking about an interactive environment. Thus, while the innovative aspect for the patient is the opportunity to have access to his/her own data, on the side of clinicians the ground-breaking element is the possibility to have access to data that is not just medical, but rather centered on values and needs of the specific patient, thorough the user's profile.

More specifically, as shown in Fig. 3, while waiting for the medical visit, the patient fills a questionnaire on an electronic device connected with the patient's personal health record. The outcome of the questionnaire is immediately stored in the PHR and re-coded in a patient's profile that is used by the physician to adjust content, level and modalities of verbal information to the patient's needs. The combined clinical and psycho-cognitive information becomes patient's and doctor's common knowledge, around which they can build together long term efficient decision-making plans (Fig. 4).

In order to cope with the previous challenges, p-medicine proposes the Interactive Empowerment Service (IEmS). The IEmS architecture is shown in Fig. 3 and combines the progress of technology and biomedical research with patients' personal needs. Patients and doctors can access the Interactive Empowerment Service through the p-medicine portal. Then, questionnaires and intelligent profiling mechanism are used in order to construct patient profiles. Thanks to collaboration among physicians, psychologists and ITs, the patient's profile can be combined with the Patient's Health Record System (PHR). PHR is one of the critical components of the IEmS, and where clinical information is patient-tagged and combined with his/her psychological, social and cognitive characteristics (patient's profile). Its importance is double: Firstly, it is a container of all the information related to the patient, such as information on diagnosis and biobank data as well as information on the possible clinical trials. Secondly, the patient his/herself has the opportunity to have access to such information. In the PHR, the patient can follow the journey of his/her data and based on it begin the decision process.

Furthermore, by having this information, the patient has the possibility to manage the consent for clinical trials and to track his biomaterials. Providing patients with "consent management" allows patients to control their own data and enhances the interaction between patients and doctors when a request for a new consent is needed, by increasing efficiency and again by involving the patient actively.

In the following we give a short description of these building blocks.

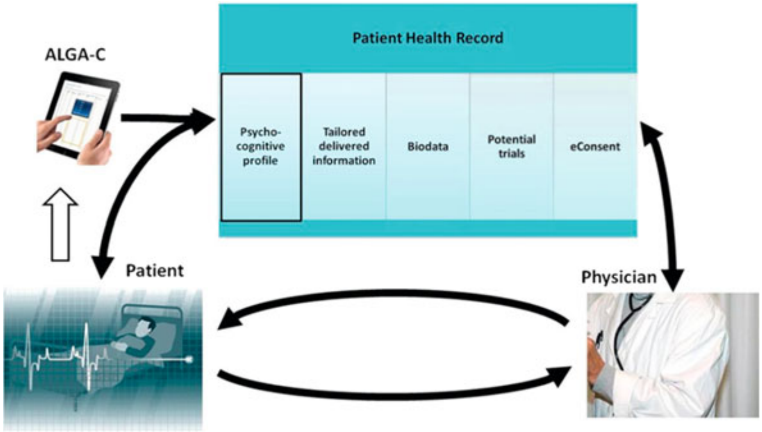


Fig. 3 Patient health record (PHR) and the information cycle in the doctor-patient interaction

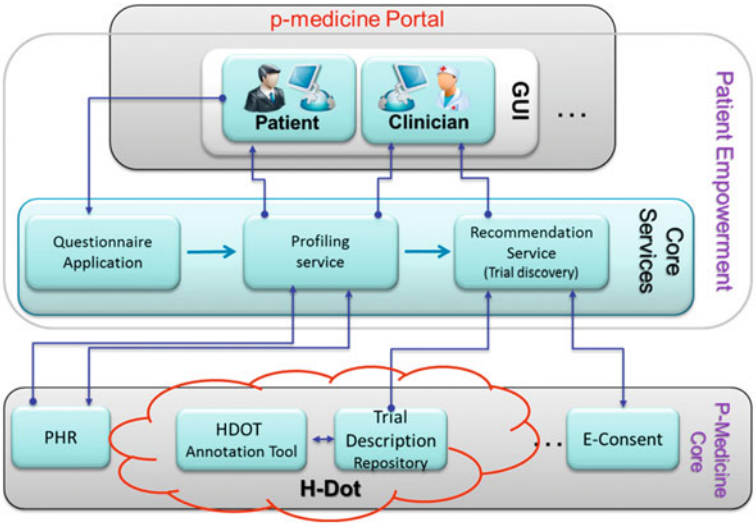


Fig. 4 High-level architecture of the p-medicine smart recommendation service

4.1 p-Medicine Portal

The p-medicine portal will be the common access point to all services provided. Through this gateway the end users will be able to search for specific tools, models, services and data based on their semantic annotations and user generated metadata.

4.2 PHR

The advancements in the health care field create a continuing demand for electronic health systems (EHS) everywhere. In the field of EHS, electronic health records (EHRs) are considered critical for improving the delivery of healthcare services. EHR systems improve accessibility to health records, support continuing treatment as repositories of information during and after the treatment, and can be used as a knowledge base for further medical research.

In p-medicine we intend to adopt an open source, widely used EHR system, Tolven. It is a free, open source, web based, robust platform that complies with many technical and operational standards and provides numerous extendable and configurable services. The Tolven[5] software environment is composed of two basic UI components, an electronic Personal Health Record solution (ePHR) and an electronic Clinician Health Record solution (eCHR). The ePHR interface enables consumers to record and selectively share healthcare information about themselves and their loved ones in a secure manner and the eCHR interface enables physicians and other healthcare providers to securely access healthcare information (collated from any number of trusted sources) relating to an individual patient in a structured and easily accessible way. The healthcare data are stored in a healthcare informatics platform and are accessed via the ePHR and eCHR solutions.

From operational point of view Tolven complies with Meaningful Use criteria that have been established by the U.S. Office of the National Coordinator for Health IT[6] (ONC). Meaningful Use criteria define high level requirements for functionality, privacy and security. It's worth mentioning that security applies at many different layers and components in the system. Tolven uses LDAP[7] for authentication, SSL for protecting data at transit while it also applies encryption on data at rest (maintained in database) through symmetric and asymmetric key cryptography. Moreover, authorization is addressed in Tolven through Role Based Access Control (RBAC) [43].

From technical point of view Tolven utilizes state of the art, industry-standard technologies such as Java, EJB3, Faces, Facelets, AJAX and relational database and supports various data formats such as CDD, CCR, and CDA documents and standardized Health Level 7[8] (HL7) messages. Moreover, Tolven is built upon an architecture that is plugin-based down to the core module. This is an important advantage for maintenance, customizations and extensions that are often required from electronic health systems in order to meet specific healthcare environment needs.

---

5 http://home.tolven.org/

6 http://healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov__home/1204

7 http://www.ietf.org/rfc/rfc2251.txt

8 http://www.hl7.org/implement/standards/index.cfm?ref=nav

4.3 HDOT Components

The p-medicine technological platform is a framework comprised by tools and services aimed at biomedical researchers and biostatisticians. The platform includes a federated Data Warehouse (DW) for storing heterogeneous data stemming from external repositories. These repositories range from private databases within hospitals and research institutions to public biomedical databases accessible through the Internet.

To achieve data transfer and semantic integration, a middle layer ontology, called HDOT, will act as global schema of the integrated data sources. Then, tools for harmonizing data sources with HDOT will be provided. The semantic layer will support the annotation of existing heterogeneous data sources with HDOT as well as the HDOT-compliant set up of new data sources for clinical trials. The Ontology Annotator Tool will be provided, which is aimed at external users (mainly database administrators) who wish to include their databases in the project framework. So Trial Description Repositories can be annotated using the HDOT annotation tool and their data will be extracted, transformed and loaded to the data ware house.

4.4 Recommendation Service

Currently, registering patients into clinical trials and finding eligible trials for patients require manual search and clinicians may be overwhelmed by the number of clinical trials and the exclusion and eligibility criteria.

However, having both PHR data and Trial Descriptions in the data warehouse p-medicine will allow the efficient recruitment of eligible patients for clinical trials. Recommendation service will use semantic matching to identify and provide suggestions of potentially eligible patients according to the available eligibility and exclusion criteria. This service will be demonstrated in the context of concrete clinical trials, with realistic data set including longitudinal EHR and clinical trial data. By automatic matching, we expect to reduce the search space with respect to the number of patients, CTs and exclusion/inclusion criteria that need to be manually reviewed to approximately 20 % of the original search space.

4.5 e-Consent

The patient's written informed consent is mandatory for research use of human biomaterial. "Multi-layered" consent, which requests from patients to make different choices on research that might or must not be performed on their samples, is increasingly recommended by ethics experts as a participative tool for patients.

Table 1 List of relevant resources

Related EU projects	
p-medicine	http://www.p-medicine.eu/
PHR systems	
Tolven	http://home.tolven.org/
Representation of profiling information	
Weighted vectors or keywords	[11]
Topic hierarchies	[12]
Scoring models	[13, 14]
Bayesian networks	[15]
Obtaining profiling information	
Questionnaires	[16, 17]
Programming by example	[11]
Simple questions	(Google "+1", facebook "like")
Data mining	[18]
Information retrieval	[19]
Machine learning	[20]
Stereotypes	[21]
Knowledge discovery tools	
WEKA	http://www.cs.waikato.ac.nz/ml/weka/
R-package	http://www.r-project.org/
BioConductor	http://www.bioconductor.org/
BioMoby	http://biomoby.open-bio.org/
Knowledge discovery methods	
Bayesian networks	[18]
Association rules mining	[19]
Case-based reasoning	[20]
Genetic algorithms	[28]
Neural networks	[29]
kNN-algorithms	[30]
Clustering and classification	[31]
Fuzzy logic	[31]

Management of multi-layered consent forms is to some extent already implemented in Biobanks Information Systems. Hence, tools for aggregating and integrating this ethical-legal data into meta biobanks and synchronizing patient's consent with scientific information will be developed in p-medicine.

A key feature of the service proposed is interactivity. With this term we refer to the possibility given to a patient to view data organized according to her/his perception of the domain, to retrieve patient-understandable information and, finally, to state decisions. IEmS will entail the tool developed in p-medicine to put people in control over the use of their data (such as type text, state decisions, upload and consult video materials). Providing patients with "consent management" offers a dual benefit: first of all there is the direct empowerment aspect of controlling one's own data; and secondly, it facilitates interaction with patients in order to ask for new consent (for new trials, secondary use of data) both increasing efficiency and again involving the patient actively. All these features need a scientific evaluation and validation before an effective use. To reach this goal, a series of experimental tests will be performed on individuals classified by age, computer skills and specific expertise through empirical user-based tests.

5 Conclusion

In this Chapter we argue that the integration of psychological and personal variables into multi-scale data systems containing heterogeneous data from a patient will greatly improve the predictive power of decision support systems developed on the basis of these data systems. Besides this, the patient might feel better cared which positively influence his/her emotional functioning. Moreover, the communication between the physician and the patient is facilitated and improved.

As a proof of concept we designed a modular patient empowerment environment where intelligent profiling techniques capture patient profile. Then the profiling information along with smart information technology resources are used to provide personalized information to the patient, and to support decision support and trial enrollment.

The intelligent environment designed and presented in this Chapter, remains to be fully implemented and tested using real patients when all technology infrastructure will be ready to support it. Potential stakeholders will evaluate the environment and developed tools will be validated to be compliant with evaluation criteria. Hopefully our platform will act as a decision support infrastructure, supporting the communication, interaction and information delivery process between doctors and patients.

Acknowledgments The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement N° 270089.

A.1 6 Appendix

Table 1.

References

1. Cawsey, A., Grasso, F., Paris, C.: Adaptive information for consumers of healthcare. In: Peter, B., Alfred, K., Wolfgang, N. (eds.) The Adaptive Web, pp. 465–484. Springer-Verlag (2007)
2. Luciano, J.S., Andersson, B., Batchelor, C., Bodenreider, O., Clark, T., Denney, C.K., Domarew, C., Gambet, T., Harland, L., Jentzsch, A., Kashyap, V., Kos, P., Kozlovsky, J.,

Lebo, T., Marshall, S.M., McCusker, J.P., McGuinness, D.L., Ogbuji, C., Pichler, E., Powers, R.L., Prud'hommeaux, E., Samwald, M., Schriml, L., Tonellato, P.J., Whetzel, P.L., Zhao, J., Stephens, S., Dumontier, M.: The translational medicine ontology and knowledge base: Driving personalized medicine by bridging the gap between bench and bedside. J. Biomed. Semant **2**(Suppl 2), S1 (2011)
3. Schiaffino, S., Amandi, A.: Intelligent user profiling artificial intelligence an international perspective. In: Bramer, M. (ed.), vol. 5640, pp. 193–216. Springer Berlin (2009)
4. Horvitz, E., Breese, J., Heckerman, D., Hovel, D., Rommelse, K.: The lumiere project: Bayesian user modeling for inferring the goals and needs of software users. In: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, pp. 256–265 (1998)
5. Greer, J., Koehn, G.: The peculiarities of plan recognition for intelligent tutoring systems (1995)
6. Goldberg, L.R.: The structure of phenotypic personality traits. Am. Psychol. **48**, 26–34 (1993)
7. McCrae, R., Costa Jr., P.T: Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In: Wiggins, J.S. (ed.) The five-factor model of personality: Theoretical perspectives, pp. 51–87. Guilford (1996)
8. Eysenck, H.J.: Personality and prediction: Principles of personality assessment. In: Wiggins J.S. (ed.) Personality and individual differences 11, p. 97 (1990)
9. Arya, A., Jefferies, L.N., Enns, J.T., DiPaola, S.: Facial actions as visual cues for personality. Comput. Animation Virtual World **17**, 371–382 (2006)
10. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a better understanding of context and context-awareness. In: Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing, pp. 304–307. Springer-Verlag, Karlsruhe, Germany (1999)
11. Lieberman, H., Fry, C., Weitzman, L.: Exploring the web with reconnaissance agents. Commun. ACM **44**, 69–75 (2001)
12. Godoy, D., Schiaffino, S., Amandi, A.: Interface agents personalizing web-based tasks. Cogn. Syst. Res. **5**, 207–222 (2004)
13. Brusilovsky, P., Sosnovsky, S., Shcherbinina, O.: User modeling in a distributed e-learning architecture. In: Proceedings of the 10th International Conference on User Modeling, pp. 387–391. Springer-Verlag, Edinburgh, UK (2005)
14. Chen, C.M., Hsieh, Y.L.: Mining learner profile utilizing association rule for common learning misconception diagnosis. In: Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies, pp. 588–592. IEEE Computer Society (2005)
15. Armentano, M., Amandi, A.: A bayesian networks approach to plan recognition for interface agents. Argentine Symposium on Artificial Intelligence, pp. 1–12 (2006)
16. Valderas, J.M., Alonso, J., Guyatt, G.H.: Measuring patient-reported outcomes: Moving from clinical trials into clinical practice. Med. J. Aust. **189**, 93–94 (2008)
17. Greenhalgh, J., Meadows, K.: The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: A literature review. J Eval. Clin. Pract. **5**, 401–416 (1999)
18. Jensen, F.V., Nielsen, T.D.: Bayesian Networks and Decision Graphs (3 edn). Springer Verlag (2007)
19. Shah, D., Lakshmanan, L.V.S., Ramamritham, K., Sudarshan, S.: Interestingness and pruning of mined patterns. In: ACM SIGMOD Workshop on Research Issues in Data Mining (1999)
20. Kolodner, J.L.: Special issue on case-based reasoning: Introduction. Mach. Learn. **10**, 195–199 (1993)
21. Kay, J.: Um: A user modeling toolkit. User Modeling Workshop (1990)
22. Geissler, J.: Informations to patients debate: As if the internet was still a walled garden (2010)
23. Ratzan, S., Parker, R.: National library of medicine current bibliographies in medicine: Health literacy. In: Health, N.I.O. (ed.), (2000)
24. Nielsen-Bohlman, L., Panzer, A.M.: Health Literacy: A Prescription to End Confusion. The National Academies Press, Washington, D.C. (2004)
25. Neuman, H.B., Charlson, M.E., Temple, L.K.: Is there a role for decision aids in cancer-related decisions? Crit. Rev. Oncol. Hematol. **62**, 240–250 (2007)

26. Barnato, A.E., Llewellyn-Thomas, H.A., Peters, E.M., Siminoff, L., Collins, E.D., Barry, M.J.: Communication and decision making in cancer care: Setting research priorities for decision support/patients' decision aids. Med. Decis. Mak. **27**, 626–634 (2007)
27. Jacobsen, P.B.: Screening for psychological distress in cancer patients: Challenges and opportunities. J. Clin. Oncol. **25**, 4526–4527 (2007)
28. Yannibelli, V., Godoy, D., Amandi, A.: A genetic algorithm approach to recognise students' learning styles. Interact. Learn. Environ. **14**, 55–78 (2006)
29. Villaverde, J.E., Godoy, D., Amandi, A.: Learning styles' recognition in e-learning environments with feed-forward neural networks. J. Comput. Assist. Learn. **22**, 197–206 (2006)
30. Lathia, N., Hailes, S., Capra, L.: kNN CF: A temporal social network. In: Proceedings of the 2008 ACM Conference on Recommender Systems, pp. 227–234. ACM, Lausanne, Switzerland (2008)
31. Martin-Bautista, M.J., Vila, M.A., Larsen, H.L.: Building adaptive user profiles by a genetic fuzzy classifier with feature selection. In: The 9th IEEE International Conference on Fuzzy Systems, vol. 301, pp. 308–312 (2000)
32. Gorini, A., Pravettoni, G.: P5-medicine: A plus for a personalized approach to oncology. Nat. Rev. Clin. Oncol. **8**, 444 (2011)
33. Zabora, J., BrintzenhofeSzoc, K., Curbow, B., Hooker, C., Piantadosi, S.: The prevalence of psychological distress by cancer site. Psychooncology **10**, 19–28 (2001)
34. Kennard, B.D., Stewart, S.M., Olvera, R., Bawdon, R.E., hAilin, A.O., Lewis, C.P., Winick, N.J.: Nonadherence in adolescent oncology patients: Preliminary data on psychological risk factors and relationships to outcome. J. Clin. Psychol. Med. Settings **11**, 31–39 (2004)
35. von Essen, L., Larsson, G., Oberg, K., Sjoden, P.O.: 'Satisfaction with care': Associations with health-related quality of life and psychosocial function among Swedish patients with endocrine gastrointestinal tumours. Eur. J. Cancer Care. **11**, 91–99 (2002)
36. Arndt, V., Merx, H., Sturmer, T., Stegmaier, C., Ziegler, H., Brenner, H.: Age-specific detriments to quality of life among breast cancer patients one year after diagnosis. Eur. J. Cancer **40**, 673–680 (2004)
37. Brezden, C.B., Phillips, K.A., Abdolell, M., Bunston, T., Tannock, I.F.: Cognitive function in breast cancer patients receiving adjuvant chemotherapy. J. Clin. Oncol. **18**, 2695–2701 (2000)
38. Nolen-Hoeksema, S., Wisco, B.E., Lyubomirsky, S.: Rethinking rumination. Perspect. Psychol. Sci. **3**, 400–424 (2008)
39. Bandura, A.: Self-Efficacy: Toward a unifying theory of behavioral change. Psychol. Rev. **84**, 191–215 (1977)
40. Scholz, U., Dona, B.G., Sud, S., Schwarzer, R.: Is general self-efficacy a universal construct? psychometric findings from 25 countries. Eur. J. Psychol. Assess. **18**, 242–251 (2002)
41. Bowling, A., Ebrahim, S.: Measuring patients' preferences for treatment and perceptions of risk. Quality in health care: QHC **10**(Suppl 1):2–8 (2001)
42. Ellis, P.M., Butow, P.N., Tattersall, M.H.N.: Informing breast cancer patients about clinical trials: A randomized clinical trial of an educational booklet. Ann. Oncol. **13**, 1414–1423 (2002)
43. Sandhu, R., Coyne, E.J., Feinstein, H.L., Youman, C.E.: Role-based access control models. IEEE Comput. (IEEE Press) **29**, 38–47 (1996)

Ontologies and Cooperation of Distributed Heterogeneous Information Systems for Tracking Multiple Chronic Diseases

Amir Hajjam

Abstract Demographic changes, particularly the ageing global population, are affecting our health systems; as the European Commission claims, 'in 2051, nearly 40 % of the population of the EU will be over 65 years'. The evolution of our ageing population leads to new challenges in terms of disability and chronic diseases, the incidence of which will increase steadily over the coming years. In this chapter, we will focus on ontologies that take a central place within the Semantic Web, on the one hand relying on conceptual representations of the areas concerned, and on the other hand allowing programmes to make inferences about them. Ontologies can significantly improve the management of complex distributed health systems. Crucially, they support multidisciplinarity by enabling communication and synchronisation between different professionals and their essential tasks.

1 Introduction

Chronic diseases are ongoing and generally incurable illnesses or conditions, such as heart disease, asthma, and diabetes. They are the leading cause of death and disability in most developed countries. These diseases are often preventable, and frequently manageable through early detection, improved diet, exercise, and treatment. Today, there are more than 15 million patients suffering from such diseases in France, a number that is expected to grow to over 20 million by 2020 [1]. It is unlikely that we will be able to place all these patients in institutions, such as hospitals and nursing homes. Consequently, homecare [2–10] is being positioned as viable service to efficiently and effectively address this issue. Homecare

A. Hajjam (✉)
Université de Technologie de Belfort Montbéliard,
IRTES-SeT 90010 Belfort CEDEX, France
e-mail: amir.hajjam-el-hassani@utbm.fr

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*,
Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_5,

solutions involving telemonitoring provide unquestionably higher quality of care and greater security than conventional practices, and ultimately lead to better quality of life for patients. These systems incorporate the most innovative technological aspects (such as monitoring and remote transmission of vital signs, fall detection and alarms) as well as providing organisational advances to coordinate different actors contributing to care. A European study Catalan Remote Management Evaluation (CARME) [11], conducted at the Spanish Hospital Germans Trias i Pujol, has shown that patients with heart failure spent 73 % less time in hospital when they used an interactive telehealth system with motivational support tools at home.

Some innovative concepts and pilot projects have been conducted throughout the world, for example:

- e-Care [12] is an intelligent platform monitoring patients with heart failure using unobtrusive sensors. Ontologies used include patient data, environmental data, and sensor data.
- The Aware Home project [13, 14] created a home environment that monitors its occupants' whereabouts and activities. The services provided by Aware Home range from enhanced social communication, such as providing a digital portrait of an elderly person to family members, to memory aids that assist users in resuming interrupted activities by using playbacks of video recordings of past events.
- Prosafe [15] is a non-intrusive monitoring system for older or disabled people who are undergoing progressive loss of autonomy and are in need of assistance. It features include a set of human presence sensors distributed in the house, a human presence data collecting system, data processing software which provides synthesised information or warnings to caregivers, and a man-machine interface.
- The ruled-based mHealth [16] system proposes a system of decision support to remotely monitor patients with heart failure. The system is based on an ontology that includes patient data such as: posture, pulse, physical activity and alerts. The decision aid is rule-based inference engine.
- The Smart IN-HOME monitoring system [17] collects data using a suite of low cost, non-intrusive sensors. The information collected is logged and analysed in an integrated data-management system. The system essentially collects information in a passive manner and does not directly interact with the person being monitored.
- Ogawa and Togawa [18, 19] have shown that health monitoring at home can be achieved even with simple sensors and instruments.
- CarerNet [20] uses viewpoint analysis to establish the needs and requirements of both the client and the care providers within a system context. It also discusses the implications for a user-centred, technology-based solution referred to as 'CarerNet'. The implementation of CarerNet is considered in terms of the enabling technologies required within the local environment of the client.

Moreover, it identifies the monitoring, networking, and system intelligence necessary for the provision of a comprehensive telecare service.

These projects vary both in the scale of their deployment and the diseases that they monitor. However, they all include various sensors to monitor the person's home and provide relevant information on the evolution of the patient's health, including information on their daily activities. All this information is backed up daily to ensure the early detection of any abnormalities, behavioural changes or abnormal vital signs, thereby allowing an alarm to be raised. The main objective is to monitor a large number of patients simultaneously. Heart failure patients alone account for about fourteen million people in Europe, with over a million new cases diagnosed every year.

The information and resources required to assess and treat various diseases is likely to explode in volume in coming years. This problem addresses how to resolve semantic conflicts between heterogeneous data sources rather than how to structure the architecture for integration. A common strategy for the resolution of such problems involves using ontologies [17] that explicitly define schema terms, thus helping to resolve semantic conflicts. Ontologies are also an integral part of bioinformatics, because they encourage common terminology for describing complex and evolving biological knowledge [21, 22].

2 Home Telemonitoring for Patients with Chronic Diseases

Home telemonitoring represents a patient management approach combining various information technologies for monitoring patients at a distance. Advances in sensors, mobile devices and embedded devices have made patient monitoring possible. They can also provide medical treatments and other health care assistance. These support solutions usually involve innovative technological and organisational solutions for the coordination of different actors, thus allowing the patient to remain at home. The infrastructure for home telemonitoring systems (Fig. 1). Generally rely on three main components:

- a portable device with associated sensors carried by the patient;
- a communication medium to transmit the information; and
- an institutional information system where data is stored and processed.

Some home telemonitoring solutions that perform complex monitoring tasks exist. These systems tend to be focused in monitoring a specific disease, and require fairly large technological equipment (for example, sensors, computers located at home and set-top boxes for the transmission information). However, most older people have comorbidities insofar as they have several chronic disorders related to age at once. Using several monitoring systems at the same time to deal with these issues is neither economical nor technically feasible. It is therefore necessary to focus on the interoperability of these systems and thereby factor out

Fig. 1 Infrastructure for home telemonitoring systems

common elements. This will reduce the costs of deployment and operation. It should also be noted that the profusion of systems is a smaller issue than it might initially appear: this is because most of these systems use similar equipment to perform their measurements. Motion sensors, for example, are found in virtually all existing solutions. However, this still means that in order to avoid multiplication of equipment, we must adopt an architecture to pool them. Solutions need to be economical as well as meeting patient needs, such as occupying minimal space within the home. In practice, this process of mutualisation and consideration of multi-pathology will be reflected by the evolution of monitoring platforms: they will be better at integrating knowledge about various diseases. Ontologies represent a formalism, which is well adapted to improving the integration and availability of new knowledge.

3 Ontology-Based Knowledge Representation

The concept of 'ontology' has been introduced in [23]. Ontologies are widely accepted as an appropriate form for conceptualising knowledge so that it can be used in several ways to model, analyse and reason upon any given domain. Roche [24] has provided a simple and generic definition: 'An ontology is a conceptualisation of a domain to which are associated one or more vocabularies of terms. The concepts are structured into a system and participate in the meaning of terms. Ontology is defined for a particular purpose and expresses a view shared by a community. An ontology is expressed in language (representation) based on a theory (semantics) that guarantees the properties of the ontology in terms of consensus, consistency, reuse and sharing.' An ontology enables the appropriate organisation of procedural knowledge so that it can be beneficial in implementing and maintaining any complex system. Ontologies are reusable and facilitate interoperability among applications [25]. They enable easier verification and comparison, as well as ensuring comparability of results derived from applications using the same ontology [26].

Ontologies are designed in a way that allows both inference of knowledge and reasoning while preserving their own semantics. Knowledge is defined in terms of concepts linked together by relationships. The ontology is presented on this basis, usually in the form of a hierarchical organisation of concepts.

Concepts are represented by a set of properties. They can be equivalent to, unconnected with, or dependent on one another. Furthermore, they can be linked by 'relations': these are defined as the connection of concepts between entities, often expressed by a convention such as a term or literal symbol. There are two types of link used: hierarchical and semantic. The hierarchical relationship assumes a hypernym–hyponym [27] structure; the semantic relationship connects the concepts in a part-whole link, which corresponds to the holonym–meronym [27] structure. A hierarchical relationship links a higher constituent, called the hypernym, and a lower constituent, the hyponym: this has the same properties as the hypernym, plus at least one additional property. Relationships between the concepts can also have algebraic properties, such as symmetry, reflexivity, and transitivity.

3.1 Knowledge Representation Formalisms

Ontology as described above needs to be formally represented. Moreover, it must represent the semantic relations linking its concepts. To this end, many formalisms have been developed, including data structure diagrams, scripts, semantic networks, rules and description logics.

- Data structure diagrams represent complex data structures. They can be considered as prototypes in describing a situation or a standard object. They provide a benchmark for comparing objects that we may wish to recognise, analyse, or classify, and should consider all possible forms for expression of knowledge. These schemes are characterised in terms of attributes (i.e., data structure), facets (i.e., attribute semantics), and relationships (i.e., inheritance semantics).
- Scripts are data structures that contain knowledge about a situation, and which combine representations. They can be seen as a set of elementary actions or references to other scenarios, ordered according to their sequence in time.
- Semantic networks involve a graphic structure that encodes some body of knowledge and its properties. The nodes of the graph represent objects (concepts, situations, events and so on), whilst the arcs express relations between these objects. These relations can be 'kind of' links, expressing an inclusive relation, 'is a' links, showing a relationship of belonging, or 'part of' links, to denote a part-whole relationship. The promise of these graphs lies in their clarity

and ease of use. This has prompted designers of multiple applications to use them, whether in knowledge acquisition, information retrieval, or reasoning about conceptual knowledge.

- Rules reflect the notion of consequence, and are formulated as 'if–then' constructs. They allow expression of various kinds of complex statements. They can be found in logic programming systems, for example the language Prolog [28] or deductive databases [29].
- Logic is a knowledge representation formalism that represents the knowledge of an application domain. It begins by defining the relevant concepts of the domain (its terminology) and then uses these concepts to specify properties of objects and individuals occurring in the domain (the world description).

Semantic networks, rules and description logics are the most prevalent. Semantic network structures can be found in RDF graph representations [30] and Topic Maps [31], whereas a formalisation of business knowledge often comes in form of rules with some 'if–then' reading, for example in business rules or logic programming formalisms. Logic is used to realise a precise semantic interpretation for both other forms. By providing formal semantics for knowledge representation languages, logic-based formalisms lay the basis for automated deduction [32].

4 Heterogeneous Medical Knowledge

The information and resources used to assess and treat various diseases are necessarily heterogeneous, making their analysis very difficult. Sources of variation can include tacit knowledge of practitioners, clinical experiences, collaborative problem-solving discussions, published medical literature, clinical practice guidelines and so on. This is the issue of semantic interoperability. A commonly accepted definition for semantic interoperability is that 'it gives meaning to the information shared and ensures that this is common sense in all systems between which exchanges must be implemented' [20, 33]. Distributed systems, which are semantically interoperable, are able to combine received information with local information whilst treating all information consistently.

Cooperation in the context of information systems is the ability to share, combine, and exchange information between multiple information sources, as well as final receivers being able to transparently access integrated information. Several approaches have been proposed to bridge the gaps amongst heterogeneous information systems. These include database translation, standardisation, federation, mediation and web-services. However, in order to achieve semantic interoperability between heterogeneous information systems, the meaning of the information that is interchanged has to be understood across the systems [34]. To ensure semantic interoperability, information shared between systems must first be described in a formal structure in order to preserve its semantics. This is a recurring problem in the field of knowledge engineering: different methodologies and techniques have been proposed to collect, identify, analyse, organise, and share knowledge between different entities. Among these techniques, ontologies have experienced the most rapid development over the past 10 years and appear to be an effective means of knowledge representation.

4.1 Techniques for Achieving Semantic Interoperability

Medical information systems need to be able to communicate complex medical concepts unambiguously, even those expressed in different languages. This is obviously a difficult task: it requires extensive analysis of the structure and concepts of medical terminologies supported by distributed heterogeneous information systems.

Ontologies provide a promising technology to solve the semantic heterogeneity problem. A number of techniques have been proposed in the literature to achieve interoperability [35]. These techniques are often used to allow data sharing between heterogeneous knowledge bases and to make recycling these bases possible. We can distinguish three main categories of techniques for achieving interoperability:

- The 'alignment of ontologies' [36] category, in which the goal is to find correspondences between ontologies. This is usually an application of the match operator [32], whose input consists of a set of ontologies and whose output is formed of correspondences between these ontologies.
- The 'mapping of ontologies' category, which allows heterogeneous knowledge bases to be queried using a common interface, or data to be transformed between different representations.
- The 'merging of ontologies' [36] category, which creates a new merged ontology with the knowledge contained in the original ontologies. The challenge then is to ensure that all correspondences and differences between ontologies are properly reflected in the resulting ontology.

Interoperability remains a difficult problem due to several reasons. Heterogeneity and instability of information sources remain the major difficulties to achieve efficient interoperability [34].

5 Cooperation of Distributed Heterogeneous Information Systems

Ontologies have been investigated in recent times and have been used as an instrument to enable cross-organisational semantic interoperability. There are many systems that propose to solve the problem of cooperation and interoperability between distributed heterogeneous information systems using ontologies.

Recent works emphasise the need for adaptive ontologies: following evolutions in data sources [37, 38], projects often use global and non-scalable ontologies [39, 40]. The project BRITE [41] aims to explore ontologies as an instrument to effectively enable European Business Registers (BRs) to respond to changes imposed on them by European Union laws and new market requirements. In BRITE, ontologies are seen as a key instrument and technology to dissolve administrative, technical, cultural and lingual barriers of BRs on one hand, and as means to provide semantic interoperability on the other hand. Seguran et al. [42] uses distributed artificial intelligence techniques, such as defined interaction protocols and scalable domain ontology, to resolve semantic conflicts coming from databases heterogeneity.

5.1 Ontologies and Cooperation

The main roles of ontologies in information integration systems are to describe the semantics of the information sources and to make their content explicit. Ontologies can be used for the identification and association of semantically corresponding information concepts. Wache et al. [43] can distinguish three main architectures according to their method for exploiting ontologies in information cooperation: single, multiple, and hybrid ontology approaches.

The single approach is the simplest, where all information sources are related to a single global ontology. Relationships clarify the semantics of the source objects and help to find semantically corresponding objects. This approach can be applied to integration problems where all the information sources due to be integrated provide nearly the same view on a domain. One global ontology provides a shared vocabulary for the specification of the semantics; it can be a combination of several ontologies [44]. TAMBIS [45], for example, proposes a global ontology approach. The task is made difficult as soon as one information source has a different view on a domain. Also, single ontology approaches are susceptible to changes in the information sources, which can affect the conceptualisation of the domain represented in the ontology. That is, changes in one information source can imply changes in the global ontology and in mappings to the other information sources. These disadvantages led to the development of the multiple ontology approach.

In the multiple ontologies approach, each information source is described by its own ontology. Mappings between ontologies can be developed to identify semantically corresponding terms within different source ontologies and can be used to express the relationships between ontologies. Mappings also have to consider different views on a domain, such as ontologies having different aggregation and granularity within their concepts. In practice, inter-ontology mapping is very difficult to define, owing to the many semantic heterogeneity problems, which can occur. OBSERVER [46], for example, manages local ontologies and provides mechanisms to map different ontologies.

Finally, the hybrid ontology approach combines the single and multiple ontology approaches to overcome their drawbacks. The semantics of each source is described by its own ontology and, in order to make the local ontologies comparable to each other, they are built from a global shared vocabulary. In hybrid approaches, we need to develop on the one hand mappings between each information source and its local ontology, and on the other hand mappings between local ontologies and the global ontology. The advantage of a hybrid approach is that new sources can easily be added without the need for modification. It also supports the acquisition and evolution of ontologies. The use of a shared vocabulary makes the source ontologies comparable and avoids the disadvantages of multiple ontology approaches. However, the drawback of the hybrid approaches is that existing ontologies cannot be reused easily, and instead have to be redeveloped from scratch.

In summary, the simplest approach is to use a single ontology. On the other hand, the multiple ontologies approach has the advantage that it makes the addition and removal of information sources easier. However, the best compromise is the hybrid ontologies approach.

6 Ontology Construction for Monitoring Patients with Chronic Diseases

In order to monitor patients with chronic diseases at home using unobtrusive sensors and to assist caregivers, we need to automate the process of collecting information from these sensors to detect and report risk situations at an early stage.

One of the first challenges for the main application that processes the data is to choose an appropriate representational formalism and to specify the conceptual resources—in other words, to define the description language—that the system can handle. This task is the aim of knowledge engineering: it reconciles the interests of formal modelling and intelligibility of resources to provide a logical notion of ontology.

The main concern of this process is to develop a concept of ontology and methodology capable of considering all the various disciplines of the medical world. Moreover, we need to have a reasoning mechanism that allows data from sensors to be interpreted, and thereby to provide an effective warning system.

6.1 Ontology Development Methodologies

There is no single protocol on how to construct any given type of ontology. An ontology can be constructed using one of the following methods: manual, automatic, and semiautomatic. In manual method, experts create a new ontology for a domain or extend an existing ontology. In automatic method, the ontology is built using knowledge extraction techniques: that is, concepts and relations are extracted and then verified by inference rules. Finally, in the semiautomatic method, ontologies are automatically built and used thereafter to extend ontologies that were previously built manually. A person who constructs the ontology needs to have some knowledge of the domain; domain experts are usually consulted to explain the meaning of domain-specific concepts.

To ensure reusability of ontologies and interoperability among different applications, it is important to use an agreed standard methodology for ontology development. Most of these standards are based on the experiences gained during developing particular ontologies. For example:

- Methontology [47], proposed in 1997, is a well-structured ontology development methodology. It was developed in the Laboratory of Artificial Intelligence at the Polytechnic University of Madrid while building an ontology in the domain of chemistry. This development methodology provides is well structured, and details a complete life cycle associated with the techniques to carry out different activities.
- KACTUS [48] methodology, proposed by, aimed to investigate how knowledge can be reused in complex technical systems and how ontologies can be used to support this. This methodology was advanced during application development, and entails reusing existing application ontologies. The steps that it involves are very abstract and not described properly from a technical point of view. Further work is necessary to make use of the advantages that description-logic ontology has over a frame-based one.
- Ontology Development 101 [49] was proposed in 2001 to target beginners in the field of ontology development. It uses a seven-step approach to ontology development: one example of its use is a Wine Ontology built in Protégé.
- The SENSUS [50] methodology aims to derive domain-specific ontologies from a large-scale ontology called SENSUS. The ontology is mainly used at a top level: it does not cover more detailed terms specific to a particular domain.
- Unified Process for ONtology building (UPON) [51] was inspired by the software development life-cycle approach. It is based on the Unified Software Development Process, where UML diagrams are used to design and evaluate ontologies. UPON aims to derive ontologies that can be used both by human beings and automated systems.

Here we have reviewed a few ontology development methodologies. A comparatively recent survey was conducted by in 2009 where methodologies were evaluated against their own defined criteria. It found that the Ontology Development Methodology 101 fulfilled the highest number of criteria (seven out of thirteen).

6.2 Steps for Creating an Ontology

It is important to determine the sources of medical information. When building an ontology, it is common to base the ontology's vocabulary on related medical guidelines. This means that all the relevant data from the guidelines have to be represented in a systematic way, using a hierarchy of concepts and relations. Other sources of medical knowledge include medical articles, other medical ontologies or terminologies, and, most importantly, experts' knowledge.

To create an ontology we need to work closely with medical experts. The objective is to define the different bodies of knowledge associated with the problem and integrate them as the project evolves. We have to model the behaviour of medical monitoring for early detection of risks, and to diagnose the evolution of the patient's condition. This step will be long-term venture. Regular and systematic integration of ontologies will be needed to upgrade the system and to improve on specific pathologies.

For the field of medicine, including chronic disease management, the creation of an ontology should go through the following steps:

1. the establishment of a corpus of work from a thesaurus, using morphosyntactic analysis to develop a list of candidate terms;
2. semantic analysis by a medical expert to validate the candidate terms;
3. semantic grouping of the validated terms; and
4. finalisation of the process in a language based on descriptive logic.

6.3 Languages and Tools for Ontologies Reasoning

To achieve semantic interoperability amongst information systems, and to ensure reusability of ontologies, we need to use a language for ontology implementations. This language should help users with well-defined syntax and semantics, and should allow the automatic reasoning.

An ontology editor such as Protégé will help to create ontology in Web Ontology Language (OWL) format [52] that can be exploited by intelligent applications. OWL is a standardised language (W3C 2004) for describing ontologies and is the most powerful language proposed for this purpose. OWL is a superset of RDF and RDF Schema and adds more vocabulary for describing properties and classes. This includes relationships between classes, cardinality as well as property characteristics such as symmetry, reflexivity and many others.

In order to reason with ontology, it is necessary to add an additional layer of knowledge representation. An ontology analysis is an activity that does not simply require only reason about assertions in the field, but also reason about the ontology itself. In other words, ontology is not a simple reference system for the implementation of reasoning based on an assertion: the reasoning must address the ontology. To use ontology for reasoning assumes that it has added operational semantics. These semantics specify how the knowledge modelled in the ontology will be used to reason and produce new knowledge automatically. There are many automated tools for ontology reasoning which permit the checking of

- context inconsistency
- context conflicts
- contradictory concepts
- subsuming relationships between classes and instances.

The current state of knowledge suggests the adoption of Semantic Web Rule Language (SWRL) [53], which is a rule language for the Semantic Web. This language is an extension of OWL that facilitates ontology definition as defined above. SWRL adds the concept of rules to existing concepts in OWL. Furthermore it adds a reasoning engine that allows it to derive new concepts or to write queries using a reasoning mechanism.

7 Summary

In this chapter we have discussed how ontologies can be used in the field of health to overcome the problem of distributed heterogeneous information systems, and to facilitate the semantic interoperability among those systems.

Most older people are affected by comorbidities insofar as they have several simultaneous chronic disorders related to age. However, most telemonitoring systems are designed to follow a particular chronic disease. In this context, the information and resources needed to assess and treat various diseases are necessarily heterogeneous, making their analysis very difficult. It is therefore necessary to focus on the interoperability of these systems.

There are many systems that propose to solve the problem of cooperation and interoperability between distributed heterogeneous information systems. Ontologies have been investigated recently and can be used as an instrument to enable cross-organisational semantic interoperability. The main role of ontologies in information integration systems is to describe the semantics of the information sources and to make their content explicit. Ontologies can be used for identifying and associating information concepts which correspond semantically. The process of designing an ontology begins after a language and a tool have been selected, and requires working closely with medical experts.

References

1. Simon, P., Acker, D.: La place de la télémédicine dans l'organisation des soins. Rapport du Ministère de la santé et des sports—Direction de l'Hospitalisation et de l'Organisation des Soins (2008)
2. http://www.h2ad.net/fr/telemonitoring-surveillance/

3. http://www.globalmed.com/solutions/mobile.php
4. http://info.vidyo.com/20120104_EMEA_Adwords_Telemedicine_French.html?gclid=CL2IrOOpgLQCFebLtAod8h8AxA
5. http://www.corwin.eu/logiciels-sante-sanitaire.html
6. http://www.rdsm.fr/
7. http://www.cardiologie-pratique.com/axistv/video/telemonitoring-resynchronisation-actualites
8. http://www.twitoo.org/
9. http://www.diatelic.com/
10. http://www.srett.com/fr/telemedecine/telesurveillance-medicale/
11. Domingo, M., Lupón, J., González, B., Crespo, E., López, R., Ramos, A., Urrutia, A., Pera, G., Verdú, J.M., Bayes-Genis, A.: Noninvasive remote telemonitoring for ambulatory patients with heart failure: effect on number of hospitalizations, days in hospital, and quality of life. CARME
12. Hajjam, A., Ahmed Benyahi, A., Andres, E., Erve, S.: e-Care : Vers une interopérabilité des ontologies pour le monitoring des maladies chroniques. Université d'été de la e-Santé, Castres (2012)
13. Georgia Institute of Technology. Aware Home. Available from: http://awarehome.imtc.gatech.edu
14. Kientz, J.A., et al.: The Georgia tech aware home. In: CHI '08 Extended Abstracts on Human Factors in Computing Systems (Florence, Italy, April 05–10, 2008). ACM, New York (2008)
15. Bourigault, D., Fabre, C., Frérot, C., Jacques, M.P., Ozdowska, S.: Syntex, analyseur syntaxique de corpus. In: Proceedings of the TALN 2005, Dourdan, juin (2005)
16. Minutolo, A., Sannino, G., Esposito, M., De Pietro, G.: A rule-based mHealth system for cardiac monitoring. In: 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (2010)
17. University of Virginia. Smart IN-HOME monitoring system: http://marc.med.virginia.edu/projectssmarthomemonitor.html
18. Etzold, T., Argos, P.: SRS—an indexing and retrieval tool for flat file data libraries. Comput. Appl. Biosci. **9**(1), 49–57 (1993)
19. Helal, S., Mann, W., El-Zabadani, H., King, J., Kaddourra, Y., Jansen, E.: The gator tech smart house: A programmable pervasive space. Computer **38**, 50–60 (2005)
20. Jouanot, F.: Un modèle sémantique pour l'interopérabilité de systèmes d'information. INFORSID 347–364 (2000)
21. Schulze-Kremer S.: Ontologies for molecular biology. In: Proceedings of the 3rd Pacific Symposium on Biocomputing, pp. 693–704. AAAI Press (1998)
22. Stevens, R., Goble, C., et al.: Building a bioinformatics ontology using OIL. IEEE Trans. Inf. Technol. Biomed. **6**(2), 135–141 (2002)
23. Gruber, T.R.: A translation approach to portable ontology specifications. Knowl. Acquisition **5**(2), 199–221
24. Roche, C.: Terminologie et ontologie. Rev. Langages **157** (2005)
25. Rector, A.L., Horrocks, I.R.: Experience building à Large, re-usable medical ontology using a description logic with transitivity and concept inclusions. In: AAAI Technical Report SS-97-06, pp. 100–107 (1997)
26. Eccher, C., Purin, B., Pisanelli, D.M., Battaglia, M., Apolloni, I., Forti, S.: Ontologies supporting continuity of care: The case of heart failure. Comput. Biol. Med. **36**, 789–801 (2006)
27. Ittoo, A., Bouma, G.: Semantic selectional restrictions for disambiguating meronymy relations. In: Proceedings of the 19th Computational Linguistics in the Netherlands Meeting, Groningen (2009)
28. Sowa, J.F.: Building, sharing, and merging ontologies. In http://users.bestweb.net/~sowa/ontology/ontoshar.htm (2001)
29. Uschold, M., Gruninger, M.: Ontologies: Principles, methods and applications. Knowl. Eng. Rev. **11**(2), 93–155 (1996)

30. Stevens, R., Goble, C.A., Bechhofer, S.: Ontology-based knowledge representation for bioinformatics. Briefings Bioinform. **1**(4), 398–414 (2000)
31. Pepper, S., Moore, G.: XML Topic Maps (XTM) 1.0. http://www.topicmaps.org/xtm/1.0
32. Grimm, et al.: Knowledge Representation and Ontologies. Logic, Ontologies and Semantic Web Languages in Studer et al.: Semantic Web Services: Concepts, Technology and Applications, pp. 51–106. Springer, Berlin. Last accessed 27 April 2009 at http://www.aifb.uni-karlsruhe.de/WBS/phi/pub/kr-onto-07.pdf (2007)
33. Vernadat, F.: Interoperable enterprise systems: Architectures, methods and metrics. In: Rapport Technique, LGIPM, Université de Metz, France (2007)
34. Ghawi, R.: Ontology-based Cooperation of Information Systems Contributions to Database-to-Ontology: Mapping and XML-to-Ontology Mapping. Thèse l'Universit_e de Bourgogne, 15 Mar 2010 (2010)
35. Choquet, R., Teodoro, D., Mels, G., Assele, A., Pasche, E., Ruch, P., Lovis, C., Jaulent, M.C.: Partage de données biomédicales sur le web sémantique. In: Proceedings of the IC'2010, Nîmes, France, June 2010
36. de Bruijn, J., Ehrig, M., Feier, C., Martín-Recuerda, F., Scharffe, F., Weiten, M.: Ontology mediation, merging and aligning. Semant Web Technol. (2006)
37. Kahng, J., Mac Leod, D.: Dynamic classificational ontologies. In: Arbid, M.A., Grethe, J. (eds.) Computing the Brain: A Guide to Neuroinformatics. Academic press, San Diego (2000)
38. Gal, A.: Semantic interoperability in information services: Experiencing with Coop-WARE. ACM SIGMOD Rec. **28**(1) (1999)
39. Arens, y., Chee, C.Y., Hsu, C., Knoblock, C.A.: Retrieving and integrating data from multiple information sources. IJICIS **2**(2), 127–157 (1993)
40. Genesereth, M.R., Keller, A.M., Duschka, O.M.: Infomaster: An information integration system. In: Proceedings of the ACM SIGMOD Conference (1997)
41. Herborn1, T., Mondorf1, A., Mougouie2, B., Wimmer1, M.A.: Semantic interoperability in the BRITE project: Ontologies as a tool for collaboration, cooperation and knowledge management. In: The 16th International Conference on Cooperative Information Systems (CoopIS 2008), Monterrey, Mexico, 12–14 November 2008
42. Seguran, M., Boulanger, D., Talens, G.: Semantic conflict resolution in heterogeneous databases: Interaction protocols for domain ontologies evolution. In: Proceedings of the 2004 International Conference on Artificial Intelligence, IC-AI '04, pp. 1053–1059 (2004)
43. Wache, H., Vogele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., and Hubner, S.: Ontology-based integration of information: A survey of existing approaches. In: IJCAI-01 Workshop: Ontologies and Information Sharing (2001)
44. Boury-Brisset, A.-C.: Ontology based approach for information fusion. In: Sixth International Conference on Information Fusion, pp. 522–529 (2003)
45. Goble, C.A., Stevens, R., Ng, G., Beshhofer, S., Paton, N.W., Baker, P.G., Peim, M., Brass, A.: Transparent access to multiple bioinformatics information sources. IBM syst. J. **40**(2), 532–551 (2001)
46. Mena, E., Illarramendi, A., Kashyap, V., Sheth, A.: OBSERVER: An approach for query processing in global information systems based on operation across pre-existing ontologies. In: International Journal on Distributed and Parallel Databases, Vol. 8, No. 2 (2000)
47. Fernandez, M., Gomez-Perez, A., Juriston, N.: METHONTOLOGY: From ontological art towards ontological engineering, In Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering, March, Stanford, USA, pp. 33–40 (1997)
48. Schreibur, A., Wielinga, B., Jansweiger, W.: The KACTUS view on the 'O' word. In: Proceedings of the 7th Dutch National Conference on Artificial Intelligence, The Netharlands, pp. 159–168 (1995)
49. Noy, N.F., McGuinness, D.: Ontology Development 101: A guide to creating your first Ontology. Stanford KSL Technical Report KSL-01-05 (2000)

50. Swartout, W., Patil, R., Knight, K., Russ, T.: Toward distributed use of large-scale ontologies. In: Proceedings of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems (1996)
51. De Nicola, A., Missikof, M., Navigli, R.: A proposal for a unified process for ontology building: UPON. Database Expert Syst. Appl. 655–664
52. Antoniou, G. and Harmelen, F.: Web Ontology Language: OWL. In Handbook on Ontologies. Berlin; New York: Springer, pp. 67–92 (2004)
53. Clark, P., Thompson, J., Porter, B.: Knowledge patterns. In: KR2000: Principles and Knowledge Representation and Reasoning, pp. 591–600

Interpreting the Omics 'era' Data

Georgios A. Pavlopoulos, Ernesto Iacucci, Ioannis Iliopoulos and Pantelis Bagos

Abstract The analysis and the interpretation of the complex and dynamic biological systems has become a major bottleneck nowadays. The latest high-throughput "omics" approaches, such as genomics, proteomics and transcriptomics and the available data repositories hosting information concerning bioentities and their properties grow exponentially in size over time. Therefore, to better understand biological systems as a whole and at a higher level, visualization is a necessity as clear and meaningful views and intuitive layouts can give a better insight into coping with data complexity. The implementation of tools to maximize user friendliness, portability and provide intuitive views is a difficult task and still remains a hurdle to overcome. In this chapter, we present a variety of significant visualization tools as they specialize in different topics covering different areas of the broad biological spectrum varying from visualization of molecular structures to phylogenies, pathways, gene expression, networks, and next generation sequencing. We emphasize their functionality, the latest research findings, and insights into how these tools could be further developed both in terms of visualization but also in the direction of data integration and information sharing.

G. A. Pavlopoulos (✉) · E. Iacucci
ESAT-SCD/iMinds-KU.Leuven Future Health Department Katholieke,
Universiteit Leuven, Kasteelpark Arenberg 10, box 2446 3001 Leuven, Belgium
e-mail: Georgios.Pavlopoulos@esat.kuleuven.be

E. Iacucci
e-mail: Ernesto.Iacucci@esat.kuleuven.be

I. Iliopoulos
Division of Basic Sciences, University of Crete Medical School, 71110 Heraklion, Greece
e-mail: iliopj@med.uoc.gr

P. Bagos
Department of Computer Science and Biomedical Informatics,
University of Central Greece, 35100 Lamia, Greece
e-mail: pbagos@ucg.gr

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*,
Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_6,

1 Molecular Structures

Starting with proteins, we introduce the four levels of protein structures. Thus, the *primary* structure refers to the amino acid sequence of the polypeptide chain, the *secondary* structure to highly regular local sub-structures such as α-helices, β-strands or β-sheets, the *tertiary* structure to the 3D structure of a single protein molecule and the *quaternary* structure to the assemblies of several protein molecules or polypeptide chains, usually called subunits in this context. As the 3D structure defines the functionality of a protein, much effort has been made in the past years in order to precisely detect it, mainly using experimental techniques such as X-Ray crystallography, NMR and electron microscopy. Simultaneously, many computational methods try to accurately predict the 3D tertiary structure of a protein given the amino-acid sequence. Today over 60 K solved protein structures are hosted in wwPDB [1] whereas ~86 % of the structures are derived from X-ray crystallography, ~13 % from NMR spectroscopy and less than ~1 % from electron microscopy [2]. Typical resolutions vary from 1.2 to 4 **Å**. Similarly to proteins, ~4 K solved RNA 3D structures are hosted in NDB [3], whereas 8 % of them correspond to PDB entries [2]. While a great variety of reviews that comment on the visualization approaches for such cases exists [4–6], here we give an overview of what is the status of the cutting-edge research in the field.

Most of the available visualization tools currently try to picture the chemistry of the biomolecules, the atoms and the bonds among them. Different representations include ribbons, space-filling atoms, ball-and-stick and others. Coloring schemes are used in order to highlight important parts of a protein such as binding sites, atoms with certain physiochemical characteristics, SNPs, active sites, different chains, exon boundaries or whole domains. Despite it is beyond the scope of this chapter to analyze all of the available visualization tools below we give some example of representative tools that are widely used and we try to categorize them according to their functionality (Table 1). Despite presenting the tools as a non-redundant list, many of them share functionalities and characteristics. For example, PyMol [7], Jmol [8], KiNG or Mage [9] offer typical views and can be incorporated in a web page. Others such as Chimera [10], SRS 3D [11], STRAP [12], Cn3D [13–15] or PdbViewer [16, 17] are able to combine the 3D structural visualization in space with the linear amino acid sequence (Fig. 1). They are highly interactive and therefore users can select regions in any of the two views and highlight the corresponding area in the other view. For example, when a sequence region is predicted to be functional or when a part of it is aligned to another sequence of interest, the 3D structural components are highlighted. This way one can look at the region of interest either from a linear or a structural point of view. Tools such as Molscript [18], PMV [19], VMD [20], ICM-Browser [21] and plusRaster3D [22] export images at a high dpi quality to be used for scientific publications. In order to superimpose two proteins and compare them directly in 3D space, tools such as MOLMOL [23], MOE, VMD [20] or PyMol [7] are suitable. In cases where computationally expensive superimposition is required, external CPU intensive packages such as STAMP [24], STRAP [12] or THESEUS [25] are recommended. Cases that require advanced computational power exist when one wants to superimpose very large regions (high size complexity) or sequences with low sequence similarity (many possible combinations). Looking at other characteristics such as hydrophobicity, electrostatics, residue conservation or connolly surfaces, MSMS software [26] is the most widely used. In order to show annotations from databases that are related to a certain part of the structure, tools such as ProSAT2 [27], JenaLib [28], PDBsum [29], SYBYL, Swiss-PdbViewer [17] or WHAT IF [30] can be used. Tools like Relibase [31, 32] and Superligands [33] can directly compare smaller molecules such as ligands between each other simultaneously. Notably, while tools such as tCONCOORD [34] and FIRST/FRODA [35] are able to picture conformational changes, Moviemaker [36] and Yale Morph [37] server applications can show two different transition stages of the same molecule. NOMAD-ref [38] and ANM [39] are can combine many transition stages but only for low frequency events. Despite the fact that few of the aforementioned tools such as PyMol [7] are also suitable for RNA structure visualization, specialized tools such as S2S Assemble [40] are implemented for RNAs.

Despite the fact that visualization of macromolecular structures is today very mature compared to other areas in biology, the current rendering techniques still lack the computational capacity to process more complex systems such as protein complexes or protein interactions at very high resolutions. In addition, molecular dynamics, simulations and motion are difficult to picture at such levels of detail, as the current tools are CPU greedy for more advanced analysis when visualization of more than one molecule per time is required. In order to come closer to a physical model and combine the chemistry-based visualization with real images from electron or cryo-microscopy great effort should still to be done in that direction towards the generation of real and more informative prototypes. In terms of data integration, tools are still available as standalone applications but a great variety of them can run as a part of a web page and come with standardized file formats and services to increase portability and data exchange.

Table 1 Software tools in the area of proteomics

Software	URL
Chimera	http://www.cgl.ucsf.edu/chimera/
FirstGlance	http://firstglance.jmol.org/
ICM-Browser	http://tinyurl.com/icm-browser/
JenaLib	http://tinyurl.com/JenaLib/
Jmol	http://www.jmol.org/
KiNG	http://tinyurl.com/KiNGapp/
Mage	http://tinyurl.com/kinemage/
MOE	http://www.chemcomp.com/
MOLMOL	http://tinyurl.com/molmol1/
Molscript	http://www.avatar.se/molscript/
NDB	http://ndbserver.rutgers.edu/

(continued)

Table 1 (continued)

Software	URL
PDBe	http://www.ebi.ac.uk/pdbe/
PDBsum	http://www.ebi.ac.uk/pdbsum/
PMV	http://tinyurl.com/PMV-MGL/
ProSAT	http://tinyurl.com/ProSAT2/
PyMOL	http://www.pymol.org/
RasMol	http://www.rasmol.org/
Raster3D	http://tinyurl.com/raster3d/
Relibase	http://tinyurl.com/relibase/
RSCB PDB	http://www.pdb.org/
SRS 3D	http://SRS3D.org/
STRAP	http://tinyurl.com/STRAP1/
Swiss-Model	http://swissmodel.expasy.org/
Swiss-PdbViewer	http://spdbv.vital-it.ch/
SYBYL	http://tinyurl.com/triposSYBYL/
VMD	http://tinyurl.com/VMD-viewer/
Chimera	http://www.cgl.ucsf.edu/chimera/
FirstGlance	http://firstglance.jmol.org/

2 Tree Hierarchies

Tree data structures and representations are widely used in biological studies in order to show hierarchies of data [41]. These include for example the Gene Ontologies (GO) [42] to describe functional annotation of genes via a hierarchically organized set of terms or the Unified Medical Language System (UMLS) [43] which serves a similar function for biomedical notions.

Another very important area of biology raises the topic of investigating and visualizing the evolution between the species. Thus, evolutionary studies try to reveal and understand how different species evolved over time and whether two different species have a common ancestor and at which time point. To picture these evolutionary transitions, phylogenetic trees are mainly used. A prime example of such tree representations is the so-called tree of life [44] which displays such evolutionary relationships between species and how they have separated and over millennia. From about $\sim$1.7 million identified species, only $\sim$80,000 of them have been analyzed for evolutionary relationships and have been assigned into a hierarchy [45] (Fig. 2).

Other areas in biology that involve high-throughput technologies such as Chip–Chip arrays, microarrays, next generation sequencing or proteomics often use tree-based clustering algorithms to interpret and visualize their results. In the case of microarrays [46–48] for example, genes are clustered according to their expression patterns in order to see which of them are correlated or anti-correlated. When one compares a healthy with a non-healthy tissue, the purpose is to find which of them

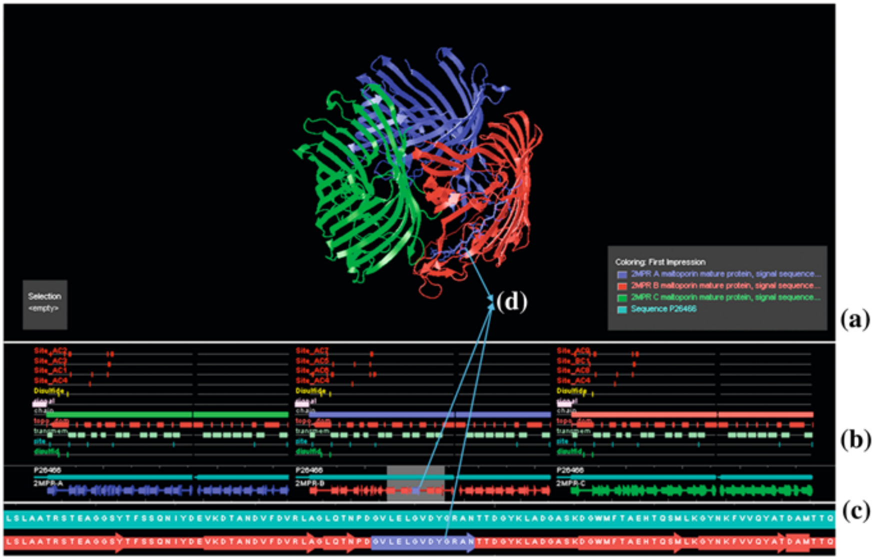


Fig. 1 P04637 (P53_HUMAN) tumor suppressor protein visualized by SRS3D application. **a** 3D structure representation of the three chains of P04637 as ribbons using three different colors. **b** Three different columns show the domains of the three different chains from different databases individually. **c** The sequence of the protein in a linear form. **d** Interactivity enables the highlighting of a chosen domain in every view (sequence and 3D structure). Switching between different representations, the 3D chemical structure of the specific domain is highlighted and visualized as a coil

are up- or down-regulated. Algorithms that are widely used, include the single linkage, average linkage, complete linkage [49], UPGMA [50], Neighbor Joining [51, 52] etc. (Figure 2).

In addition, in the field of sequence analysis, biologists try to determine the similarity between two protein or nucleotide sequences. For a given set of sequences, often a multiple alignment or an all-against-all pairwise alignment is performed constructing a distance matrix that hosts the similarity scores between every pair of genes. Notably, widely used applications that perform such analyses include the Clustal W [53], MUSCLE [54], BLAST [55], and the T-Coffee suite [56]. In order to classify these sequences in families, a clustering algorithm is applied based on the constructed similarity matrix by bringing together those sequences that are closely related to each other. The post-clustering results are visualized using a tree hierarchy (Fig. 2).

While a variety of computer readable formats exist, most phylogenetic trees are described using either the New Hampshire/Newick [57], the NHX extended Newick file format or the Nexus [58] file format. In terms of tree annotation and information sharing across repositories, Markup languages such as phyloXML [59] and NeXML are of demand.

Although the most common representation of hierarchies is a tree representation based on a 2D Euclidean drawing [60], treemaps [61] which present a tree hierarchy as nested rectangles serve as an alternative as they are often best suited for classifications rather than phylogenies [62]. While tree visualization is today a mature area, the growth of taxa is still a limiting factor as the space to represent such huge hierarchies on a single screen is insufficient. Traditional viewers that have been in use for many years such ATV [63] or TreeView [64] are nowadays weak for displaying huge taxonomies with thousands of data such as [65]. To overcome this problem, several approaches have been proposed (Table 2). One approach is the implementation of efficient zooming. Thus, as users zoom in or out, nodes collapse or expand respectively. Tools that try to compress the information into a given smaller canvas include DOI trees [66], space trees [67] and expand-ahead browsers [68]. Another approach that tools such as HyperTree [69] follow, is to project data on hyperbolic space [70]. While, this idea is very efficient in terms of visualization, in practice users find these views difficult to navigate [61]. Preferred tree visualization on the other hand involve radial layouts like those found in iTOL [71], TreeDyn [72], TreeVector [73], or Dendroscope [74]. A third

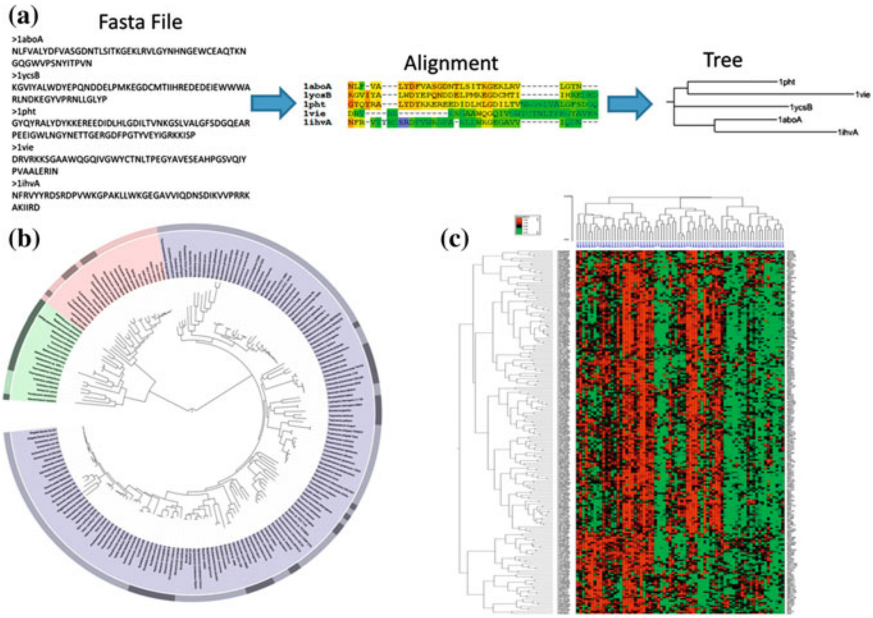


Fig. 2 Examples of tree hierarchies in Biology **a** 5 protein sequences were aligned with TCoffee and clustered according to their sequence similarity. The clustering results are shown as a tree hierarchy. **b** The Tree of Life presented in [44]. **c** Example of a gene expression heatmap. The expression levels of several genes (*tree hierarchy on the left*) were measured across several conditions (*tree hierarchy on the bottom*) using the Expander software. Genes and conditions were clustered using the average linkage hierarchical clustering. Dense red or green areas show the correlations between the genes and the experimental conditions

approach that tools such as Paloverde [75] and the Wellcome Trust Tree of Life follow, involves the utilization of 3D space. Despite the fact that such an approach is less disorientated compared to hyperbolic viewers it is still not preferred to single 2D visualization with an exception in the visualization of geophylogenies where geographical and phylogenetic information is combined towards the implementation of geographic information systems [76]. Based on such approaches, in Biology, georeferenced barcode DNA sequences are likely to become more widely used [77].

To directly compare two trees between each other so far methods such as tanglegram alignment have been proposed [78]. According to this methodology, two trees are mirrored against each other and lines connect the leaves that are equivalent to each other. Alternatively, color schemes can highlight the taxonomies that are different between each other. As tree hierarchies can vary in size and host overloaded information of thousands of taxa, direct comparison, navigation and exploration still remain a problem as the aforementioned approaches succeed in efficiently organizing the data but often fail to visually deliver them to the user in efficient ways. An example is the visualization of the tree of life versus the visualization of the forest of life [79]. While image tiling [80] methods to generate large images and then break them into smaller pieces at different resolutions (Google Earth) and recombine them could be of a solution, further opportunities for further investigation are still available in this respect.

Table 2 Tools to represent hierarchies

Software	URL
ATV	http://phylogeny.lirmm.fr/phylo_cgi/
Dendroscope	http://ab.inf.uni-tuebingen.de/software/dendroscope
Hypertree	http://kinase.com/tools/HyperTree.html
iTOL	http://itol.embl.de/
Paloverde	http://loco.biosci.arizona.edu/paloverde/paloverde.html
PhyloExplorer	http://www.ncbi.orthomam.univ-montp2.fr/phyloexplorer/
TreeDyn	http://www.treedyn.org/
TreeVector	http://supfam.cs.bris.ac.uk/TreeVector/
TreeView	http://taxonomy.zoology.gla.ac.uk/rod/treeview.html

3 Next Generation Sequencing

Recent technological improvements have led to great steps towards the understanding of the genome, its genes, their expression and their function. While the Human Genome Project (1990–2003) allowed the release of the first human reference genome by determining the sequence of $\sim$3 billion base pairs and identifying the approximately $\sim$25,000 human genes [81–83], current technologies allow the sequencing of a whole exome in a few days and at a very low cost. The first generation sequencing technique was discovered back in 1977 and is known as the Sanger (dideoxy) [24] technique. High-throughput second generation technologies have already been developed by Illumina [84], Roche/454 [85] and Biosystems/SOLiD [86] while Helicos BioSciences [87], Pacific Biosciences [88], Oxford Nanopore [89] and Complete Genomics [90] belong to the third generation of sequencing techniques. Similarly to DNA sequencing, RNA Sequencing [91, 92] which allows today the simultaneous gene expression measures in a cell and ChIP-Sequencing which uses immunoprecipitation with massively parallel DNA sequencing to mainly identify DNA regions that are binding sites for proteins such as transcription factors [93] are now more feasible and more accurate due to the rapid technological advantages as the aforementioned. Projects like the *1000 Genomes Project* (started in 2008) to sequence a large number of human genomes and provide a comprehensive resource for human genetic variation [94] and the *International HapMap Project* [95–99] to identify common genetic variations among people from different countries show the broad spectrum of the application of such technologies and the scale of the data that they can process.

Advances in high throughput next generation sequencing techniques allow the production of vast amounts of data in different formats that currently cannot be analyzed in a non-automated way. Visualization approaches are today called to cope with huge amounts of data, efficiently analyze them and deliver the knowledge to the user in a visual, easier to grasp, way. User friendliness, pattern recognition and knowledge extraction are the main targets that an optimal visualization tool should excel in. Issues such as de novo genome assemblies, SNP identification, visualization of structural variations, whole genome alignment, alignment of short reads, comparisons between several genomes simultaneously, alignment of unfinished genomes, intra/inter chromosome rearrangements, identification of functional elements and display of sequencing data and genome annotations are still open fields for visualization. Therefore, tasks like handling the overload of information, displaying data at different resolutions, fast searching or smoother scaling and navigation are not trivial when the information to be visualized consists of millions of elements and reaches an enormously high complexity. Modern libraries, able to scale millions of data points smoothly and visualize them using different resolutions are essential. While established genome browsers (Fig. 3) such as Ensembl [100, 101], UCSC Genome Browser [102] and IGV [103] are able to partially address some of the aforementioned challenges, visualization of genomic data in this respect is still an underdeveloped field.

4 Network Biology

In Systems and Integrative Biology, often bioentities are interconnected with each other and are represented as networks where nodes (bioentities) are linked with edges. Several categories of different biological networks already exist [104] such as protein–protein interactions networks, signal transduction networks, pathways, knowledge and integration networks (where bioentities are found to be related in

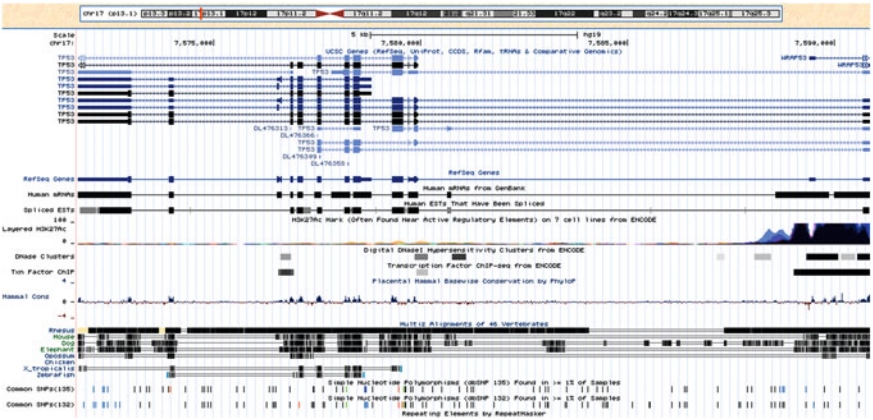


Fig. 3 P04637 (P53_HUMAN) tumor suppressor protein is found in Chromosome 17 in positions: 6,375,874–878,791,773 and visualized by the UCSC Genome browser at the highest resolutions. While a red mark shows where TP53 is in the chromosome information about alignments, SNPs, mRNA coding regions and others are shown. Notably, one can interactively zoom in and out to see the information even at the lowest nucleotide level

literature or in records of public repositories), metabolic and biochemical networks, or gene regulatory networks which picture the factors that control gene expression.

While it is not within the scope of this section to present a thorough review of available repositories for each individual network category, we shortly mention experimental, computational and high throughput techniques to detect protein–protein interactions in order to give an overview of a few of the available repositories to demonstrate the size complexity of the available data and their heterogeneity. Thus, the most widely used experimental methods include pull down assays [105], tandem affinity purification (TAP) [106], yeast two hybrid systems (Y2H) [107], mass spectrometry [108], microarrays [109] and phage display [110]. Furthermore, computational methods such as MCODE [111], jClust [112], Clique [113], LCMA [114], DPClus [115], CMC [116], SCAN [117], Cfinder [118], GIBA [119] or PCP [120] are graph-based algorithms that use graph theory to detect highly connected subnetworks. DECAFF [121], SWEMODE [122] or STM [123] have been developed to predict protein complexes incorporating graph annotations, whereas others like DMSP [124], GFA [125] and MATISSE [126] also take the gene expression data into account. A very useful review article that describes and compares the aforementioned techniques can be found in [127].

Of course, such biological networks share common characteristics but they can differ significantly in their topology and properties such as for example the number of their highly connected nodes or regions, their average eccentricity, betweeness or other types of centralities, shortest paths or their clustering coefficient. Protein–protein interaction networks tend to have hubs as signal transduction networks do not. Today, there exists a wide variety of tools (Table 3) that are network specific as reviewed in [104, 128, 129], but the field of network visualization is an active fields with many challenges to be addressed as the amount of data increases exponentially and the annotation databases expand continuously.

Currently the most widely used network representations include node-link diagrams where bioentities are represented as nodes and the interactions between them as edges forming a hairball or distance or similarity matrices which hold information about every pairwise relationship with size N(N − 1)/2 and hybrid views that combine the two previous ones. While matrices are often preferred for larger scale networks, all of the aforementioned approaches suffer in terms of scaling when the size of the network consists of few thousands of nodes and edges. In order to make large scale biological networks more informative, several layout algorithms [130] try to reveal the properties of the network such as showing the hubs using a force-directed algorithm and simultaneously try to minimize the crossovers between the lines. Similar to node-link diagrams, various ordering algorithms try to efficiently order the columns and the rows of a distance matrix to

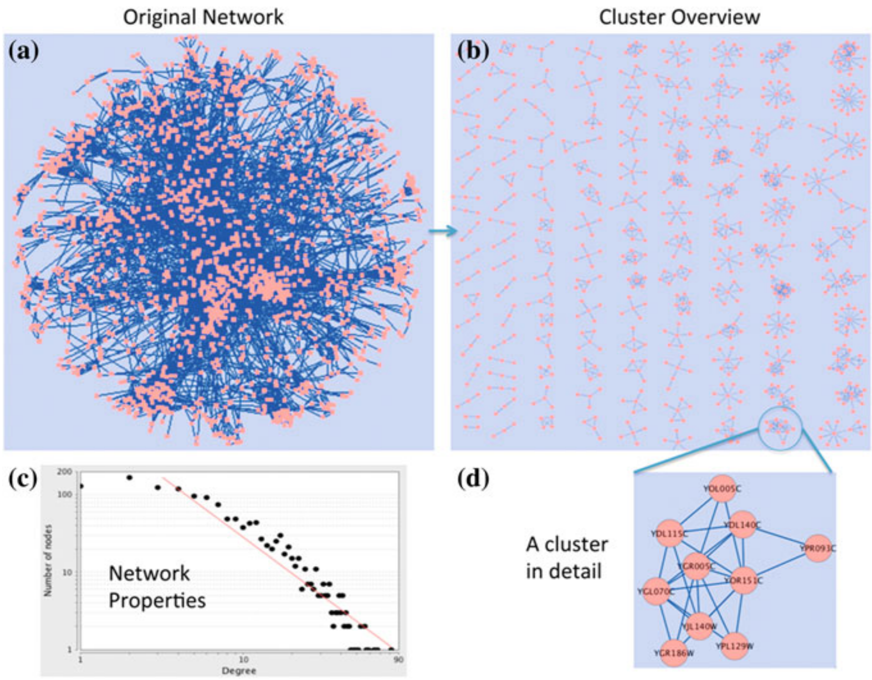


Fig. 4 A Yeast protein–protein interaction network [147] consisting of ~1600 proteins is analyzed by Cytoscape. **a** A force-directed layout algorithm is applied on the network. **b** The network was clustered according to MCL algorithm and ~240 clusters were produced. **c** The zooming functionality enables the user to the cluster and the node's labels in detail. **d** Connectivity degree versus number of nodes is plotted to show some network characteristics. Notably, different combinations of node properties can also be plotted such as the clustering coefficient (0.42 for the specific network) versus network centralization (0.053 for this network)

make highly connected regions more visible. While established tools such as Ondex [131], Pajek [132], Cytoscape [133], Medusa [134], VisANT [135] for real world-like networks or iPath [136], PATIKA [137], PathVisio [138] for pathways or EXPANDER [139], HCE [140], ExpressionProfiler [141] for expression data implement such layouts most of them try to project data a 2D plane using advanced navigation techniques to make data exploration easier. Other tools such as Arena3D [142, 143] or BioLayout Express 3D [144] take advantage of 3D space to show data in a universe. While still very few of the aforementioned tools try to fill the gap between analysis and visualization, efforts have been made the past years. ClusterMaker [145] Cytoscape's plugin and jClust [112] applications for example try to cluster the data within the application without the help of an external application (Fig. 4). Similarly, CentiBiN application [146] tries to compute and visualize different vertex and graph centrality measures.

While network visualization is a developing area, there is much space for improvements as for example visualization of time series data, network evolution and dynamics are still important features to be visually represented. Similarly, node aggregation, edge bundling, faster and more efficient layout algorithms and their extension into 3D space, multi-dimensional data visualization, semantic zooming, interactivity and data integration still remain open problems in network biology.

Table 3 Tools in network biology

Software	Type	URL
Arena3D	Network	http://www.arena3d.org/
BioLayout Express 3D	Network	http://www.biolayout.org/
Cytoscape	Network	http://www.cytoscape.org/
Medusa	Network	https://sites.google.com/site/medusa3visualization/
Ondex	Network	http://www.ondex.org/
Osprey	Network	http://tinyurl.com/osprey1/
Pajek	Network	http://pajek.imfm.si/
STITCH	Network	http://stitch.embl.de/
VisANT	Network	http://visant.bu.edu/
iPath	Pathways	http://pathways.embl.de/
Patika	Pathways	http://www.patika.org/
PathVisio	Pathways	http://www.pathvisio.org/
EXPANDER	Expression	http://acgt.cs.tau.ac.il/expander/
ExpressionProfiler	Expression	http://tinyurl.com/exprespro/
HCE	Expression	http://tinyurl.com/HCExplorer/

5 Visualization in Biology—the Present and the Future

In the aforementioned sections we widely discussed visualization tools which may be applied on different biological areas of the "omics" spectrum. These mainly include software for genome analysis, microarrays, molecular structures, phylogenies, alignments and network biology. Despite the tremendous efforts for the development of better, more efficient, more interactive and user friendlier visualization tools which has been going on over approximately the past 20 years [148] and despite the fact that all of these tools share common characteristics, the future challenges partially overlap and many difficulties still need to be addressed.

So far, there is a tendency to produce tools that mainly run as standalone applications being able to read their own file formats. While this has slowly changed over time, it is still a limiting factor, as integration needs to come to the foreground. Thus, visualization tools should share common human and computer readable file formats in order to easier exchange information. Such a demand for integration can be partially solved whenever each tool is launched with its own API or by implementing specific web services for data exchange. In addition, it is highly recommended to make tools directly available through a web interface (i.e. Flash, JavaFX, Processing.org, Applets) or directly make them downloadable through other technologies such as JNLP (Java web start) in the case of a Java implementation. Such an effort for integration would greatly help to further bridge the gap between analysis and visualization as visualization tools often use external packages to perform a typical analysis that is not embedded in the tool. A visible example of such a gap can be observed whenever one works with network biology where the nodes usually represent bioentities and the edges the connections between those. As such networks can increase in size and complexity, clustering analysis to categorize data and investigate the clusters individually is often in demand. Unfortunately, today very limited number of visualization tools hosts such functionality to cluster data within a visual application. Another example can be given for genomic data analysis where tasks such as SNP and variation calling, genome assembly or genome alignments should initially be performed individually and sequentially, the results of the analysis should be visualized by different tools after reformatting them to the tool-specific input format. In conclusion, it is expected in the future, the visualization tools will follow golden standards both in terms of data storage, analysis and integration (as to manually merge software packages and combine their functionalities requires some expertise, something that is tedious and time consuming). A first step would provide tools with a pluggable architecture where users can implement their own plugin for a tool based on their own expertise.

During the past 10 years, a progress has been made to move away from static images and cope with the increasing size complexity by handling biological data interactively. This includes operations such as efficient zooming, panning and navigation. Noticeably, multi-touch screens today encourage more modern and less conservative interfaces to handle multiple events simultaneously to increase interactivity. Similarly, vibrations could potentially be used to get the attention of the user when a property or a characteristic of the system changes. A characteristic example is the MacOS systems where icons start to vibrate in order to indicate that the corresponding process is running. Apart from these operations, in biology, often data need to be explored at multiple-scales and at different resolutions. Similarly to GoogleEarth application, which can be used to explore maps from different heights, one could imagine the biological world as a universe that can be observed from the organism to a cell or to an atomic level. In the case of genome browsers for example, a genome can be explored at a chromosome, at a gene or even at a single nucleotide resolution. Similarly, in biological networks, node aggregation or edge bundling methodologies could be applied on the network while exploring it at different levels. In order to explore multi-scaled data, pre-processing and pre-indexing is normally required, as the enormous amount of data does not allow such calculations on the fly. GoogleEarth application is a great approach to be followed as real-world images that refer to a specific resolution are indexed and stored in a database and get loaded on the fly upon users request.

Besides user interface challenges, progress in biological data management has been made over the past years. Current technologies, infrastructures and architectures allow the parallel processing of information at significantly lower costs. Unfortunately, not many visualization tools for biology today are engaged to these technologies with an exception being the tools that are implemented for biological image analysis as in the case of microscopy. Taking advantage of libraries like CUDA, which allows parallel calculations at multiple Graphical Processing Units (GPUs), protocols like Message Passing Interface (MPI) to distribute computational tasks to computers over the network or other multi-core supercomputers with multiple CPUs are ways to significantly reduce the processing time and the running time complexity of huge-scale data. Similar to architectures, display hardware such as large screens, tiled arrays or virtual reality environments, which take advantage of a very large space to project data, should be taken into consideration by programmers and designers as they become more and more affordable over time. A great advantage of such technologies is that they allow the representation of the dataset as a whole without the need of algorithms to project data at lower dimensions, something that can lead to information loss.

As visualization in biology evolves rapidly, a great variety of new visualization concepts and representations appear. While this is encouraging and it can become a source of inspiration for other fields such as economics, physics, environmental or social studies, golden standards concerning the design, the interactivity and the prototyping should be strictly defined, aiming to maximize human–computer interaction. In addition, as visualization tools are designed for a broad range of users, prototypes should take into consideration rare cases like the careful choice of color schemes as 10 % of the population suffers from color-blindness.

As biological systems are highly dynamic, visualization tools to capture the behavior and property changes of such systems and how they evolve over time is a necessity. Approaches that picture how the properties of a system change, currently include parallel coordinates, 3D representations using multi-layered graphs or animations. The effectiveness of the animation-approach is however often very low and limited by human perception capabilities, mainly due to changes in the user's mental map of the structure. More efficient approaches should be implemented to tackle this problem, as time-series visualization for biology is still a very immature field.

Finally, an ideal visualization system of the future should be able to track users preferences and learn users behavior while he or she explores specific data types. After training, such a system could guess and suggest possible solutions that anticipate the users preference, something that would minimize the time–cost to solve a problem. SVMs, SOMs, neural networks and other approaches have significantly evolved and can be used as initial steps for such user profiling. Concerning data parameterization, today visual analytics approaches that require human judgment are followed as data properties and results can vary significantly as one changes the parameters of a software or workflow. Automation of such procedures like optimal parameterization finding and profiling still remain a bottleneck.

Acknowledgments We would like to acknowledge Dr. Theodoros Soldatos, Dr. Charalampos Moschopoulos and Mgr. Izabella Januszewska for their valuable input. This work was supported by the Greek State Scholarship Foundation (I.K.Y—http://www.iky.gr/IKY/portal/en).

References

1. Berman, H., Henrick, K., Nakamura, H.: Announcing the world wide Protein Data Bank. Nat. Struct. Biol. **10**(12), 980 (2003)
2. O'Donoghue, S.I., Goodsell, D.S., Frangakis, A.S., Jossinet, F., Laskowski, R.A., Nilges, M., Saibil, H.R., Schafferhans, A., Wade, R.C., Westhof, E., Olson, A.J.: Visualization of macromolecular structures. Nat. Methods **7**(Suppl 3), S42–S55 (2010)
3. Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.H., Srinivasan, A.R., Schneider, B.: The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. Biophys. J. **63**(3), 751–759 (1992)
4. Tate, J.: Molecular visualization. Methods Biochem. Anal. **44**, 135–158 (2003)
5. Olson, A.J., Pique, M.E.: Visualizing the future of molecular graphics. SAR QSAR Environ. Res. **8**(3–4), 233–247 (1998)
6. Goddard, T.D., Ferrin, T.E.: Visualization software for molecular assemblies. Curr. Opin. Struct. Biol. **17**(5), 587–595 (2007)
7. PyMol: http://www.pymol.org/
8. jMol: http://jmol.sourceforge.net/
9. Richardson, D.C., Richardson, J.S.: The kinemage: A tool for scientific communication. Protein Sci. **1**(1), 3–9 (1992)
10. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E.: UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. **25**(13), 1605–1612 (2004)
11. O'Donoghue, S.I., Meyer, J.E., Schafferhans, A., Fries, K.: The SRS 3D module: Integrating structures, sequences and features. Bioinformatics **20**(15), 2476–2478 (2004)
12. Gille, C., Frommel, C.: STRAP: Editor for STRuctural alignments of proteins. Bioinformatics **17**(4), 377–378 (2001)
13. Porter, S.G., Day, J., McCarty, R.E., Shearn, A., Shingles, R., Fletcher, L., Murphy, S., Pearlman, R.: Exploring DNA structure with Cn3D. CBE Life Sci. Educ. **6**(1), 65–73 (2007)
14. Wang, Y., Geer, L.Y., Chappey, C., Kans, J.A., Bryant, S.H.: Cn3D: Sequence and structure views for Entrez. Trends Biochem. Sci. **25**(6), 300–302 (2000)

15. Hogue, C.W.: Cn3D: A new generation of three-dimensional molecular structure viewer. Trends Biochem. Sci. **22**(8), 314–316 (1997)
16. Guex, N., Peitsch, M.C., Schwede, T.: Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis **30**(Suppl 1), S162–S173 (2009)
17. Guex, N., Peitsch, M.C.: SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis **18**(15), 2714–2723 (1997)
18. Esnouf, R.M.: An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J. Mol. Graph. Model. **15**(2), 132–134, 112–133 (1997)
19. Sanner, M.F.: A component-based software environment for visualizing large macromolecular assemblies. Structure **13**(3), 447–462 (2005)
20. Humphrey, W., Dalke, A., Schulten, K.: VMD: Visual molecular dynamics. J. Mol. Graph. **14**(1), 33–38, 27–38 (1996)
21. ICM-Browser: http://www.molsoft.com/icm_browser.html
22. Merritt, E.A., Murphy, M.E.: Raster3D version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **50**(Pt 6), 869–873 (1994)
23. Koradi, R., Billeter, M., Wuthrich, K.: MOLMOL: A program for display and analysis of macromolecular structures. J. Mol. Graph. **14**(1), 51–55, 29–32 (1996)
24. Russell, R.B., Barton, G.J.: Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels. Proteins **14**(2), 309–323 (1992)
25. Theobald, D.L., Wuttke, D.S.: THESEUS: Maximum likelihood superpositioning and analysis of macromolecular structures. Bioinformatics **22**(17), 2171–2172 (2006)
26. Sanner, M.F., Olson, A.J., Spehner, J.C.: Reduced surface: An efficient way to compute molecular surfaces. Biopolymers **38**(3), 305–320 (1996)
27. Gabdoulline, R.R., Ulbrich, S., Richter, S., Wade, R.C.: ProSAT2–protein structure annotation server. Nucleic Acids Res. **34**(Web Server issue), W79–W83 (2006)
28. Reichert, J., Suhnel, J.: The IMB Jena image library of biological macromolecules: 2002 update. Nucleic Acids Res. **30**(1), 253–254 (2002)
29. Laskowski, R.A.: Enhancing the functional annotation of PDB structures in PDBsum using key figures extracted from the literature. Bioinformatics **23**(14), 1824–1827 (2007)
30. Vriend, G.: WHAT IF: A molecular modeling and drug design program. J. Mol. Graph. **8**(1), 52–56, 29 (1990)
31. Gunther, J., Bergner, A., Hendlich, M., Klebe, G.: Utilising structural knowledge in drug design strategies: Applications using Relibase. J. Mol. Biol. **326**(2), 621–636 (2003)
32. Hendlich, M., Bergner, A., Gunther, J., Klebe, G.: Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. J. Mol. Biol. **326**(2), 607–620 (2003)
33. Michalsky, E., Dunkel, M., Goede, A., Preissner, R.: SuperLigands—a database of ligand structures derived from the Protein Data Bank. BMC Bioinf. **6**, 122 (2005)
34. Seeliger, D., De Groot, B.L.: tCONCOORD-GUI: Visually supported conformational sampling of bioactive molecules. J. Comput. Chem. **30**(7), 1160–1166 (2009)
35. Thorpe, M.F., Lei, M., Rader, A.J., Jacobs, D.J., Kuhn, L.A.: Protein flexibility and dynamics using constraint theory. J. Mol. Graph. Model. **19**(1), 60–69 (2001)
36. Maiti, R., Van Domselaar, G.H., Wishart, D.S.: MovieMaker: A web server for rapid rendering of protein motions and interactions. Nucleic Acids Res. **33**(Web Server issue), W358–W362 (2005)
37. Flores, S., Echols, N., Milburn, D., Hespenheide, B., Keating, K., Lu, J., Wells, S., Yu, E.Z., Thorpe, M., Gerstein, M.: The Database of Macromolecular Motions: New features added at the decade mark. Nucleic Acids Res. **34**(Database issue), D296–D301 (2006)
38. Lindahl, E., Azuara, C., Koehl, P., Delarue, M.: NOMAD-Ref: Visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. Nucleic Acids Res. **34**(Web Server issue), W52–W56 (2006)

39. Eyal, E., Yang, L.W., Bahar, I.: Anisotropic network model: Systematic evaluation and a new web interface. Bioinformatics **22**(21), 2619–2627 (2006)
40. Jossinet, F., Westhof, E.: Sequence to structure (S2S): Display, manipulate and interconnect RNA data from sequence to structure. Bioinformatics **21**(15), 3320–3321 (2005)
41. Pavlopoulos, G.A., Soldatos, T.G., Barbosa-Silva, A., Schneider, R.: A reference guide for tree analysis and visualization. BioData Min. **3**(1), 1 (2010)
42. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G.: Gene ontology: Tool for the unification of biology. The gene ontology consortium. Nat. Genet. **25**(1), 25–29 (2000)
43. Bodenreider, O.: The unified medical language system (UMLS): Integrating biomedical terminology. Nucleic Acids Res. **32**(Database issue), D267–D270 (2004)
44. Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B., Bork, P.: Toward automatic reconstruction of a highly resolved tree of life. Science **311**(5765), 1283–1287 (2006)
45. Pennisi, E.: Modernizing the tree of life. Science **300**(5626), 1692–1697 (2003)
46. Schmidt, M.T., Handschuh, L., Zyprych, J., Szabelska, A., Olejnik-Schmidt, A.K., Siatkowski, I., Figlerowicz, M.: Impact of DNA microarray data transformation on gene expression analysis—comparison of two normalization methods. Acta Biochim. Pol. **58**(4), 573–580 (2011)
47. Geller, S.C., Gregg, J.P., Hagerman, P., Rocke, D.M.: Transformation and normalization of oligonucleotide microarray data. Bioinformatics **19**(14), 1817–1823 (2003)
48. Quackenbush, J.: Microarray data normalization and transformation. Nat. Genet. **32**(Suppl), 496–501 (2002)
49. Johnson, S.C.: Hierarchical clustering schemes. Psychometrika **32**(3), 241–254 (1967)
50. Sneath, P.H.A., Sokal, R.R.: Unweighted pair group method with arithmetic mean. Numer. Taxonomy, 230–234 (1973)
51. Gascuel, O., Steel, M.: Neighbor-joining revealed. Mol. Biol. Evol. **23**(11), 1997–2000 (2006)
52. Saitou, N., Nei, M.: The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**(4), 406–425 (1987)
53. Thompson, J.D., Higgins, D.G., Gibson, T.J.: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**(22), 4673–4680 (1994)
54. Edgar, R.C.: MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**(5), 1792–1797 (2004)
55. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J.: Basic local alignment search tool. J. Mol. Biol. **215**(3), 403–410 (1990)
56. Notredame, C., Higgins, D.G., Heringa, J.: T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. **302**(1), 205–217 (2000)
57. Cardona, G., Rossello, F., Valiente, G.: Extended Newick: It is time for a standard representation of phylogenetic networks. BMC Bioinf. **9**, 532 (2008)
58. Maddison, D.R., Swofford, D.L., Maddison, W.P.: NEXUS: An extensible file format for systematic information. Syst. Biol. **46**(4), 590–621 (1997)
59. Han, M.V., Zmasek, C.M.: phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinf. **10**, 356 (2009)
60. Procter, J.B., Thompson, J., Letunic, I., Creevey, C., Jossinet, F., Barton, G.J.: Visualization of multiple alignments, phylogenies and gene family evolution. Nat. Methods **7**(Suppl 3), S16–S25 (2010)
61. Schneiderman, B.: Tree visualization with tree-maps: 2-d space-filling approach. ACM Trans. Graph. **11**, 92–99 (1992)
62. Page, R.D.: Space, time, form: Viewing the tree of life. Trends Ecol. Evol. **27**(2), 113–120 (2012)

63. Zmasek, C.M., Eddy, S.R.: ATV: Display and manipulation of annotated phylogenetic trees. Bioinformatics **17**(4), 383–384 (2001)
64. Page, R.D.: TreeView: An application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12**(4), 357–358 (1996)
65. Goloboff, P.A., Catalano, S.A., Mirande, J.M., Szumik, C.A., Arias, J.S., Källersjö, M., Farris, J.S.: Phylogenetic analysis of 73060 taxa corroborates major eukaryotic groups. Cladistics **25**(3), 211–230 (2009)
66. Heer, J., Card, S.K.: DOITrees revisited: Scalable, space-constrained visualization of hierarchical data. In: AVI '04 Proceedings of the Working Conference on Advanced Visual Interfaces (2004)
67. Parr, C.S., Lee, B., Campbell, D., Bederson, B.B.: Visualizations for taxonomic and phylogenetic trees. Bioinformatics **20**(17), 2997–3004 (2004)
68. Michael, J.M., Gord, D., Ravin, B.: Expand-ahead: A space-filling strategy for browsing trees. In: INFOVIS '04 Proceedings of the IEEE Symposium on Information Visualization 2004, pp. 119–126. (2004)
69. Bingham, J., Sudarsanam, S.: Visualizing large hierarchical clusters in hyperbolic space. Bioinformatics **16**(7), 660–661 (2000)
70. Hughes, T., Hyun, Y., Liberles, D.A.: Visualising very large phylogenetic trees in three dimensional hyperbolic space. BMC Bioinf. **5**, 48 (2004)
71. Letunic, I., Bork, P.: Interactive tree of life v2: Online annotation and display of phylogenetic trees made easy. Nucleic Acids Res. **39**(Web Server issue), W475–W478 (2011)
72. Chevenet, F., Brun, C., Banuls, A.L., Jacq, B., Christen, R.: TreeDyn: Towards dynamic graphics and annotations for analyses of trees. BMC Bioinf. **7**, 439 (2006)
73. Pethica, R., Barker, G., Kovacs, T., Gough, J.: TreeVector: Scalable, interactive, phylogenetic trees for the web. PLoS ONE **5**(1), e8934 (2010)
74. Huson, D.H., Richter, D.C., Rausch, C., Dezulian, T., Franz, M., Rupp, R.: Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinf. **8**, 460 (2007)
75. Sanderson, M.J.: Paloverde: An OpenGL 3D phylogeny browser. Bioinformatics **22**(8), 1004–1006 (2006)
76. Kidd, D.M., Ritchie, M.G.: Phylogeographic information systems: Putting the geography into phylogeography. J. Biogeogr. **33**(11), 1851–1865 (2006)
77. Kidd, D.M.: Geophylogenies and the map of life. Syst. Biol. **59**(6), 741–752 (2010)
78. Buchin, K., Buchin, M., Byrka, J., Nöllenburg, M., Okamoto et al. : Drawing (complete) binary tanglegrams: Hardness, approximation, fixed-parameter tractability. 16th international symposium on graph drawing, Heraklion,Lecture Notes in Computer Science 5417 (2008)
79. Puigbo, P., Wolf, Y.I., Koonin, E.V.: Search for a 'Tree of Life' in the thicket of the phylogenetic forest. J. Biol. **8**(6), 59 (2009)
80. Tanner, C.C., Migdal, C.J., Jones, M.T.: The clipmap: A virtual mipmap. In: SIGGRAPH '98 Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (1998)
81. Collins, F.S., et al.: Finishing the euchromatic sequence of the human genome. Nature **431**(7011), 931–945 (2004)
82. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W.,

McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J.: Initial sequencing and analysis of the human genome. Nature **409**(6822), 860–921 (2001)

83. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G.,

Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigo, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., Zhu, X.: The sequence of the human genome. Science **291**(5507), 1304–1351 (2001)

84. Bennett, S.: Solexa Ltd. Pharmacogenomics **5**(4), 433–438 (2004)
85. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M.: Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**(7057), 376–380 (2005)
86. Barabási, A.-L., Gulbahce, N., Loscalzo, J.: Network medicine: A network-based approach to human disease. Nat. Rev. Genet. **12**, 56–68 (2011)
87. Xie, C., Tammi, M.T.: CNV-seq, a new method to detect copy number variation using high-throughput sequencing. BMC Bioinf. **10**, 80 (2009)
88. Keravala, A., Lee, S., Thyagarajan, B., Olivares, E.C., Gabrovsky, V.E., Woodard, L.E., Calos, M.P.: Mutational derivatives of PhiC31 integrase with increased efficiency and specificity. Mol. Ther. **17**(1), 112–120 (2009)
89. Chiang, D.Y., Getz, G., Jaffe, D.B., O'Kelly, M.J., Zhao, X., Carter, S.L., Russ, C., Nusbaum, C., Meyerson, M., Lander, E.S.: High-resolution mapping of copy-number alterations with massively parallel sequencing. Nat. Methods **6**(1), 99–103 (2009)
90. Kim, T.M., Luquette, L.J., Xi, R., Park, P.J.: rSW-seq: Algorithm for detection of copy number alterations in deep sequencing data. BMC Bioinf. **11**, 432 (2010)
91. Wang, Z., Gerstein, M., Snyder, M.: RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet. **10**(1), 57–63 (2009)
92. Morin, R., Bainbridge, M., Fejes, A., Hirst, M., Krzywinski, M., Pugh, T., McDonald, H., Varhol, R., Jones, S., Marra, M.: Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques **45**(1), 81–94 (2008)
93. Johnson, D.S., Mortazavi, A., Myers, R.M., Wold, B.: Genome-wide mapping of in vivo protein-DNA interactions. Science **316**(5830), 1497–1502 (2007)
94. Durbin, R.M., Abecasis, G.R., Altshuler, D.L., Auton, A., Brooks, L.D., Gibbs, R.A., Hurles, M.E., McVean, G.A.: A map of human genome variation from population-scale sequencing. Nature **467**(7319), 1061–1073 (2010)
95. Buchanan, C.C., Torstenson, E.S., Bush, W.S., Ritchie, M.D.: A comparison of cataloged variation between International HapMap consortium and 1000 genomes project data. J. Am. Med. Inform. Assoc. **19**(2), 289–294 (2012)
96. Tanaka, T.: International HapMap project. Nihon Rinsho **63**(Suppl 12), 29–34 (2005)
97. Thorisson, G.A., Smith, A.V., Krishnan, L., Stein, L.D.: The international HapMap project web site. Genome Res. **15**(11), 1592–1593 (2005)

98. Integrating ethics and science in the international HapMap project. Nat. Rev. Genet. **5**(6), 467–475 (2004)
99. The international HapMap project. Nature **426**(6968), 789–796 (2003)
100. Stalker, J., Gibbins, B., Meidl, P., Smith, J., Spooner, W., Hotz, H.R., Cox, A.V.: The Ensembl web site: Mechanics of a genome browser. Genome Res. **14**(5), 951–955 (2004)
101. Birney, E., Bateman, A., Clamp, M.E., Hubbard, T.J.: Mining the draft human genome. Nature **409**(6822), 827–828 (2001)
102. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D.: The human genome browser at UCSC. Genome Res. **12**(6), 996–1006 (2002)
103. Integrative Genomics Viewer: http://www.broadinstitute.org/software/igv/
104. Pavlopoulos, G.A., Secrier, M., Moschopoulos, C.N., Soldatos, T.G., Kossida, S., Aerts, J., Schneider, R., Bagos, P.G.: Using graph theory to analyze biological networks. BioData Min. **4**, 10 (2011)
105. Vikis, H.G., Guan, K.L.: Glutathione-S-transferase-fusion based assays for studying protein–protein interactions. Methods Mol. Biol. **261**, 175–186 (2004)
106. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., Seraphin, B.: The tandem affinity purification (TAP) method: A general procedure of protein complex purification. Methods **24**(3), 218–229 (2001)
107. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., Sakaki, Y.: A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. U S A **98**(8), 4569–4574 (2001)
108. Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga, G.: Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature **415**(6868), 141–147 (2002)
109. Stoll, D., Templin, M.F., Bachmann, J., Joos, T.O.: Protein microarrays: Applications and future challenges. Curr. Opin. Drug Discov. Devel. **8**(2), 239–252 (2005)
110. Willats, W.G.: Phage display: Practicalities and prospects. Plant Mol. Biol. **50**(6), 837–854 (2002)
111. Bader, G.D., Hogue, C.W.: An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinf. **4**, 2 (2003)
112. Pavlopoulos, G.A., Moschopoulos, C.N., Hooper, S.D., Schneider, R., Kossida, S.: jClust: A clustering and visualization toolbox. Bioinformatics **25**(15), 1994–1996 (2009)
113. Spirin, V., Mirny, L.A.: Protein complexes and functional modules in molecular networks. Proc. Natl. Acad. Sci. U S A **100**(21), 12123–12128 (2003)
114. Li, X.L., Tan, S.H., Foo, C.S., Ng, S.K.: Interaction graph mining for protein complexes using local clique merging. Genome Inf. **16**(2), 260–269 (2005)
115. Altaf-Ul-Amin, M., Shinbo, Y., Mihara, K., Kurokawa, K., Kanaya, S.: Development and implementation of an algorithm for detection of protein complexes in large interaction networks. BMC Bioinf. **7**, 207 (2006)
116. Liu, G., Wong, L., Chua, H.N.: Complex discovery from weighted PPI networks. Bioinformatics **25**(15), 1891–1897 (2009)
117. Mete, M., Tang, F., Xu, X., Yuruk, N.: A structural approach for finding functional modules from large biological networks. BMC Bioinf. **9**(Suppl 9), S19 (2008)
118. Adamcsek, B., Palla, G., Farkas, I.J., Derenyi, I., Vicsek, T.: CFinder: Locating cliques and overlapping modules in biological networks. Bioinformatics **22**(8), 1021–1023 (2006)
119. Moschopoulos, C.N., Pavlopoulos, G.A., Schneider, R., Likothanassis, S.D., Kossida, S.: GIBA: A clustering tool for detecting protein complexes. BMC Bioinf. **10**(Suppl 6), S11 (2009)
120. Chua, H.N., Ning, K., Sung, W.K., Leong, H.W., Wong, L.: Using indirect protein–protein interactions for protein complex prediction. J. Bioinf. Comput. Biol. **6**(3), 435–466 (2008)

121. Li, X.L., Foo, C.S., Ng, S.K.: Discovering protein complexes in dense reliable neighborhoods of protein interaction networks. Comput. Syst. Bioinf. Conf. **6**, 157–168 (2007)
122. Lubovac, Z., Gamalielsson, J., Olsson, B.: Combining functional and topological properties to identify core modules in protein interaction networks. Proteins **64**(4), 948–959 (2006)
123. Cho, Y.R., Hwang, W., Ramanathan, M., Zhang, A.: Semantic integration to identify overlapping functional modules in protein interaction networks. BMC Bioinf. **8**, 265 (2007)
124. Maraziotis, I.A., Dimitrakopoulou, K., Bezerianos, A.: Growing functional modules from a seed protein via integration of protein interaction and gene expression data. BMC Bioinf. **8**, 408 (2007)
125. Feng, J., Jiang, R., Jiang, T.: A max-flow-based approach to the identification of protein complexes using protein interaction and microarray data. IEEE/ACM Trans. Comput. Biol. Bioinf. **8**(3), 621–634 (2011)
126. Ulitsky, I., Shamir, R.: Identification of functional modules using network topology and high-throughput data. BMC Syst. Biol. **1**, 8 (2007)
127. Li, X., Wu, M., Kwoh, C.K., Ng, S.K.: Computational approaches for detecting protein complexes from protein interaction networks: A survey. BMC Genomics **11**(Suppl 1), S3 (2010)
128. Pavlopoulos, G.A., Wegener, A.L., Schneider, R.: A survey of visualization tools for biological network analysis. BioData Min. **1**, 12 (2008)
129. Gehlenborg, N., O'Donoghue, S.I., Baliga, N.S., Goesmann, A., Hibbs, M.A., Kitano, H., Kohlbacher, O., Neuweger, H., Schneider, R., Tenenbaum, D., Gavin, A.C.: Visualization of omics data for systems biology. Nat. Methods **7**(Suppl 3), S56–S68 (2010)
130. Suderman, M., Hallett, M.: Tools for visually exploring biological networks. Bioinformatics **23**(20), 2651–2659 (2007)
131. Kohler, J., Baumbach, J., Taubert, J., Specht, M., Skusa, A., Ruegg, A., Rawlings, C., Verrier, P., Philippi, S.: Graph-based analysis and visualization of experimental results with ONDEX. Bioinformatics **22**(11), 1383–1390 (2006)
132. Breitkreutz. BJ, Stark. C, M, T.: Pajek—program for large network analysis. Connections **21**, 47–57 (1998)
133. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T.: Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res. **13**(11), 2498–2504 (2003)
134. Pavlopoulos, G.A., Hooper, S.D., Sifrim, A., Schneider, R., Aerts, J.: Medusa: A tool for exploring and clustering biological networks. BMC Res. Notes **4**(1), 384 (2011)
135. Hu, Z., Hung, J.H., Wang, Y., Chang, Y.C., Huang, C.L., Huyck, M., DeLisi, C.: VisANT 3.5: Multi-scale network visualization, analysis and inference based on the gene ontology. Nucleic Acids Res. **37**(Web Server issue), W115–W121 (2009)
136. Letunic, I., Yamada, T., Kanehisa, M., Bork, P.: iPath: Interactive exploration of biochemical pathways and networks. Trends Biochem. Sci. **33**(3), 101–103 (2008)
137. Dogrusoz, U., Erson, E.Z., Giral, E., Demir, E., Babur, O., Cetintas, A., Colak, R.: PATIKAweb: A web interface for analyzing biological pathways through advanced querying and visualization. Bioinformatics **22**(3), 374–375 (2006)
138. van Iersel, M.P., Kelder, T., Pico, A.R., Hanspers, K., Coort, S., Conklin, B.R., Evelo, C.: Presenting and exploring biological pathways with PathVisio. BMC Bioinf. **9**, 399 (2008)
139. Shamir, R., Maron-Katz, A., Tanay, A., Linhart, C., Steinfeld, I., Sharan, R., Shiloh, Y., Elkon, R.: EXPANDER–an integrative program suite for microarray data analysis. BMC Bioinf. **6**, 232 (2005)
140. Seo, J., Gordish-Dressman, H., Hoffman, E.P.: An interactive power analysis tool for microarray hypothesis testing and generation. Bioinformatics **22**(7), 808–814 (2006)
141. Kapushesky, M., Kemmeren, P., Culhane, A.C., Durinck, S., Ihmels, J., Korner, C., Kull, M., Torrente, A., Sarkans, U., Vilo, J., Brazma, A.: Expression Profiler: Next generation–an online platform for analysis of microarray data. Nucleic Acids Res. **32**(Web Server issue), W465–W470 (2004)

142. Secrier, M., Pavlopoulos, G.A., Aerts, J., Schneider, R.: Arena3D: Visualizing time-driven phenotypic differences in biological systems. BMC Bioinf. **13**(1), 45 (2012)
143. Pavlopoulos, G.A., O'Donoghue, S.I., Satagopam, V.P., Soldatos, T.G., Pafilis, E., Schneider, R.: Arena3D: Visualization of biological networks in 3D. BMC Syst. Biol. **2**, 104 (2008)
144. Freeman, T.C., Goldovsky, L., Brosch, M., van Dongen, S., Maziere, P., Grocock, R.J., Freilich, S., Thornton, J., Enright, A.J.: Construction, visualisation, and clustering of transcription networks from microarray expression data. PLoS Comput. Biol. **3**(10), 2032–2042 (2007)
145. Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D., Ferrin, T.E.: clusterMaker: A multi-algorithm clustering plugin for Cytoscape. BMC Bioinf. **12**, 436 (2011)
146. Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M.A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A.M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B., Superti-Furga, G.: Proteome survey reveals modularity of the yeast cell machinery. Nature **440**(7084), 631–636 (2006)
147. Junker, B.H., Koschutzki, D., Schreiber, F.: Exploration of biological network centralities with CentiBiN. BMC Bioinf. **7**, 219 (2006)
148. O'Donoghue, S.I., Gavin, A.C., Gehlenborg, N., Goodsell, D.S., Heriche, J.K., Nielsen, C.B., North, C., Olson, A.J., Procter, J.B., Shattuck, D.W., Walter, T., Wong, B.: Visualizing biological data-now and in the future. Nat. Methods **7**(Suppl 3), S2–S4 (2010)

Personalisation Systems for Cultural Tourism

Katerina Kabassi

Abstract This paper makes a review on personalisation systems specialised for cultural tourism. Tourists interested in cultural heritage have different requirements from tourist recommendation systems than other users. Therefore, emphasis is given on recommendation systems for city tours and museum guides. More specifically, systems for PC, PDAs and mobile phones are discussed as well as the methods and the technologies used.

1 Introduction

Nowadays people do not travel only for resting on a beach and enjoy the sun but to combine rest with their interests in culture, religion or the environment. This result in different kinds of tourism: cultural tourism, religious tourism or ecotourism. The tourists with such interests use the Information and Communication Technologies (ICTs) for searching information about their destination or taking information on site. Indeed, ICTs enable tourists to access reliable and accurate information as well as to undertake reservations and plans in a fraction of time, cost and inconvenience that may be required by conventional methods [37].

These services were further influenced by the Internet and related technologies. However, ICTs and the Internet have increased the number of choices so dramatically that is very difficult for the consumers to find what they are looking for. An effective solution for reducing complexity when searching information over the Internet has been given by recommendation systems [1]. Recommendation systems have been used for finding books [34], movies [56], tv-programs [61], music [53], etc. The main characteristic of the recommender systems is that they can

K. Kabassi (✉)
Department of Protection and Conservation of Cultural Heritage,
TEI of Ionian Islands, 2 Kalvou Square 29100 Zakynthos, Greece
e-mail: kkabassi@teiion.gr

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*,
Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_7,

personalize their interaction to each individual user. Personalization involves the design of enabling systems to capture or infer the needs of each person and then to satisfy those needs in a known context [45].

Personalized recommendation systems have been gaining interest in tourism to assist users during their city tours [29, 47] or museum tours [46, 59]. The users that make a city tour or a museum tour have different interests and needs. A remedy for the negative effects of the traditional 'one-size-fits-all' approach is to develop systems with an ability to adapt their behavior to the goals, tasks, interests and other features of individual users and groups of users [55].

Therefore, these recommendation systems use information about the user to personalize the interaction with each individual user. A personalization system is based on three main functionalities: content selection, user model adaptation and presentation of results [14, 15, 36]. By content selection, one may refer to selecting destination, tourist attractions, museum "artifacts" or all the above for planning a whole trip. By user model adaptation, one may refer to techniques used for maintaining updated user models. Finally the presentation of results involves the technologies used (e.g. multimedia, GIS etc.) for improving the interactivity of the systems and, therefore, human–computer interaction.

In view of the above, different approaches have been proposed for helping the user during his/her cultural pursuit. Ellis, Patten and Evans [16] explore a variety of more or less social museum media, and point to the continuing need to "target personalised offerings at specific users." Such media involve guidance systems for mobile phones or PDAs, multimedia, etc. More innovative approaches include robots that guide users through museums [54]. However, these are not appropriate for individual use and are difficult to adapt to different environments.

2 Personalising City Tours

New technologies are used for organizing different aspects of a trip, e.g. selecting destination, accommodations, restaurants, routes or all the above for planning a whole trip. Among others, new technologies and the Internet are used for selecting the tourist attractions and sights that the tourists are planning to visit, if they are interested on the cultural heritage of the area. Indeed, many researchers support that tourist attractions are often the reason driving travelers to visit destinations [21, 27, 33, 44]. In view of this, Traveller [52] takes into account the touristic sites that may be of interest of the particular user to suggest package holidays and tours.

Interesting work is also that of [25], who have developed a recommendation system for suggesting specific tourist attractions over the Internet. The proposed system combine a multi-criteria decision making theory, the Analytic Hierarchy Process (AHP), with a Bayesian network for finding over the Internet which is the tourist attraction that would interest the user interacting with the system.

On a different basis, AVANTI [18, 19] personalizes the presentation of information about specific touristic sites. In this case, the user is not only proposed with touristic sites that may interests him/her but the information provided about each sight is adapted to his/her interests and knowledge. These systems usually help the user by personalizing interaction with the personal computer while surfing on the Internet to locate information about the cultural heritage of a city or a country. Indeed, many users search on the Internet about cultural sights of a city or a country that may visit prior to their visit. Other systems such as the INTRIGUE guide recommends sightseeing destinations by taking into account the preferences of heterogeneous tourist groups [3].

Table 1 refers to some systems that personalizing information about sights and attraction.

However, another way to find out information about the cultural sights of a city or a country is to have mobile Internet, either with a palmtop or a mobile phone and search information about it while you are on sight. Such systems are Speta [20], PinPoint [47], m-ToGuide prototype [29] and UMT [58].

Additionally to promoting touristic sights' CRUMPET [40] uses advertisements to promote shops, restaurants, entertainment places, events as well as information, reservation, booking and payment services that may be helpful to any tourist. [39] (CATIS) and [11] propose systems that take into account the physical location of the user to provide a set of request-related services in the surrounding area.

Special requirements are also addressed by systems that are designed to assist specific types of tourists during their tours. Such systems that provide personalized information on specific cities and their cultural sights, Lancaster, Heidelberg,

Table 1 Some systems that personalize information about attractions and sights

Attractions—Sights
• Huang and Bian [25]
• Gunn [21]
• Lew [33]
• Jafari [27]
• Richards [44]
• *WebGuide* [17, 63]
• *MastroCARonte* [12]
• *CRUMPET* [40]
• *Traveller* [52]
• *Travel Planner* [10]
• *AVANTI* [18, 19]
• *INTRIGUE* [3]
• Hinze and Voisard [22]
• *Speta* [20]
• *Gulliver's Genie* [23, 38]
• *MobiDENK* [32, 6]
• *PinPoint* [44]
• *m-ToGuide prototype* [29]
• *UMT* [58]
• *ITR* [41]

Oldenburg and Vienna, are the GUIDE system [9], WebGuide [17, 62], Sightseeing 4U [51] and LoL@ [2], respectively.

In order to personalize interaction, these systems use specific criteria for evaluating the different alternatives. The criteria used for evaluating the packages and tours are summarized in Table 2.

3 Personalising Museum Tours

Roes et al. [46] have identified four types of museum tours: human-guided tours, audio tours, online/virtual tours, and multimedia tours. Several museums, e.g. Tate Modern, Science Museum Boston, already explored the potential of bridging the Web and the physical museum spaces. Indeed, several technologies such as multimedia, mobile and web technologies have been used for this purpose. However, the main problem with such approaches is that a human tour may be more interesting as it is live and it can be adapted to the audience. A solution to this may be given with the incorporation of personalization services in museum guides in order to enhance the tourist experience in the museum. A context-aware system for intelligent museum collects information of visitors and surroundings, recognizes visitors' purposes, and then assists visiting, while striving to be minimally intrusive through this process [49]. A visitor may enter the system by any device, desktop computer or mobile devices.

In view of this, [59] propose a context-aware intelligent museum system, namely iMuseum, that provides visitors with customized relic context usage through an underlying context server. IMuseum uses interests, to adapt the context presented to users. The user interests are also taken into account in the approach of [48], which personalizes user interaction in a semantically annotated museum collection. More characteristics of the user and not just his/her interests are taken into account in the Rijksmuseum project [5], which personalizes users' museum experiences within the virtual and physical collections.

However, the above mentioned approaches do not emphasize much in simulating the live experiences in a museum. The value of multimedia for a museum guide in order to simulate better the live experience is discussed by Proctor and Tellis (2003) who present an extended user study conducted at the Tate Modern in

Table 2 Criteria used for recommending tours

	Intrigue [3]	Traveller [52]
Destination	√	√
Duration	√	√
Season		
Price	√	√
Category/Type	√	√
Historical or artistic value	√	

2002. However, this study emphasizes just on mobile museum guides. Some projects that have taken place towards this direction include the Multimedia Tour [57] and the Interactive Museum Guide Bay et al. [7]. The latter is not addressed to mobile museum guides. More specifically, it uses a PC with a touch screen, a webcam and a bluetooth receiver. The guide recognizes objects in the museum based on images of particular artifacts and provides additional information on the subject.

A rather interesting and complete work on the subject of museum guides is the Cultural Heritage Information Personalization (CHIP) project, which demonstrates how Semantic Web technologies can be deployed to provide personalized access to digital museum collections. More specifically, CHIP personalizes the selection of artworks for the museum visitor based on their underlying semantic relations, e.g. related styles, artists, themes, or locations and the strength of the user interest in those semantically enriched properties [46].

4 Technology

The systems used for supporting cultural tourism use different methods of personalisation or intelligence to become useful and, therefore, attract users. To this direction, many systems are developed for mobile phones or PDAs. However, the problems addressed in such technology are quite different due to the limited space in the screen. For this purpose, the MoMo project [26] proposes a mechanism for browsing large collections of explanatory items on PDAs. This project aims at providing users with social interaction within museums, but has not achieved to make services intelligent enough. Cheverst, Davies and Mitchell [9], on the other hand, propose personalized and, therefore, in a way intelligent, tours for PDA users during his/her physical visit in the sight of interest. A quite different approach is used in the Exploratorium [24] and Peabody Essex Museum's ART scape [28] allows a visitor to bookmark an exhibit during the physical visit and then later search related information about this exhibit from the website.

Another way to attract users by using different technologies are incorporating multimedia into the systems. Such projects include 3D virtual reality representations of galleries and other geographical areas of cultural interest [60]. A more innovative approach is proposed by [31] who personalize content into a tourist's mobile in a Multimedia Messaging Service (MMS). Additionally, to the profile of the user, this approach also takes into account the physical location of the user. For this purpose, in other systems, researchers use agents to monitor the transportation of cultural assets. In one scenario, users visit Villa Adriana, an archaeological site in Tivoli, Italy and the agents discover users' movements via a Galileo satellite signal [13]. The agents elicit users' habits and preferences and personalize interaction according to the information that has been extracted implicitly.

5 User Modeling

For a system to be able to provide personalized recommendations it should make inferences about the users' preferences. Such information as well as information about the users' previous experiences is stored in a user model [20, 52]. Indeed, as Schafer et al. [50] point out, recommender systems offer guidance based on users' profiles or visiting background. Therefore, every recommender system builds and maintains a collection of user models [35].

A recommender system may maintain an individual user model or some user models that represent classes of users [42, 43]. When commercializing complex customizable products online, there may be various classes of users of the configurator that differ in properties such as skills, needs and knowledge level [4]. These classes are called stereotypes. Stereotypes [30, 42, 43] are used in user modeling in order to provide default assumptions about individual users belonging to the same category according to a generic classification of users that has previously taken place. This method has the advantage of providing personalized recommendations from the first interaction of the user with the system. However, a main disadvantage of this approach is that users may be similar in some characteristics but differentiate in many others. Furthermore, a user's characteristics may change over time. Some systems that use stereotypical techniques for personalizing the presentation of cultural sights to tourists are AVANTI [18, 19], INTRIGUE [3] and UMT [58]. However, a main problem that such systems encounter is that each user differentiates from all the others in many ways. Therefore, many systems use individual user modeling (e.g. Sightseeing4U [51], PinPoint [47], m-ToGuide prototype [25, 29]. Individual user modeling has many advantages it can not be used before the user has to interact with the system for a long time without any personalization so that the system collects adequate information. This disadvantage is addressed in many systems by using a combination of the two methods (e.g. WebGuide [17, 63], MastroCARonte [12], Traveller [52], Speta Garcia-Crespo et al. [20]).

Systems can also be categorized taking into account the way of information acquisition for the user model. Information about the user may be acquired explicitly or may be inferred implicitly from the user's previous interactions (e.g. WebGuide [17, 63], AVANTI [18, 19], INTRIGUE [3], Gulliver's Genie [23, 38], m-ToGuide prototype [29], UMT [58] or both (TravelPlanner [10], MastroCARonte [12], Speta [20], [25]). Enabling consumers to develop their online profile and to include personal data that indicate their reference can support tourism organizations to provide better service [8]. The main problem with explicit user models is that users may have to answer too many questions. Furthermore, users may not be able to describe themselves and their preferences accurately. In this respect, implicit user modeling has been considered as more reliable and non-intrusive than explicit user modeling. However, one main problem of this approach is that the hypotheses generated by the system for each user may not be accurate. Furthermore, there may not be sufficient time for the system to observe the user for producing accurate hypotheses about him/her. In view of the above advantages and disadvantages, some systems, such as TravelPlanner [10], MastroCARonte [12], Speta [20, 25], use a combination of explicit and implicit user modeling. More specifically, TravelPlanner selects the most useful queries to present to the user and all the other information is acquired implicitly. Similarly, in SPETA [20], in the beginning the user provides explicitly information about his/her interest, the kind of places s/he prefer to visit, and the ratings given to attractions. Additionally, a huge amount of information can be extracted from the social networks they belong to and the user behavior.

Finally, the last dimension that is taken into account user modeling systems involves short-term versus long-term user models. Almost all recommender systems maintain long-term user models as the previous interactions of the particular user or other users with similar characteristics are essential for content-based, collaborative or demographic filtering.

6 Discussion

There are many tourists that do not prefer to travel in groups with guides as they choose to have their own pace in a museum or a city. For this purpose, many tourists use guide books or other systems. More specifically, different kinds of hardware equipment using specialised software systems or the Internet have been used into the museum environment such as PDA's, mobile phones, tablet PC's or even robots. However, a main problem with the above means is that the books have limited information while Internet has so many links to visit that a user may be frustrated. Software systems for a city or a museum, on the other hand, they are usually static and difficult to use. These are the main problems addressed by adaptive recommender systems for museums or city tours.

The main focus of the paper is to make a state-of-the-art on the adaptive systems for tourists interested in cultural heritage. More specifically, emphasis is given on the adaptive systems as well as the hardware technology used. In order to make their interaction adaptive, system use user modeling techniques for capturing and using the users' interests and background knowledge. The complexity of such systems makes it difficult and time consuming to implement. Therefore, in this paper we refer to the problems of the systems implemented and the special requirements of each hardware device used for this purpose.

References

1. Adomavicius, G., Tuzhilin, A.: Towards the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)
2. Anegg, H., Kunczier, H., Michlmayr, E., Pospischil, G., Umlauft, M.: LoL@: Designing a location based UMTS application. ÖVE-Verbandszeitschrift e&i, Springer, Heidelberg (2002)

3. Ardissono, L., Goy, A., Petrone, G., Signan, M., Torasso, P.: Intrigue: personalized recommendation of tourism attractions for desktop and handset devices. Appl. Artif. Intell. **17**(8–9), 687–714 (2003)
4. Ardissono, L., Felfernig, A., Friedrich, G., Jannach, D., Schafer, R., Zanker, M.: Intelligent interfaces for distributed web-based product and service configuration. In: Zhong, M., et al. (eds.) Web Intelligence 2001, LNAI 2198, pp. 184–188. Springer, Heidelberg (2001)
5. Aroyo, L., Brussee, R., Rutledge, L., Gorgels, P., Stash, N., Wang, Y: Personalized museum experience: the rijksmuseum use case. In: Trant, J., Bearman, D. (eds). Museums and the Web 2007: Proceedings, Toronto
6. Baldzer, J., Boll, S., Klante, P., Krösche, J., Meyer, J., Rump, N., Scherp, A., Appelrath, H.: Location-aware mobile multimedia applications on the niccimon platform. In: Braunschweiger Symposium—Informationssysteme für mobile Anwendungen (2004)
7. Bay, H., Fasel, B., Van Gool, L., Interactive museum guide: fast and robust recognition of museum objects. In: Proceedings of the International Workshop on Mobile Vision. (2006)
8. Buhalis, D., Law, R.: Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. Tourism Manage. **29**, 609–623 (2008)
9. Cheverst, K., Mitchell, K., Davies, N.: The role of adaptive hypermedia in a context-aware tourist guide. Commun. ACM **45**(5), 47–51 (2002)
10. Chin D, Porage A.: Acquiring user preferences for product customization, in: Bauer, M., Gmytrasiewicz, P., Vassileva, J. (eds.) Proceedings of the 8th International Conference on User Modeling 2001; LNAI 2109, 95–104 (2001)
11. Choi, D.-Y.: Personalized local internet in the location-based mobile web search. Decis. Support Syst. **43**, 31–45 (2007)
12. Console, L., Lombardi, I., Gioria, S.: Personalized and adaptive services on board a car: an application for tourist information. J. Intell. Inf. Syst. **21**(3), 249–284 (2003)
13. Costantini, S., Mostarda, L., Tocchio, A., Tsintza, P.: DALICA: agent-based ambient intelligence for cultural-heritage scenarios. IEEE Intell. Syst. **23**(2), 34–41 (2008)
14. Diaz, A., Gervas, P.: Personalization in news delivery systems: Item summarization and multi-tier item selection using relevance feedback. Web Intell. Agent Syst. **3**(3), 135–154 (2005)
15. Diaz, A., Garcia, A., Gervas, P.: User-centered versus system-centered evaluation of a personalization system. Inf. Process. Manage. **44**, 1293–1307 (2008)
16. Ellis, M., Patten, D., Evans, D.: Getting the most out of our users, or, the science museum lab: how the dana centre lets us play. In: Trant, J., Bearman, D. (eds.) Museums and the Web 2005: Proceedings (2005)
17. Fink, J., Kobsa, A.: User modeling for personalized city tours. Artif. Intell. Rev. **18**, 33–74 (2002)
18. Fink, J., Kobsa, A., Nill, A.: User-oriented adaptivity and adaptability in the AVANTI project. Designing for the Web: Empirical Studies. Redmond, WA, 1996
19. Fink, J., Kobsa, A., Nill, A.: Adaptable and adaptive information provision for all users, including disabled and elderly people. New Rev. Hypermed. Multimed. **4**, 163–188 (1998)
20. Garcia Cerspo, A., Chamizo, J., Rivera, I., Mncke, M., Colomo Palaciosm, R., Gomez Berbis JM.: SPETA: Social pervasive e-Tourism advisor. Telematics and Informatics 2009; to appear
21. Gunn, C.: Amendment to leiper: the framework of tourism. Ann. Tourism Res. **7**, 253–255 (1980)
22. Hinze, A., Voisard, A.: Location and time-based information delivery in tourism. In: Proceedings of the 8th International Symposium on Advances in Spatial and Temporal Databases 2003
23. Hristova N, O'Hare GMP, Lowen T. Agent-based ubiquitous systems: 9 lessons learnt. In: Workshop on System Support for Ubiquitous Computing (UbiSys'03), 5th International Conference on Ubiquitous Computing (UbiComp), Seattle 2003

24. Hsi, S., Fait, H.: RFID enhances visitors' museum experience at the Exploratorium. Commun. ACM **48**(9), 60–65 (2005)
25. Huang, Y., Bian, L.: A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet. Expert Syst. Appl. **36**(1), 933–943 (2009)
26. Jaen, J., Esteve, J.M., Mocholi, J.A., et al.: MoMo: enabling hybrid museums. IEE Proc. Softw. **152**(5), 245–251 (2005)
27. Jafari, J. (Ed.) Encyclopedia of tourism. Routledge; 2000
28. Johnson, B.: (2004). Beyond on-line collections: putting objects to work. Museums and the Web 2004
29. Kamar, A.: Mobile tourist guide (m-ToGuide). Deliverable 1.4, Project Final Report. IST-2001-36004 (2003)
30. Kobsa, A., Koenemann, J., Pohl, W.: Personalized hypermedia presentation techniques for improving on-line customer relationships. Knowl. Eng. Rev. **16**, 111–115 (2001)
31. Koch, F., Sonenberg, L.: Using multimedia content in intelligent mobile services, Proceedings of the WebMedia and LA-Web 2004 Joint Conference 10th Brazilian Symposium on Multimedia and the Web 2nd Latin American Web Congress (LA-Webmedia'04), 2004, pp. 41–43
32. Krösche, J., Baldzer, J., Boll, S.: MobiDENK-Mobile multimedia in monument conservation. IEEE MultiMed, **11**(2), 72–77 (2004)
33. Lew, A.: A framework of tourist attraction research. Ann. Tourism Res. **14**, 533–575 (1987)
34. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput. **7**(1), 76–80 (2003)
35. Montaner, M., Lopez, B., de la Rosa, J.L.: A taxonomy of recommender agents on the internet. Artif. Intell. Rev. **19**(4), 285–330 (2003)
36. Mirazzo, S., Tasso, C.: Ephemeral and persistent personalization in adaptive information access to scholarly publications on the Web. Adaptive hypermedia and adaptive web based systems. In: Proceedings of AH 2002, 2002; 2347, 302–316
37. O'Connor, P.: Electronic information distribution in tourism and hospitality. CAB, Wallingford (1999)
38. O'Grady, MJ., O'Hare, GMP.: Gulliver's Genie: agency, mobility and adaptivity. computers & graphics, special issue on pervasive computing and ambient intelligence—mobility, ubiquity and wearable's get together 2004; 28, 4, Elsevier
39. Pashtan, A., Blattler, R., Heusser, A., Scheuermann, P.: CATIS: a context-aware tourist information system. In: Proceedings of IMC 2003, 4th International Workshop of Mobile Computing, Rostock (2003)
40. Posland, S., Laamanen, H., Malaka, R., Nick, A., Buckle, P., Zipf, A.: Creation of user-friendly mobile tourism services personalized for tourism. In: Proceedings of the 2nd International Conference on 3G Mobile Communication Technologies, London, pp. 28–32 (2001)
41. Ricci, F., Venturini, A., Cavada, D., Mirzadeh, N., Blaas, D., Nones, M.: Produce recommendation with interactive query management and twofold similarity. In: Ashley, K.D., Bridge, D.G. (eds.) Proceedings of the 5th international conference on case-based reasoning, ICCBR 2003, LNCS vol. 2689, pp. 479–493 Springer, Heidelberg (2003)
42. Rich, E.: Stereotypes and user modeling. In: Kobsa, A., Wahlster, W., (eds.) User Models in Dialog Systems, p. 199–214 (1989)
43. Rich, E.: Users are individuals: individualizing user models. Int. J. Hum. Comput. Stud. **51**, 323–338 (1999)
44. Richards, G.: Tourism attraction systems-exploring cultural behavior. Ann. Tourism Res. **29**(4), 1048–1064 (2002)
45. Riecken, D.: Personalized views of personalization. Commun. ACM **43**(8), 27–28 (2000)
46. Roes I., Stash N., Wang Y., Aroyo L. A personalized walk through the museum: the chip interactive tour guide. In: CHI 2009—Proceedings of the 27th International Conference CHI 2009, pp. 3317–3322 (2009)

47. Roth, J.: Context-aware web applications using the pinpoint. IADIS International Conference WWW/Internet 2002, Lissabon (Portugal), IADIS Press 2002; pp. 3–10
48. Rutledge, L., Aroyo, L., Stash, N., Interactive user profiling in semantically annotated museum collections, In: Proceedings the 5th International Semantic Web Conference (ISWC'06), Athens (2006)
49. Satyanarayanan, M.: Interview on mobile and pervasive computing. IEEE Distrib. Syst. Online **2**(6): (2001)
50. Schafer, J.B., Konstan, J., Riedl, J.: Electronic commerce recommender applications. J. Data Min. Knowl. Discovery **5**(1–2), 115–152 (2000)
51. Scherp, A., Boll, S.: Generic support for personalized mobile multimedia tourist applications ACM Multimedia 2004. In: Proceedings of the 12th ACM International Conference on Multimedia, pp. 178–179 (2004)
52. Schiaffino, S., Amandi, A.: Building an expert travel agent as a software agent. Expert Syst. Appl. **36**, 1291–1299 (2009)
53. Sotiropoulos, D.N., Lampropoulos, A.S., Tsihrintzis, G.A.: MUSIPER: a system for individualization of music similarity perception based on objective feature subset selection. User Model. User-Adap. Inter. **18**, 315–348 (2008)
54. Thrun, S., Beetz, M., Bennewitz, M., Burgard, W., Cremers, A., Dellaert, F., Fox, D., Hahnel, D., Rosenberg, C., Roy, N., Schulte, J., Schulz, D.: Probabilistic algorithms and the interactive museum tour-guide robot minerva. Int. J. Robot. Res. **19**(11), 972–999 (2000)
55. Virvou, M.: Intelligence and adaptivity in human-computer interaction concerning biotechnological software users. In: Proceedings of the 5th International Workshop on Mathematical Methods in Scattering Theory and Biomedical Technology (2001)
56. Virvou, M., Savvopoulos, A.: User modelling for adaptive e-shopping. In: Proceedings of the 11th International Conference on Human Computer Interaction HCII'2005, July 22–27 2005, Las Vegas (2005)
57. Wilson, G., Multimedia tour programmer at Tate Modern. In: Proceedings of Museums and the Web (2004)
58. Yang, Y., Marques, N.C.: User group profile modeling based on user transactional data for personalized systems. Lect. Notes Comput. Sci. (LNCS **3808**, 337–347 (2005)
59. Yu, Z., Zhou, X., Yu, Z., Park, J.H., Ma, J.: iMuseum: a scalable context-aware intelligent museum system. Comput. Commun. **31**, 4376–4382 (2008)
60. Zára, J.: Virtual reality and cultural heritage on the web. In: Proceedings of the 7th International Conference on Computer Graphics and Artificial Intelligence, Limoges, France (2004)
61. Zimmerman, J., Kurapati, K., Buczak, A., Schafer, D., Gutta, S., Martino, J.: TV personalization system: design of a TV show recommender engine and interface. In: Personalized Digital Television: Targeting Programs to Individual Viewers. Kluwer Academic Publishers, Dordrecht, pp. 27–51 (2004)

Resource List

62. WebGuide 2001, WebGuide: A city guide for the internet. European media lab. http://www.eml.org/english/research/deepmap/deepgis/webguide.html
63. Museums and the web. The International Conference about museums and the Web http://www.museumsandtheweb.com
64. Tate modern multimedia tours for temporary exhibitions http://www.vitakapa.com/tatemoderntempshows.html and http://www.tate.org.uk/about/projects/tate-modern-multimedia-tour
65. Amazon http://www.amazon.com/
66. Rijksmuseum–Het museum van Nederland https://www.rijksmuseum.nl/

67. Ricci, F.; Rokach, L.; Shapira, B.; Kantor, P.B. (Eds.): Recommender Systems Handbook. Springer, Berlin (2011)
68. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender Systems: An Introduction. Cambridge University Press (2011)

Educational Recommender Systems: A Pedagogical-Focused Perspective

Salvador Garcia-Martinez and Abdelwahab Hamou-Lhadj

Abstract With the growing number of students in the classroom and the switch to online environments, instructors are beginning to integrate collaborative learning approaches in the classroom. However, many times in large collaborative environments and large social networks, students are overwhelmed by the amount of available information; it is often challenging to select the most appropriate sources of information. A promising way to deal with this challenge and enhance social interaction in collaborative learning environments is by introducing recommender systems. The main goal of this article is, through a literature review, explore the differences between general recommender systems and educational recommender systems, and to provide a general overview about the benefits, challenges and limitations of recommender systems in educational settings.

1 Introduction

Lecturing is one of the most used teaching styles in education. In this teaching style, usually the instructor stands up in front of the class and explains the topics using the blackboard or slides [1]. However, with the growing number of students and the introduction of online learning, instructors are becoming aware of the importance of supplementing or switching their teaching to emphasize collaborative learning, where social interaction plays a central role.

S. Garcia-Martinez (✉)
Department of Education, Concordia University, Montreal, Canada
e-mail: sa_garci@education.concordia.ca

A. Hamou-Lhadj
Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
e-mail: abdelw@ece.concordia.ca

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*, Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_8,

Collaborative learning is defined as “a situation in which two or more people learn or attempt to learn something together” [2]. This type of learning is frequently supported by computers and related technologies, which are often used to “facilitate, augment, and even redefine interaction among members of a group” [3]. When using web-based collaborative environments, students have virtual access to a wide variety of sources of information being able to communicate their ideas within and outside of the classroom [3]. In virtual environments, this social interaction is facilitated by social networks, which are defined as “a social structure of nodes that represent individuals (or organizations) and the relationships between them within a certain domain” [4].

On the importance of social interaction and collaborative learning, prominent theorists such as Berger and Luckmann argue that knowledge is maintained by social interactions, which are mostly based on social relationships. Additionally, they stress that human beings create society, construct social life, and that they are a social product [5]. Similarly, Vygosky concurs that it is possible to make meaning of knowledge within a social context, which is influenced by time, experience and culture Vygotsky and Cole [6]. In the same way, Cole states that in social learning environments, human cognition is distributed among the group of members [7].

However, in large collaborative environments and social networks, students are overwhelmed by all the available information and the quantity of users [8, 9]. The amount of content is hard to handle; and, information cannot be classified in just one category anymore [10, 11]. Students tend to be confused when selecting the most appropriate sources of information, especially when the number of choices increases [9, 12]. Students do not know exactly what they are looking for [13], and, they face difficulties when building teams or when peer-collaborating [8].

A promising way to deal with this problem and enhance social interaction in collaborative learning environments is by introducing *recommender systems*. The main goal of recommender systems is to “help users to deal with information overload by providing personalized recommendations, content and services” [14]. Recommendations are based on the user’s profile (i.e., a user model) and recommendation from other users (collaborative filtering). Recommender systems are mostly used in E-commerce for recommending music, movies, books, tv shows or different types of items [15]. Examples of commercial recommender systems are those used in online stores such as Amazon.com and Barnes and Noble [14].

The main goal of this article is to provide a general overview of the benefits recommender systems in educational settings. More precisely, we aim to answer the following questions:

- What are the main differences between general-purpose recommender systems and educational recommender systems?
- What are the benefits of introducing recommender systems in education?

At this end, we conducted search queries using keywords such as “recommender systems” or “recommendation systems” and “education” in the ED/ITLib, Engineering village, EBSCOHOST, Web of Science, IEEEXPlore, and ACM

Digital databases. We also retrieved articles from related journals and conferences (see *Appendix A* for a selected list of resources that were also considered for this review). In order to gather information about the cutting edge in the literature, the search ranged between the years 2000 and 2012. We found a total of 302 articles and four books. However, after deleting unrelated articles, we gathered a total of 98 relevant articles. From these articles, 24 focused on the development of the recommender algorithms, 24 were system design proposals, and 50 were focused on user's satisfactions, a project in development, usability studies, theoretical articles or literature reviews.

In the first part of this article, we present general-purpose recommender systems and how they differ from educational recommender systems. In the second section, we explore the main advantages and limitations of using recommender system in educational settings. Finally we conclude our work.

2 Recommender Systems and Educational Recommender Systems

The main purpose of a recommender system is to help users deal with information overload by pre-selecting information that could be helpful for reaching the desired goal [16, 17]. Recommender systems have their origins in fields such as ecommerce in which their main principle is to "use justifications to recommend products to consumers, and ensure to costumers like these products" [18]. However, their application has been extended to fields such as movies and music suggestion, bookstores, and education [15, 19]. Table 1 shows the main factors that influence the results of recommender systems.

Table 1 Factors that influence the recommendation for recommender systems

User profile	This includes all basic characteristics such as background, demographics and language [21, 40, 43, 49, 63, 71, 72]
Personal factors	Behavior, flexibility to accept or reject recommendations, interest, mood, motivation, trust, intuition and honesty [23, 27, 28, 61] privacy, awareness of other options, bookmarks, needs, interaction weight, interaction preferences, interactions between users [41, 58, 73, 74]
Resources	Content, thesaurus, taxonomy, tags, keywords, ratings, reviews, summary, contributors, date, number of votes [28, 29, 34, 39, 46, 48, 49, 60, 65, 71, 73]
Recommendation	Quality, credibility, measurability and weight of the recommendation, reliability, classification, date and time. It is recommendable to also include an explanation about why a resource is recommended and who the contributors are [28]
System	Accessibility, usability, parameters, goal, initial data, data analysis techniques, design, architecture, graphical interface [10, 42]

Recommender systems can be classified in three types Balabanovic and Shoham as cited in [14]:

- Content-based recommendation systems: The recommendations are based on choices that the user has done in the past. In education, examples of such systems are presented in [12, 19–22].
- Collaborative recommendations systems. Recommendations are related to user with similar profiles and preferences. Some educational collaborative recommender systems can be found in [23, 24–30].
- Hybrid approaches, which combine methods from both types of recommender systems.

However, there are important differences and factors between general-purpose recommender systems and educational recommender systems. The first difference that should be considered in educational recommender systems is the *goal*. In fields such as e-commerce, a user is looking for buying a product, with a specific quality and in a specific price range [16]. The goal of educational recommender systems is to help the user or a group of users to find suitable resources and learning activities for a better achievement of the learning goal and the development of competences in less time, [16, 31, 32]. However, even though principles of recommender systems may fit well with those in the learning sciences, it is important to note that recommender systems often need to be adapted in order to facilitate learning [33].

Additionally, in education, recommender systems are also commonly employed to suggest material that the instructor or the course designer can use for improving the course [34–37], help the instructor detect common misconceptions and to identify students who present difficulties [21], help students select their courses [38]; and assist with peer-matching and group formation [23, 30, 39, 40].

The second difference is the *context* [17, 36, 41, 42]. Even though most recommender systems share factors such as networks and peer information, the context in educational recommender systems is pedagogically related. Factors that should be considered as part of the context are pre and post requisites, timeframe, and instructional design [10, 43], pedagogical scenarios [9, 42], and social networks [44].

In other fields, recommender systems are mostly based on user tastes, personal preferences, or what a user likes or dislikes [9, 43]. The third difference is that educational recommender systems are highly influenced by *pedagogical factors* such as the learning history, knowledge, preferences, processes, strategies, styles, patterns, activities, feedback, misconceptions, weaknesses, progress, and expertise [10, 12, 13, 22–50, 16].

The fourth difference is *user's classification*. In a pedagogical context users are classified according to the function of the user (student, teacher, courseware developer) [34, 35] or the level of knowledge (beginner-intermediate-advanced); learning styles [23, 51–53, 16]. Additionally, it seems that pedagogical users tend to be more tolerant than in other fields such as e-commerce [48].

Regardless of the context where it is used, a recommender system has three main phases: (1) Input sources of the user's profile, data gathering, and data modeling; (2) recommending method as the interface mechanism; (3) output recommendations [39–49, 54]. Recommendations are usually activated when an event derived from user's behavior and interaction to the system takes place [19]. Educational recommender systems are also influenced by the *cycle of learning*, which is the last important difference. This cycle involves the following steps: (1) Design of the learning experience; (2) Administration of the environment; (3) Runtime; and, (4) Feedback [43]; learning process [49]; and evaluation [16].

3 Advantages of Introducing Recommender Systems in the Classroom

As we mentioned in the previous section, the main goal of recommender systems in education is to help users find suitable resources for a better achievement of the learning goal in less time. However, the benefits of introducing educative recommender systems go beyond learning goals. Based on the literature, we propose to classify the benefits from three points of view: (1) students' performance, (2) social learning enhancement, and (3) increased motivation.

3.1 Student Performance

From a student's perspective, the main benefit is to find better quality resources and to reach the learning goal [25, 39, 55–57]. Educational recommender systems can also help to identify students with problems and weakness [19, 58], to detect student's misconceptions [46], to help students to navigate in knowledge hyperspace, and to get a good quality information feedback [59]. That is useful for monitoring students and for adjusting the course content if necessary [58].

Educational recommender systems can also help promote personalized learning [12]. Personalization or adaptation is the process of adapting the application to the needs of the users based on what the systems knows about them [35]. In education, adaptive systems adjust the presentation and the content according to the student's profile and other factors [35]. Recommender systems can suggest, for example, learning paths taken by successful learners [24]. However, it has been shown that, even though recommender systems can increase student's performance, they help mostly students that are in most need for assistance [46].

3.2 Social Learning Enhancement

One of the main features of collaborative recommender systems is the inclusion of social interaction and social navigation [60]. where navigation history and bookmarks are visible to the others [61]. In an educational context, this social attribute promotes student collaboration [23], helps to find like-minded people, propagates the "word-of-mouth" from trusted and high quality resources [39], and enhances virtual community experiences [62].

3.3 Increased Motivation

There is also evidence that educational recommender systems have a positive feedback on student's motivation [63] by keeping them interested in the learning experience [23]. Additionally, these systems have been shown to improve the atmosphere of the learning environment [61], and enhance the interaction within the learning environment [45, 57].

4 Challenges

There are many challenges that should be overcome for the implementation of education recommender systems to be effective. In this section, we summarize the main challenges found in the literature and discuss some ways to address them.

Sparcity: It refers to "a situation in which transactional or feedback data is sparse and insufficient to identify similarities in consumer interests" [64]. It mostly influences the accuracy and quality of the recommendation and the interaction between members of a network [18].

Lack of structure in data: One of the main challenges and essential attributes of a recommender system is the way the data is structured [65, 66]. For example, in social learning environments, information tends to be classified in just one category reducing the number of options to the user [10]. In addition, there is not a predefined structure in the social network [11], and information cannot be reused in other systems because of a lack of structure, which hinders the interoperability among recommender systems [67].

First starter problem and cold start: This problem occurs when there are no ratings for new resources or when a new user has not rated any item [48]. Some ways to overcome this challenge are: (1) A knowledge provider can be the first starter; consequent users can contribute to this elaboration [28]; (2) Use of artificial learners [42, 48]; (3) Use of information related to completion of activities and similar preferences [11].

Information overload: This problem refers when in an environment such as internet, the amount of pedagogical content is overwhelming and widely spread over the network generating [22]. This leads to information overload, making hard to students to find and evaluate quality of the most suitable learning resources [12, 35].

Cognitive overload: Required effort in the process of selecting useful resources or assigning accurate ratings [28]. This takes place especially when there is raw data, the user is unable to ask the right question, pedagogical resources are not properly defined by the expert [13], resources are not classified; and, when there are not existing summaries, keywords or other types of descriptors [63]. A way to overcome this issue is to use content analysis techniques such as data mining for finding keywords or structures [63].

System-related: Misconceptions about recommender systems, unsatisfactory results, speed [28], lack of evaluation criteria [35], the system is not integrated to learning management systems [68], and scalability issues [8, 28].

Quality of the recommendation and trust: Another problem is that users do not trust the system and the recommendations. The probabilities that a user performs an action based on the recommendations many times is too low [8, 45]. For that, it is suggested to always define the quality of a recommendation [30], make clear when a recommendation should be precise or simply relevant [60], reduce biased recommendations as much as possible [8], and make clear where are the recommendations coming from [28] or how new items are added [28].

User Privacy: It is important to consider the privacy of the user and the inclusion of reflections and feedback [66]. However, it is important to be careful on the negative or positive effects on feedback [8].

5 Conclusions

Based on existing literature, we presented recommender systems and explored their unique characteristics in education. Additionally, we discussed how recommender systems can improve students' performance and enhance collaborative learning.

We can conclude that recommender systems can positively help students to reach the desired pedagogical goal. However, from a content perspective, the strength of the recommendations depends on the content-analysis tool. Data mining and other content analysis techniques have been from moderated to highly successful [15, 20, 21, 34, 69]. That depends on the algorithm, factors that are considered, and the context.

From a social perspective, the strength of the recommendations also depend on other users and social factors such as who is reviewing, rating, tags, the quality of the recommendation, and how much users trust it. In addition it has been proven that recommender systems also help to increase student's motivation, improve student's interest and to enhance the atmosphere in the learning environment. There is no right or wrong combination of factors. Depending on the case, they can be adjusted to take different values [10].

The main limitation of this work is the lack of empirical evidence. From the 98 articles that we considered for this literature review, just 37 presented empirical evidence based on human participants. Questions on the factors that should be considered in an educational system and how they should be contextualized depending on the learning experience remain open. As future work, it is suggested to conduct an evidence-based review. As noted by Abrami et al. these reviews "are important not only in meeting the needs of policy-makers and practitioners, but also in providing students and researchers with an overview of the evidence" [70].

Appendix A: Selected Resources

Databases

- Association for the Advancement of Computing in Education (AACE)/Education and Information Technology Digital Library (ED/ITLib): http://editlib.org/
- Association for Computing Machinery (ACM) Digital library: http://dl.acm.org/
- EBSCO: http://web.ebscohost.com/
- Engineering Village: http://www.engineeringvillage.com
- Institute of Electrical and Electronics Engineers (IEEEXPlore) digital database: http://ieeexplore.ieee.org
- Web of Science/Web of Knowledge: http://Wokinfo.com

Books

- Jannach (2011). *Recommender systems: An introdu*ction. New York: Cambridge University Press.
- Manouselis, Nikos, Drachsler, Hendrik, Verbert, Katrien, & Duval, Erik. (2012). *Recommender Systems for Learning*. Springer-Verlag New York Inc.
- Ricci, Rokach, Shapira and Kantor (2011). *Recommender Systems Handbook*. Boston, MA: Springer Science + Business Media, LLC.
- Santos and González (2012). *Educational recommender systems and technologies: Practices and challenges*. Hershey PA: Information Science Reference.

Journals and Conferences

- ACM Conference on Recommender Systems (Rec Sys)
- Computers in Human Behavior: http://www.journals.elsevier.com/computers-in-human-behavior/
- Computers and Education: http://www.journals.elsevier.com/computers-and-education/
- Expert Systems with Applications: http://www.journals.elsevier.com/expert-systems-with-applications/
- International Conference on Advanced Learning Technologies
- International Journal on E-Learning: http://www.aace.org/pubs/ijel/
- International Journal of Learning Technology: http://www.inderscience.com/jhome.php?jcode=ijlt
- Journal of Digital Information: http://journals.tdl.org/jodi/index.php/jodi
- Journal of Educational Technology and Society: http://www.ifets.info/
- Journal of Interactive Learning Research: http://www.aace.org/pubs/jilr/

- Knowledge-Based Systems: http://www.journals.elsevier.com/knowledge-based-systems/
- World Conference on Educational Multimedia, Hypermedia and Telecommunications
- World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education

References

1. Garcia-Martinez, S., van-Thiel, A.: Exploring the teaching and learning methods for data structure courses. In: Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, pp. 2709–2720. AACE, Chesapeake (2010)
2. Dillenbourg, P.: Collaborative Learning: Cognitive and Computational Approaches. Pergamon, Amsterdam (1999)
3. Driscoll, M.P.: Psychology of Learning for Instruction. Pearson A and B, Boston (2005)
4. Liccardi, I., Ounnas, A., Pau, R., Massey, E.P. et al.: The role of social networks in students' learning experiences. SIGCSE Bull. **39**(4), pp. 224–237 (2007)
5. Berger, P.L., Luckmann, T.: The Social Construction of Reality: A Treatise in the Sociology of Knowledge. Doubleday, Garden City (1966)
6. Vygotskii, L.S., Cole, M.: Mind in Society: The Development of Higher Psychological Processes. Harvard University Press, Cambridge (1978)
7. Cole, M., Y, Engeström.: A cultural-historical approach to distributed cognition. In: G. Salomon (Ed.), Distributed cognitions: psychological and educational considerations. Cambridge University Press, Cambridge, New York (1993)
8. Geyer-Schulz, A., Hahsler, M., Jahn, M.: Educational and scientific recommender systems: designing the information channels of the virtual university. International Journal of Engineering Education, **17**(2), pp. 153–162 (2001)
9. Tang, T.Y., McCalla, G.: A multidimensional paper recommender: Experiments and evaluations. IEEE Internet Comput. **13**(4), 34–41 (2009)
10. Abel, F., Bittencourt, I.I., Henze, N., Krause, D., Vassileva, J.: A rule-based recommender system for online discussion forums. Lect. Notes Comput. Sci. **5149**, 12–21 (2008)
11. Drachsler, H., Hummel, H., Koper, R.: Navigation support for learners in informal learning environments. In: Proceedings of the 2008 ACM conference on Recommender systems, ACM, pp. 303–306 (2008)
12. Yang, Y.J., Wu, C.: An attribute-based ant colony system for adaptive learning object recommendation. Expert Syst. Appl. **36**(2), 3034–3047 (2009)
13. DeLong, C., Desikan, P., Srivastava, J.: USER: user-sensitive expert recommendations for knowledge-dense environments. Advances in Web Mining and Web Usage Analysis, pp. 77–95 (2006)
14. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of- the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)
15. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. Expert Syst. Appl. **39**(11), 10059–10072 (2012)
16. Drachsler, H., Hummel, H.G.K., Rob, K.: Identifying the goal, user model and conditions of recommender systems for formal and informal learning. JODI. J. Digital Inf. **10**(2), 1–17 (2009)
17. Hu, J., Zhang, W.: Community collaborative filtering for E-Learning. In: International Conference on Computer and Electrical Engineering, IEEE, pp. 593–597 (2008)

18. Lu, J.: A personalized E-learning material recommender system. In 2nd International Conference on Information Technology for Application (ICITA 2004), pp. 374–379. Macquarie Scientific Publishing, Harbin, China (2004)
19. Hsu, M.H.: Proposing an ESL recommender teaching and learning system. Expert Syst. Appl. **34**(3), 2102–2110 (2008)
20. Bendakir, N., Aïmeur, E.: Using association rules for course recommendation. In: Proceedings of the AAAI Workshop on Educational Data Mining, pp. 31–40 (2006)
21. Romero, C., Ventura, S., Zafra, A., de Bra, P.: Applying web usage mining for personalizing hyperlinks in web-based adaptive educational systems. Comput. Educ. **53**(3), 828–840 (2009)
22. Wang, F.H.: Content Recommendation Based on Education-Contextualized Browsing Events for Web-Based Personalized Learning. Edu. Technol. Soc. **11**(4), 94–112 (2008)
23. Reategui, E., Zattera, C.: Do learning styles influence the way students perceive interface agents? Paper presented at the IHC '08: Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems, Porto Alegre, Brazil (2008)
24. Berg, vd, Eshuis, J., Waterink, W., Nadolski, R., Berlanga, A., Boers, N., et al.: Effects of the ISIS Recommender System for Navigation Support in Self-Organised Learning Networks. Edu. Technol. Soc. **12**(3), 115–126 (2009)
25. Abramowicz, W., Malyszko, J., Weckowski, D.G.: Discovering of users' interests evolution patterns for learning goals recommendation. Perspect. Bus. Inf. Res. **90**, 231–238 (2011)
26. Mitsuhara, H., Ochi, Y., Kanenishi, K., Ogata, H., Yano, Y.: An adaptive web-based learning system for providing new findings. Paper presented at the Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (2003): Honolulu. Hawaii, USA (2003)
27. O'Mahony, M.P., Smyth, B.: A recommender system for on-line course enrolment: an initial study. In: Proceedings of the 2007 ACM conference on Recommender systems, ACM, pp. 133–136 (2007)
28. Rafaeli, S., Dan-Gur, Y., Barak, M.: Social recommender systems: recommendations in support of E-learning. Int. J. Distance Edu. Technol. **3**(2), 30–47 (2005)
29. Saloun, P., Velart, Z.: Concept space rating for personalization of learning materials based on relations. In 4th International Workshop on Semantic Media Adaptation and Personalization, SMAP'09. IEEE, pp. 67–72 (2009)
30. Zheng, Y., Li, L.: A three-dimensional context-awareness model for peer recommendation in the E-learning context. Int. J. E-Learning **7**(1), 153–168 (2008)
31. Drachsler, H., Hummel, H.G.K., Koper, R.: Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model. Int. J. Learn. Technol. **3**(4), pp. 404–423 (2008)
32. Drachsler, H., Pecceu, D., Arts, T., Hutten, E., Rutledge, L., Van Rosmalen, P., Koper, R.: ReMashed–recommendations for mash-up personal learning environments. Learning in the synergy of multiple disciplines, pp. 788–793 (2009)
33. Buder, J., Schwind, C.: Learning with personalized recommender systems: A psychological view. Comput. Hum. Behav. **28**(1), 207–216 (2012)
34. Garcia, E., Romero, C., Ventura, S., Castro, C.: An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. User Model. User-Adap. Inter. **19**(1), 99–132 (2009)
35. Gasparini, I., Lichtnow, D., Pimenta, M.S., de Oliveira, J.P.M.: Quality ontology for recommendation in an adaptive educational system. In: Intelligent Networking and Collaborative Systems, 2009. INCOS'09, IEEE, pp. 329–334 (2009)
36. Ramadoss, B., Balasundaram, S.R.: Management and selection of visual metaphors for courseware development in web based learning. In: IEEE Conference on Cybernetics and Intelligent Systems. IEEE, pp. 1–6 (2006)
37. Staab, S., Studer, R., Stumme, G., Tane, J.: Accessing distributed learning repositories through a courseware watchdog. Paper presented at the Proceedings of World Conference on E-learning in Corporate, Government, Healthcare, and Higher Education, Montreal, Canada (2002)

38. Prins, F.J., Nadolski, R.J., Berlanga, A.J., Drachsler, H., Hummel, H.G.K., Koper, R.: Competence description for personal recommendations: the importance of identifying the complexity of learning and performance situations. Edu. Technol. Soc. **11**(3), 141–152 (2008)
39. Recker, M.M., Walker, A.: Supporting "Word-of-Mouth" Social networks through collaborative information filtering. J. Interact. Learn. Res. **14**(1), 79–98 (2003)
40. Yang, A., Li, J., Tang, Y., Wang, J., Zhao, Y.: The similar scholar recommendation in Schol@t. 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2012), pp. 666–70 (2012)
41. Santos, O., Boticario, J.: Guiding learners in learning management systems through recommendations. Learn. Syn. Multiple Disciplines, pp. 596–601 (2009)
42. Tang, T.Y., McCalla, G.: Laws of attraction: In search of document value-ness for recommendation. Res. Adv. Technol. Digital Libraries, pp. 269–280 (2004)
43. Santos, O.C.: A recommender system to provide adaptive and inclusive standard-based support along the eLearning life cycle. In: Proceedings of the 2008 ACM conference on Recommender systems, ACM, pp. 319–322 (2008)
44. Yang, Q., Sun, J., Wang, J., Jin, Z.: Semantic web-based personalized recommendation system of courses knowledge research. 2010 International Conference on Intelligent Computing and Cognitive Informatics (ICICCI 2010), pp. 214–217 (2010)
45. Chang, C-W., Liu, C-C., Horng, J-T., Chen, G-D.: Students' trust-aware learning material recommendation system. Paper presented at the Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Montreal, Canada (2005)
46. Masters, J., Madhyastha, T., Shakouri, A.: ExplaNet: A collaborative learning tool and hybrid recommender system for student-authored explanations. J. Interact. Learn. Res. **19**(1), 51–74 (2008)
47. Prieto, M., Menéndez, V., Segura, A., Vidal, C.: A recommender system architecture for instructional engineering. Emerging Technologies and Information Systems for the Knowledge Society, pp. 314–321
48. Tang, T.Y., McCalla, G.: Smart recommendation for an evolving E-learning system: Architecture and experiment. Int. J. E-Learning **4**(1), 105–129 (2005)
49. Wan, X., Okamoto, T.: Building learner profile for group learning recommender system from learning process. In: International Workshop on Intelligent Systems and Applications, ISA 2009, IEEE, pp. 1–4 (2009)
50. Wang, T.I., Kun Hua, T., Ming Che, L., Ti Kai, C.: Personalized learning objects recommendation based on the semantic-aware discovery and the learner preference pattern. J. Edu. Technol. Soc. **10**(3), 84–105 (2007)
51. Bobadilla, J., Serradilla, F., Hernando, A.: Collaborative filtering adapted to recommender systems of E-learning. Knowledge-Based Syst. **22**(4), 261–265 (2009)
52. Coffey, J.W.: Categories of inference in a multi-faceted, educational, knowledge-based recommender system. Interaction **12**, 207–241 (2004)
53. Mangina, E., Kilbride, J.: Evaluation of key phrase extraction algorithm and tiling process for a document/resource recommender within E-Learning environments. Comput. Educ. **50**(3), 807–820 (2008)
54. Khribi, M. K., Jemni, M., Nasraoui, O.: Toward a hybrid recommender system for E-learning personalization based on web usage mining techniques and information retrieval. In: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, **2007**(1), pp. 6136–6145 (2007)
55. Huang, C., Ji, Y., Duan, R.: A semantic web-based personalized learning service supported by on-line course resources. 2010 6th International Conference on Networked Computing (INC), p. 7 (2010)
56. Ghauth, K.I., Abdullah, N.A.: Measuring learner's performance in E-learning recommender systems. Australasian J. Edu. Technol. **26**(6), 764–774 (2010)
57. Wang, P., Yang, H.: Using collaborative filtering to support college students' use of online forum for English learning. Comput. Educ. **59**(2), 628–637 (2012)

58. Hsu, M.H., Chang, J.H., Yu, S.Y.: Recommending english grammar teaching strategy for EFL teachers. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications, **2009**(1), 2854–2861 (2009)
59. Pinter, R., Marusic, T., Radosav, D., Cisar, S. M.: Recommender system in E-student web-based adaptive educational hypermedia system. In MIPRO, 2012 Proceedings of the 35th International Convention, IEEE, pp. 1270–1273 (2012)
60. Ruiz-Iniesta, A., Jimenez-Diaz, G., Gomez-Albarran, M.: Promoting strong personalization in content-based recommendation systems of learning objects. IEEE-RITA Latin Am. Learn. Technol. J. **5**(1), pp. 43–50 (2008)
61. Kurhila, J., Miettinen, M., Nokelainen, P., Tirri, H.: EDUCO: social navigation and group formation in student-centred E-learning. J. Interact. Learn. Res. **18**(1), 65–83 (2007)
62. Leino, J.: Case study: Recommending course reading materials in a small virtual learning community. Int. J. Web Based Communities **8**(3), 285–301 (2012)
63. Jie Chi, Y., Yi Ting, H., Chi Cheng, T., Chung, C.I., Yu Chieh, W.: An automatic multimedia content summarization system for video recommendation. J. Edu. Technol. Soc. **12**(1), 49–61 (2009)
64. Huang, Z., Chen, H., Zeng, D.: Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. ACM Trans. Inf. Syst. (TOIS) **22**(1), 116–142 (2004)
65. Dron, J.: The blog and the borg: a collective approach to E-Learning. Paper presented at the Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003, Phoenix, Arizona, USA (2003)
66. Pearce, J.: A system to encourage playful exploration in a reflective environment. Paper presented at the Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Chesapeake (2008)
67. Manouselis, N., Vuorikari, R.: What if annotations were reusable: a preliminary discussion. Adv. Web Based Learning–ICWL 2009, 255–264 (2009)
68. Santos, O.C., Boticario, J.G.: Recommendation strategies for promoting eLearning performance factors for all. In 6th Workshop on Intelligent Techniques for Web Personalization & Recommender Systems in conjunction with the 23rd AAAI Conference on Artificial Intelligence (2008)
69. Ghauth, K.I., Abdullah, N.A.: Building an E-learning recommender system using vector space model and good learners average rating. In 9th IEEE International Conference on Advanced Learning Technologies, IEEE, pp. 194–196 (2009)
70. Abrami, P.C., Borokhovski, E., Bernard, R.M., Wade, C.A., Tamim, R., Persson, T., Bethel, E.C., Surkes, M.A.: Issues in conducting and disseminating brief reviews of evidence. Evidence Policy **6**(3), 371–389 (2010)
71. Castillo, d., Peis, E., Moreno, J.M., Enrique Herrera-Viedma, Moreno, J.M.: D-Fussion: A semantic selective disssemination of information service for the research community in digital libraries. Inf. Res.: An Int. Electron. J., **14**(2), p. 398 (2009)
72. Wan, X., Ninomiya, T., Okamoto, T.: A learner's role-based multi dimensional collaborative recommendation (LRMDCR) for group learning support. In: Neural Networks. IJCNN 2008. (IEEE World Congress on Computational Intelligence), IEEE, pp. 3912–3917 (2008)
73. Vuorikari, R.: Can social information retrieval enhance the discovery and reuse of digital educational content?. In: Proceedings of the 2007 ACM conference on Recommender systems, ACM, pp. 207–210 (2007)
74. Wan, X., Ninomiya, T., Okamoto, T.: LRMDCR: A learner's role-based multi dimensional collaborative recommendation for group learning support. In: Eighth IEEE International Conference on Advanced Learning Technologies. ICALT'08, IEEE, pp. 603–605 (2008)

Melody-Based Approaches in Music Retrieval and Recommendation Systems

Bartłomiej Stasiak and Mateusz Papiernik

Abstract Analysis and description of digital audio, being one of the fundamental information sources in the modern world, becomes surprisingly complex when it comes to music information retrieval. In addition to the numerous known classic problems, such as pitch and rhythm extraction, tempo analysis, audio key detection or chord estimation, there exist further considerable difficulties in higher-level analysis involving e.g. emotional content or genre categorization. These issues are of utmost interest for the rapidly growing music industry and new e-commerce solutions which necessitate the need for precise estimation of users' musical preferences. Out of the many proposed approaches, those based on modelling the social context and behavior of a user via collaborative filtering techniques, although potentially effective, do not offer real insight into the decision-making process. The content-based methods, on the other hand, enable to build refined models of users' preferences with adjustable weights attributed to different musical elements and varied description levels. This chapter presents an evaluation of the latest efforts and development in fields related to music information retrieval, as well as a study of melody based approaches for use with recommendation systems. The evaluation concentrates on algorithms extracting audio features relevant to the idea of similarity analysis and recommendation proposals. Only content-based audio analysis approaches are taken into account. Tasks such as audio key detection, chord estimation, tempo analysis or genre classification are described, latest achievements and algorithms are presented and ways in which they can be effectively used in recommendation systems are estimated. The concept of audio melody extraction is explored and the authors' own research results in this subject are presented. Possible usage scenarios of melody analysis for music

B. Stasiak (✉) · M. Papiernik
Institute of Information Technology, Technical University of Łódź, ul.Wólczańska 215, 90-924 Łódź, Poland
e-mail: bartlomiej.stasiak@p.lodz.pl

M. Papiernik
e-mail: mati@maticomp.net

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*, Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_9,

recommendation services are evaluated. Application in conjunction with query by humming and query by tapping systems is analysed as an example of a complete recommender solution.

1 Introduction

Music is everywhere. Not only has it the potential to fill up all the available spaces we live in—at home, at work, in our cars, in shops, offices, clubs and concert halls, but it can also build bridges or walls between people, being an important element of our social space. Musical preferences may reveal a lot about our character and other aspects of our individual “inner space” as, in general, they constitute an inseparable part of the human personality itself.

Getting down to more measurable spaces—millions of hours of music available in music stores all over the world are enough to fill up every second of human life, making the problem of *selection* a real challenge. This is where the music recommender systems come into play. Similarly to the tools used to anticipate the users’ preferences and needs for other types of goods and multimedia, most notably for films and books, the systems used for music recommendations are based on either of two main paradigms—collaborative filtering or content-based approach [1].

The first method or group of methods is generally based on the analysis of preferences of large groups of users inferred from their purchase history. Discovering similarities and application of clustering techniques enables to formulate recommendations for users, such as e.g. the Amazon’s famous “customers who bought this item also bought...”. A list of the most known music recommendation engines based on this approach include i.a. *Genius*—the feature implemented by Apple in *iTunes* and *Audioscrobbler* included in *Last.fm* internet radio station service. In collaborative filtering, the more a system knows about a particular user, i.e. the more products he or she has bought, the more accurate preference model may be built. Similarly, popular songs which have many ratings may be precisely included or excluded from the playlist suggested by the system. On the other hand, new users or new songs may often generate inadequate recommendations, which is known as the *cold start* problem. This problem may be avoided, at least in the context of the analysis of unknown songs, by direct audio analysis performed by systems from the second group, i.e. those representing the content-based approach. One of the best known systems of this kind is created by Pandora Media Inc on the basis of the Music Genome Project in which every piece of music is encoded with several hundred attributes (“genes”). In fact, this system relies on human experts which have to manually label every recording with appropriate attribute values. Possible solutions to this apparent limitation may include, apart from pure automatic content-based analysis, also social tagging and music labeling games [2].

It is interesting how both approaches, although based on completely different models, can serve as an effective tool for construction of music recommendation systems. The collaborative filtering is probably more eagerly used as more general, based on well-established methods that proved successful in many other fields of e-commerce (e.g. selling movies and books). The cold start problem may be partially solved on registering a new user, e.g. by uploading the list of songs he or she already has in his/her collection. As for the scalability, the size of multimedia databases and the number of users of the most popular music recommendation systems is counted in millions which actually requires substantial computation power. On the other hand, exploiting such a rich information source may lead to significant enhancement in accuracy of the generated recommendations. The systems based solely on collaborative filtering approach may behave as if they performed detailed analysis of the content of the audio files. This phenomenon was studied in Barrington et al. [3], where Apple's Genius engine was compared to a content-based approach. However, the known weakness of methods based on popularity and social context analysis is that they tend to fail in case of e.g. unknown songs recorded by an obscure band, irrespective of the actual similarities of the audio content. This is one of the reasons for the hybrid systems, including both content-based and content-independent solutions, becoming more and more popular recently.

In this chapter we will focus on a certain group of the content-based approaches—we will try to consider the music similarity problem from the *melody similarity* point of view. This is somewhat specific. It is true that melody is one of the most elementary ingredients of almost every composition or recording which a~ user may like or dislike. As an analogy, it may be compared to a line drawing or a~ contour being the part of an image which contains most information about the depicted objects, such as shapes, positions or e.g. facial expression of people. On the other hand, there exist pieces in which other components play the dominant role, such as most of the art of the Impressionists who generally refrained from using explicit lines or contours, both in music and in visual arts. We will therefore briefly present also other important aspects of music description, formulated as separately solved Music Information Retrieval (MIR) tasks, which may collectively contribute to content-based music recommendation solutions.

One of the best forums for observing the current development in the MIR domain is the Music Information Retrieval Evaluation eXchange (MIREX) [4], an annual event accompanying the International Society (and annual Conference) for Music Information Retrieval (ISMIR) [5]. It is hosted by the International Music Information Retrieval Systems Evaluation Laboratory (IMIRSEL) at the Graduate School of Library Information Sciences—a part of the University of Illinois at Urbana-Champaign. MIREX, organized every year starting from 2005, is basically a kind of benchmarking contest, offering the researchers from all over the world a~ possibility of testing and comparing their algorithms on large real-world datasets, thus giving insight into the bleeding edge technology and the latest research efforts of the MIR community. Several MIR tasks, their goals and objectives, as well as the latest results reported by the MIREX participants will be presented in the following part of this chapter.

2 Music Similarity and Classification

One of the major groups of MIR related tasks can be labelled as music similarity and classification tasks. They concentrate on extracting information from audio signal that will enable algorithms to perceive certain pieces as similar to each other, or belonging to a certain style or genre. It is one of the most popular groups of tasks related to Music Information Retrieval, due to multitude of possible applications in plagiarism detection, cover song identification and recommender systems, very important factor being the ability to include knowledge about musical likings (therefore pieces already associated with the listener) as a reasoning base.

2.1 Train/Test Tasks

Music is a very complex phenomenon and its numerous aspects may be described on a variety of levels. This truism has to be remembered when automatic music analysis is considered, as the goal of audio signal processing significantly influences the applied methods and tools. There is however a common methodology which is used mostly for the highest levels of music description and recognition, i.e. the train/test procedure in which a large collection of real-world audio data is used for teaching the classifier and the consecutive testing stage is based on similar but different dataset. The list of tasks which use this scheme includes:

1. Artist identification
2. Composer (classical) identification
3. Music genre classification
4. Mood classification.

The task details depend primarily on the dataset construction and content. Music genres are organized in a hierarchy [6] or a tree structure and only a part of this tree, e.g. classical music, jazz, etc. may be considered [7]. In MIREX competition two subtasks, each including 10 genres, are formulated: Audio US Pop Music Genre Classification (7,000 files) and Audio Latin Music Genre Classification (3,227 files). Separate Audio Classical Composer Identification subtask includes 11 composers from Baroque to Romantic era. Audio Mood Classification subtask comprises 5 categories labeled by human experts with a set of descriptive adjectives.

2.2 Methods and Tools

Human unique abilities to recognize the style of the music, its mood or the author/composer are experience-based and dependent on various cultural influences and, as such, difficult to imitate with some simple criteria or rules. General machine learning methods are therefore used on the basis of standard feature sets known for their robustness and classification potential.

The input signal is segmented into fixed-duration, partially overlapped frames (usually from 10 to 80 ms) and for each frame a set of features is computed. In some problems a more refined approach, in which the signal is segmented according to e.g. beat positions, is preferred (cf. Sect. 3.2). Apart from some simple descriptors, such as zero-crossing rate (ZCR) or signal energy, which may be computed in the time domain, most features computed for a single frame are frequency-based. The DFT transform, preceded by application of Hamming, Blackman or other windowing function and sometimes also pre-emphasis filter, is typically performed as the first step of computation. Several spectral features are known to yield good results in MIR problems [6]:

- Spectral centroid (the center of gravity of the DFT magnitude spectrum):

$$C_i = \frac{\sum_{n=0}^{N-1} nX_i(n)}{\sum_{n=0}^{N-1} X_i(n)} \tag{1}$$

  and higher-order moments (kurtosis, skewness, etc.) represent basic information on the shape of the spectrum.
- Spectral rolloff, defined by a threshold frequency dividing the spectrum into two parts containing predefined percentage of the signal energy, which also describes the spectrum shape.
- Spectral flatness measure—proportion of the geometric mean of the magnitude spectrum versus its arithmetic mean:

$$F_i = \frac{\left(\Pi_{n=0}^{N-1} X_i(n)\right)^{1/N}}{\frac{1}{N}\sum_{n=0}^{N-1} X_i(n)}, \tag{2}$$

  which may be used to distinguish between a harmonic sound or a mixture of harmonic sounds with distinct partials (sharp peaks in the spectrum) from noise-like sounds which have flat spectrum.
- Mel-Frequency Cepstral Coefficients (MFCCs) which represent the general shape of the spectrum in a robust way, resembling human perception owing to logarithmic scaling of the frequency domain. The filterbank used for MFCC coefficients computation comprises triangular filter functions spaced equidistantly on the mel scale [8, 9]. Each filter function $H_k(f)$ is defined as:

$$H_k(f) = \begin{cases} \frac{f-l_k}{c_k-l_k}; & \text{for } f \in [l_k, c_k], \\ \frac{r_k-f}{r_k-c_k}; & \text{for } f \in [c_k, r_k], \\ 0; & \text{otherwise}, \end{cases} \quad (3)$$

where f is the frequency and the parameters l_k, c_k, r_k are given as:

$$\begin{aligned} c_k &= \mu(kd), \\ l_k &= \mu((k-1)d), \\ r_k &= \mu((k+1)d), \end{aligned} \quad (4)$$

where $k = 1$, ..., $k$is the filter function number and $\mu(m)$ is a mel-to-Hertz conversion function:

$$\mu(m) = 700\left(10^{\frac{m}{2595}} - 1\right). \quad (5)$$

The parameter d in formulas (4), defines the distance between consecutive filters and their width in the mel scale (Fig. 1) The output sequence of the filterbank (3)—for any given sound frame—is transformed with logarithmic function and the discrete cosine transform (DCT) is computed. The MFCC coefficients are defined by the initial output values of the DCT.

Feature selection and/or some additional operations, such as e.g. PCA whitening may be applied on the obtained feature set to obtain more compact representation. An important decision to take is how to use the features representing a~ single frame to classify the whole piece. Each frame may be classified separately

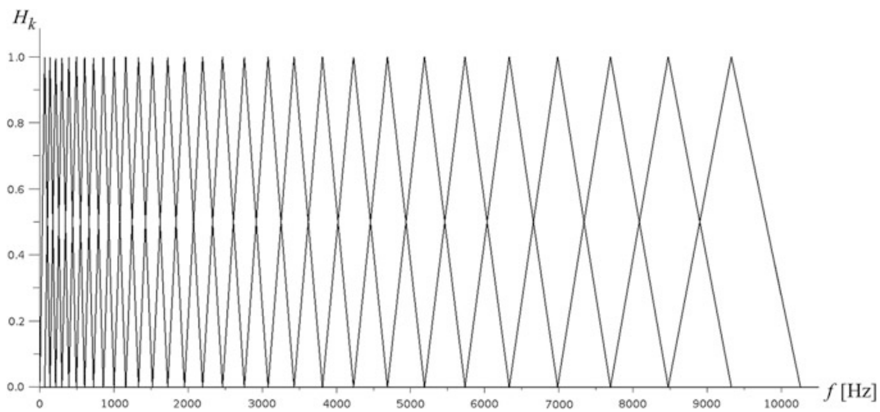


Fig. 1 MFCC filterbank with triangular filter functions spaced equidistantly on the mel scale ($K = 30$, $d = 100$)

and a simple majority rule may be applied afterwards [7]. It is however convenient to apply some kind of pooling based on statistics (e.g. mean, variance and higher-order central moments) of several consecutive frames. The authors of the best solution submitted to MIREX 2011 in genre and composer recognition subtasks tested two models [10]: Pooled Features Classifier (PFC) and Multi-Time-Scale Learning model (MTSL), obtaining slightly better results with the second one. A~ different approach based on block-level audio features extraction has been proposed in Seyerlehner et al. [11].

The classification is usually performed with standard machine-learning tools, such as multi-layer neural networks [10] or support vector machines (SVM) [12]. The authors of the best algorithm in Audio Mood Classification subtask (MIREX 2011), based on Gaussian Mixture Model (GMM), proposed to integrate also some elements of visual analysis into the audio classification system. The point is to consider the sound spectrogram as an image or, more precisely, as a set of 7 subimages corresponding to octave-based subbands of the spectrum. Image filtering is then performed with Gabor filterbank composed of 6 different orientations and 5 scales and the final features are defined by the mean and standard deviation of the filtering result.

2.3 Audio Tag Classification

This task is quite similar to the problems presented in Sect. 2.1, as it concerns a very subjective matter—how to describe a piece of music with a small number of appropriate *tags* (usually nouns or adjectives, e.g. "drums", "instrumental", "soft"). However, in this case a recording may be associated with multiple tags, instead of a single label.

There are two datasets used for this task in MIREX competition:

- MajorMiner game collection [2]
- Mood tag dataset [13].

The MajorMiner game is based on the following rules:

- A user is presented with a short music clip
- He/she describes the clip with a tag or a list of tags
- Each tag may earn a point in a way that promotes innovative, yet descriptive tags:
  - If many other users have used the same tag for this clip previously: 0 points
  - If exactly one other user has used the same tag for this clip: 1 point
  - If nobody has used the same tag: initially 0 points, but if somebody uses the same tag later—the user gets 2 points (so every time a user scores one point, another user has just scored two).

The first dataset comprises 2,300 short (10s) music clips representing 1,400 different tracks on 800 different albums by 500 different artists. These clips has been tagged with 73,000 tags by the MajorMiner participants, although only 12,000 of them has been verified by at least two users. For the MIREX audio tag classification task only a group of the most popular 43 tags, that have been verified at least 35 times, has been chosen. The second dataset contains 3,469 songs tagged by human experts with 135 unique tags from 18 mood groups.

There are two types of evaluation applied to algorithms submitted to the MIREX contest. Binary evaluation is used to assess the classification potential of the proposed methods and affinity (ranking) evaluation enables to asses the relevance of each tag, so that higher scores are given when correct tags appear at the beginning of the generated tag list.

The solutions proposed by MIREX participants are mostly coherent with methods and algorithms used for problems presented in Sect. 2.1. In fact, many authors submit the same algorithms for both genre or mood classification and for tag prediction [12].

2.4 Cover Song Classification

This task is in fact a replacement for the artist identification task mentioned in Sect. 2.1. The important difference between the two lies in a general approach to music description and analysis. This task was introduced in response to observations that simple timbral features, such as Mel Frequency Cepstral Coefficients (MFCC) do not precisely reflect the way the music is perceived by a human listener [14]. This is reflected in a so-called "album effect" [15], i.e. a situation when secondary qualities of the sound, common to all the tracks of a given CD due to the same recording equipment parameters and acoustics, influence the efficiency of classification.

In order to avoid this problem the dataset must be properly constructed so that no track from the same album appears in both the training and testing sets simultaneously. Introducing this *artist filtering* mechanism into MIREX evaluation led to significant decrease in artist recognition results in MIREX 2007 (48.14 %) with respect to the results from MIREX 2005 (72.45 %) [14]. It also demonstrated the need of searching for music representations which would go beyond the timbral qualities into the music information domain.

The Audio Cover Song database used in MIREX contains 30 cover songs, each represented by 11 different versions. The versions of the same cover song may be recorded by different artists with changed instrumentation or otherwise substantially altered, yet being undoubtedly variations of the same song for a human listener. Additional 670 "noise" tracks are added to increase the search space. The algorithms are evaluated with the top-ten ranking scheme, i.e. for any version of a~ given cover song the remaining 10 versions should be returned.

2.4.1 Methods and Tools

As the role of general spectral characteristics is deliberately confined in this task, the music content-level analysis must be performed instead. The methods are often coincident with those applied for polyphonic, symbolic music similarity assessment.

The tonality of a song and subsequences of consecutive chords are among the most potentially useful descriptors which may be used for song identification. *Chromagram* or a *pitch class profile* [16, 17] is a basic tool for compact representation of a chord. It is defined as a 12-dimensional vector where each dimension corresponds to one of the 12 semitones in an octave. More precisely, these 12 values represent the intensities of all 12 semitones from C to B, irrespective of the octave they appeared in [17]. More compact representation may be obtained with Tonal Centroid Transform [18] which reduces the dimensionality by half. Thresholding operation enables to obtain 6-element binary vectors, which may be effectively encoded as symbols from 64-element alphabet. A whole recording may be therefore represented in a robust way with a sequence of symbols and string matching algorithms may be finally applied for similarity assessment.

2.5 Audio Music Similarity and Retrieval

The concept of evaluating similarity between two distinct musical pieces is closely related to the basic idea of recommender systems and recommendations in general. Listeners, and in particular—casual listeners, tend to discover new music based on what they already like. In other words—following similarity patterns. There is much smaller possibility that one would choose a next piece to listen to randomly, from a genre he/she never listened to before, than the possibility one would look for music similar to what he/she already likes. That becomes even more apparent nowadays, when the size of digital music collections are incomparable to what we have known in the past. Popular live streaming services such as Last.fm [13], Spotify [19] or Grooveshark [20] provide people with hundreds of thousands of musical pieces they can choose from on everyday basis. Listeners are now not limited by what CDs they have on their shelves.

On one hand the virtually limitless possibilities of music choice is a blessing, on the other hand it becomes increasingly difficult to browse such large musical databases. Nobody will stop on a track-by-track basis to look for another piece to listen to. People expect the digital platforms to provide them with a sensible choice based on their previous playbacks. The same is equally important in e-commerce platforms, considering the steady growth of online music sales, especially on a per-track basis. Customers expect sensible recommendations based on their previous purchases. All these are best achieved by solving the problem of automated music similarity evaluation.

The task related to Audio Music Similarity and Retrieval has been explored during MIREX 2006, 2007, 2009, and 2010 evaluations and has returned in the 2012 edition of MIREX [4]. The inherent problem with evaluating the success rate of similarity evaluation systems is the fact that similarity is a subjective characteristic. There exists no ground truth that can be used as a basis for evaluation. The combination of objective similarity by genre and a large scale human judgement is used as an evaluation base, i.e. the human committee is responsible for judging the output of the algorithms solving the posed problem.

2.5.1 Methods and Tools

Several approaches have been explored during the evaluations to provide similarity results. The submission with the highest success rate in the 2010 evaluation is based on audio features extracted at a block-level [21]. A block is defined as a~ consecutive set of frames spanning the block width, each of which is transformed into spectral representation using STFT. Thus, a block can be perceived as a~ matrix with as many columns W as many frames there are in the block width, and with rows representing H subsequent frequency bins of the Fourier transform.

$$\begin{matrix} b_{H,1} & \cdots & b_{H,W} \\ \vdots & \ddots & \vdots \\ b_{1,1} & \cdots & b_{1,W} \end{matrix}$$

Several block-level features are taken into account. Spectral pattern that represents the timbral contents of the audio block, delta spectral pattern that involves a difference between two blocks shifted by several frames to emphasise on-sets, a logarithmic fluctuation pattern extracting information about rhythmical patterns, as well as spectral contrast pattern that approximates the *tone-ness* of the signal.

The extracted block-level features are used with Manhattan distance to compute pairwise distances for each set of features. Then a variant of distance space normalization (DSP) method is used to calculate single similarity estimate from the distances, taking into account the fact that different distances are measured on different scales.

Other highly successful approaches involve using MFCC coefficients to model the timbral characteristics, variations on Fluctuation Patterns to model the rhythmical aspect of the pieces as well as inharmonicity, odd to even harmonic energy ratio, spectral centroid, spread, skewness, kurtosis, spectral energy bands, zero-crossing rate and others. Different distance measures have been evaluated to provide best accuracy with such a broad range of audio features taken into account in the comparisons [22, 23].

3 Rhythmical Analysis

From the broadly defined rhythm, various valuable information can be extracted. It is also inherently difficult to analyse rhythmical structures, as it has always been an ambiguous task even for a trained musician.

3.1 Tempo Estimation

The information about the tempo is, apart from melody and harmonic structures, the most important factor describing a musical piece. Various kinds of information can be guessed from the tempi alone, making the tempo a very valuable descriptor to extract from an unknown audio signal. The tempo dictates the perceived *feel* of the piece. It can be used, among other applications, to recommend pieces with a similar beat to prepare a coherent listening experience, or to aid an automated DJ equipment. The tempo estimation is also a basis for rhythmical analysis, which in turn can be a principal component of query by tapping expert systems.

The problem of estimating the tempo from musical audio is a complex task, given that even a skilled musician can deem the tempo ambiguous in certain fragments of the piece he/she is unfamiliar with. Often multiples of the base tempo are chosen in tempo predictions, as well as completely different metrical levels may be chosen among the listeners.

Several broadly different approaches have been evaluated at MIREX in recent years. The most successful submission published in Gkiokas et al. [24] involves a combined approach. First the MFCC coefficients are used as a filterbank vector for each time frame. It is assumed that MFCC vectors will share similar characteristics on the beats. Alongside the MFCC coefficients, tonal characteristics are extracted using constant Q transform followed by a post-processing, resulting in 12-dimensional feature vectors for each frame, describing the chroma strength of each of 12 tones in an octave. Each feature vector set—both based on the MFCC coefficients and chroma vectors—is convolved with a set of resonators, each corresponding to a different base tempo. The convolution results correspond to the strength of a~ given tempo in the feature set. Results from chroma-based and MFCC-based analysis are then combined to form the final result.

3.2 Beat Tracking

In this task the precise locations of beats are to be found within a music recording. Similarly to the tempo estimation problem, the periodicity of the signal must be determined but with an additional goal of finding the beginning of the period. An associated problem of *downbeat* tracking is to determine the position and role of a given beat within a bar.

The problem significantly depends on the genre of music under consideration. Better results are usually obtained for rock, pop or dance music, while classical music or jazz analysis may be problematic [25].

Two datasets were used for MIREX 2011 contest. The first one contains 160 short (30s) music clips with stable tempi and meters, annotated manually by 40 different listeners. The second dataset is the collection of 367 Chopin Mazurkas (whole pieces). The tempo in many recordings from the second set is highly variable which enables to verify the ability of the submitted algorithms to track tempo changes within a single audio file.

3.2.1 Methods and Tools

The beat tracking problem is usually considered a higher-level approach to music analysis w.r.t. tempo and meter estimation. The tempo of a recording, expressed in beats-per-minute, and meter (e.g. 2/4 or 3/8) constitute therefore a part of the input to a beat tracking algorithm, in addition to raw audio data.

Several different approaches have been proposed in literature, including probabilistic approach [26–28], application of a bank of resonating comb-filters [29], dynamic programming [30] and HMM [25, 26, 31]. Some kind of onset-energy-function [25] or onset detection function (ODF) [26] is usually the starting point for further analysis. The periodicities of the ODF are detected by means of autocorrelation [32], comb filter resonators [31] or short-time Fourier transform [32]. Time-frequency reassignment (TFR) technique [26, 32] is commonly used here to reduce the problem of the time-frequency trade-off.

The downbeat tracking is often based on the observation that harmonic changes occur most often on downbeat positions. Chord analysis, usually exploiting some kind of chroma-based representation, such as chromagram [25] or chroma variation function (CVF) [26] is therefore applied.

The authors of one of the best solutions submitted to MIREX 2011 [26] proposed to describe rhythmic events associated with beat/downbeat segments location with a dictionary of words, where each word is based on the alphabet composed of 5 units: beat pre-attack, beat attack, downbeat pre-attack, downbeat attack and no beat. Hidden Markov model (HMM) without self-transitions (which would make periodicity modeling difficult) and an additional word sequence modeling layer is applied for beat/downbeat tracking afterwards.

A different approach to HMM application for periodic events representation is proposed in Peeters and Papadopoulos [25]. The hidden states $s_{i,j}$ are defined as "time t_i is a beat and it is in a given β_j", where β_j is defined as the position of the jth beat relative to the beginning of a bar which contains it. A special "reverse" Viterbi algorithm must be then applied for decoding. An important step in HMM model building proposed in Peeters and Papadopoulos [25] is application of beat-template models which are obtained by training based on linear discriminant analysis (LDA).

4 Structural Analysis

Apart from rhythm, other types of higher-level information can be extracted from the audio signal. These can be jointly classified as structural information—the melody, tonality, chord structures, or musical and logical structures (such as, for example, chorus parts). Different approaches exist depending on the particular task considered.

4.1 Key Detection

The majority of music around us is tonal, with *tonality* referring primarily to *functional tonality*. In functional tonality there exists a concept of musical key, which refers to the tonic note and a certain chord. Every chord found in a particular key has a relationship with other chords in this key. Tonal analysis is therefore dependent on the knowledge of the key a piece has been written in [33]. Tonal analysis, in turn, opens the doors to analysis of musical structures, such as chords, cadences, ornaments or harmonic progressions. All of these can be used to help determine the mood and a genre of a piece. As such, there is a need for automated key detection, to aid further analysis of the piece.

Several methods of automated key detection have been proposed during evaluations in consecutive years. The algorithm with highest success rate in the 2011 evaluation strongly relies on the Harmonic Pitch Class Profiles (HPCP) [34]. Harmonic Pitch Class Profiles extend the concept of Pitch Class Profile descriptor proposed by Fujishima [16]. It is a sequence of feature vectors that describe the intensity of each of the 12 pitch classes located on the well-tempered scale within a~ single analysis frame, thus depicting the tonality.

HPCP profiles are extracted from the training database of key-labeled musical pieces, and from the extraction results a sequence of 24 PDFs (probability density functions) is formed, 12 for major and 12 for minor keys. The generated PDFs are then used to determine the musical key of a piece on the basis of the known HPCP profiles of that piece. The pseudolikelihood that the extracted HPCP belongs to one of the 24 musical keys is calculated and the most probable key along with its mode (major or minor) is chosen as the predominant key of the piece. All computations are carried on a pre-processed signals—the most important factor being the analysis confined to the first 30 s of the audio signal to avoid possible key modulations within a piece.

Other approaches to key detection also base mostly on different descriptions of pitch profiles (chroma features) characteristic for certain minor/major keys, e.g. those proposed by Krumhansl [35] or Temperley [36]. Pitch profiles describe the average distribution of pitches for pieces in a given key, the important factor in creating these models being the presence of characteristic chords for a given key.

The match between spectral characteristics of an analysed piece and the key pitch profiles is being decided using various methods, including machine learning algorithms and Hidden Markov Models.

4.2 Chord Estimation

Estimating chord sequences within a musical piece is crucial for many application in music information retrieval. Chords are the principal harmonic structures in tonal music, thus the sequence of chords in a piece can be used to infer a variety of information about the genre, style or period from their characteristics alone. The information gathered from extraction of harmonic structure of a musical piece can be used for segmenting, semantic analysis as well as finding similar pieces (therefore being applicable for recommender systems).

The most robust solution in terms of the achieved accuracy, proposed in Ni et al. [37], is based upon complex Hidden Markov Model structures trained to determine keys, chords and bass notes at the same time. The signal processing used in this method involves a novel approach of loudness based chromatogram that resembles how humans perceive loudness. The loudness perception is not linear to the power spectrum neither is it uniform among the frequency bands. The proposed chromatogram uses the constant Q transform as its basis:

$$X_{s,t} = \sum_{n=t-\frac{L_s}{2}}^{t+\frac{L_s}{2}} x_n w_n \exp(\frac{-2\pi st}{L_s}) \tag{6}$$

where L_s is a frequency-dependent bandwidth. The power spectrum from the constant Q transform is taken, which is further normalized with reference power factor p_{ref}. The logarithm $\log_{10}$ is taken to scale the power spectrum into a sound power level.

$$L_{s,t} = 10\log_{10}(\frac{||X_{s,t}||^2}{p_{ref}}) \tag{7}$$

The loudness-based chroma is taken independently for bass and treble bands, which are in turn additionally compensated with pitch-based weighting to closely resemble human hearing.

This chroma information calculated from the learning database is used to teach the HMMs. The process is repeated during the decoding stage, in which an unknown piece is processed in the same way and HMMs are used to determine the probable chord sequences.

Different approaches involve learning-based systems using other variations of the Hidden Markov Models methods, as well as Gaussian Mixture Models (GMM) [38] and Support Vector Machines. Several attempts have been made to implement a chord detection algorithms with no prior training needed, the most successful case being a feature smoothing method proposed in Cho and Bello [39] based on predefined binary templates.

4.3 Audio Melody Extraction

The results of the MIREX 2011 exchange on the task of Audio Melody Extraction is the representative evaluation of the latest developments in the field.

The most successful submissions in overall accuracy measure were proposed by Salomon and Gomez [40] and Eokhwan Jo and Yoo [41]. No algorithm scored over 70 % in the overall accuracy measure, clearly illustrating that there exists a~ great research field to pursue. The results in raw pitch accuracy measure, where voicing (the presence or absence of perceived melody) was not taken into consideration, are different—best results being yield by Liao et al. [42].

The best algorithm in terms of raw pitch accuracy explores heavily the concepts of psychoacoustics. Perceptual qualities such as perceived loudness, timbre and masking effect are taken into account.

The primary algorithm for selecting spectral peaks representing fundamental frequencies that are selected as F_0 candidates for melody estimation rely mainly on the masking effect. The masking effect represents how an audible tone can mask other spectral components in its neighbourhood—the effect observed e.g. when a~ passing car can make a discussion inaudible. For each frequency a mask contour is estimated on the bark psychoacoustic scale [43] using the spread function.

The complete masking curve is calculated from maximum masks overlaying at each bin in STFT time frames. For F_0 candidates, the frequencies standing out above the masking curve are selected.

Utilising psychoacoustic factors in determining the F_0 candidates leads to increased accuracy in pitch detection and has been widely pursued in other forms by other contributors as well.

As for the overall accuracy estimation, the most successful algorithm explores the concepts of psychoacoustics and human perception as well.

The signal is first pre-processed with equal loudness filter in an attempt to better approximate human perception of loudness depending on frequency. In the preprocessing stage the phase spectrum from STFT is used to filter out noise components from the spectrum leaving a cleaner representation.

From the spectral peaks representing the F_0 candidates a salience function is generated, which is a representation of pitch salience over time. The method for generating the salience function proposed by the authors involve both chroma feature analysis (the similarity in spectral characteristics between the same tone in different octaves) as well as psychoacoustic assumptions regarding the octave range in which the melody is most probably to be found.

The signal is then segmented into melody segments in which further iterative computations take place, estimating the pitch mean for the melody. The iterative process involves detecting and removing octave errors, removing pitch outliers (pitches far from the most recently found one) and filtering based on other aural characteristics.

Most algorithms submitted to MIREX evaluation favour the use of *Short Time Fourier Transform* (STFT) over more complex approaches to spectral analysis. In one of the variants submitted by Justin Salamon and Emilia Gómez a complex multi-resolution FFT (MRFFT) has been used with no significant improvement to the overall accuracy. Other approaches such as auditory correlograms or YIN frequency estimators had been used in the past but they have not been submitted to MIREX 2011.

Many of the presented algorithms rely heavily upon psychoacoustic features, including chroma characteristics, loudness equalisation, timbral distance measurement and noise models.

For selecting the final melody line based on F_0 candidates several methods are used, mostly being a form of melody tracking and pitch proximity calculations. The use of Hidden Markov Models and Support Vector Machines have been explored as well [44].

4.4 Multiple F_0 Estimation

The complexity of reliable fundamental frequency (F_0) estimation has been taken even further with the concept of multiple F_0 estimation. The algorithms for multiple F_0 estimation assume more than one harmonic source in the signal, usually accompanied by the noise signal as the residual to the sum of harmonic sources. Possible solutions to the problem of multiple F_0 estimation are fundamental to solving the problems of automatic music transcription, finding occurrences of similar patterns in different musical pieces and any other field of complex harmonic structure analysis. Estimation of multiple F_0s from the signal involves addressing several major issues.

4.4.1 Overlapping Partials

Multiple harmonic signals summed together interfere with each other. As all harmonic sources exhibit multiple harmonics in the spectrum, equally spaced in the frequency domain, they may overlap partially or completely. The fundamental frequencies of most musical notes are related, resulting in high probability of overlapping harmonics in the signal. Distinguishing the sound source for two notes distant from each other by an octave may be impossible, as the harmonics of these two notes produced by two distinct harmonic sources may overlap completely.

Octave, being the most obvious example, is not the only relation with high risk of overlapping harmonics.

4.4.2 Diverse Spectral Characteristics

Different musical instruments exhibit diverse spectral characteristics. Most of natural instruments, apart from clean harmonic structures, produce inharmonicities. For many musical instruments additional harmonics are present in the spectrum, mostly produced by the excitation of resonant body elements of the instruments or the strings. Some harmonics for certain instruments may be totally absent or weak enough to be completely overpowered by other instruments. The sound decay with time resulting in change in harmonic forms observed in the spectrum. The harmonic variations during note on-set (attack) increase the complexity as well.

4.4.3 Other Noise Sources

For signals that are difficult to analyse by themselves, any additional noise source or a source capable of changing the harmonic structures of the signal increases the complexity of F_0 estimation analysis dramatically. The most common source of the difference between the actual signal and the sum of single harmonic sources is the room acoustics. The recorded signal is a mixture of direct sounds, reflected ones and sounds coming from the reverberation. The reverberated parts are often non-stationary, resulting in even more increased complexity [45].

4.4.4 Evaluation

The task of multiple F_0 estimation has been covered in MIREX evaluations as well for many years now. One of the most successful approaches was proposed in Yeh et al. [45]. It involves the estimation of the noise level to distinguish the noise from the sinusoids in the observed spectrum. Taking the noise level into consideration, a~ combined approach of iterative F_0 estimation (in which consecutive F_0 candidates are estimated upon cancellation of prior findings in the spectrum) and joint estimation (in which several F_0 candidates are estimated at once from the observed partials) is used to determine harmonically related and unrelated F_0 candidates. A note contour tracking is used to follow full note trajectories instead of single (per-frame) candidates as a method of final estimation of the number of sources in the signal (*polyphony inference*).

4.5 Structural Segmentation

This relatively new task[1] aims at identification of the general structure of a song with respect to its main sections, such as e.g. verse or chorus and their potential subsections and repetitions. The music collection used in MIREX 2011 contest is composed of ca. 1,500 pieces representing a variety of musical styles.

4.5.1 Methods and Tools

There exists some similarity between this task and the cover song identification problem (Sect. 2.4), expressed in the level of description of the audio content. The proposed approaches are also often based on harmonic content characteristics, such as harmonic pitch class profile (HPCP) [46] or chroma-based features [47]. Some authors additionally combine these with timbral features, such as MFCCs, spectral flatness measure, etc. [47, 48].

The unique feature of the music structure discovery task is the need to build a~ hierarchical representation reflecting the correspondence between different parts of a given song. Some kind of tree structure is therefore constructed [46]. Several approaches including e.g. hierarchical agglomerative clustering have been proposed for this purpose [46–48].

5 Query-Based Music Identification

The ability to measure similarity between audio signals, classify them into certain categories and extract information that can be used to precisely create feature vectors that uniquely identify a piece can be used as a base for other group of tasks. These tasks can be summarized as query-based music identification, i.e. all processes that enable the listener to find a certain musical piece based on a limited information he/she possesses—for example a melody fragment or tappable rhythmical structure. In such tasks, at first feature vectors are extracted for a~ database of musical pieces to form a knowledge base for the algorithm. Then every time a listener asks the recommender system a question, for example providing a~ hummed melody, the same information is extracted from the provided query and matched against the database to find possible matches. The most popular groups of query-based music identification tasks include query by tapping (based on rhythmical searches) and query by humming (based solely on melody).

---

[1] introduced in MIREX 2009

5.1 *Query by Tapping*

This task is associated with a special input method for music searching in which a user specifies the rhythmic content of an audio file by e.g. tapping on a microphone or computer keyboard. The preprocessing of the submitted query consists basically of energy computation and thresholding. The audio collection to search through must be represented in symbolic form, such as e.g. MIDI format used in MIREX competition.

This task is somewhat similar to the Query by Humming problem (cf. Sect. 5.2) with the important difference that no pitch information is available here. Nevertheless, the same problems of tempo variations and missing notes may occur and the methods to solve them are also similar. In particular, given the note onsets and duration, approximate string matching algorithms [49, 50] or DTW-based dynamic programming approaches [51] may be applied for similarity assessment.

5.2 *Query by Humming (QbH) and Melody Similarity*

In the field of human-computer interaction the Query by Humming approach (also called Query by Singing/Humming, QbSH) lies definitely on the human side. This is natural. This is how we search for a piece of music not only in multimedia databases but also in the memory of our friends who may help us identify a melody stuck in our head.

The main problem in QbSH task is basically *melody matching* where the melody is understood as a sequence of notes with given pitches and durations. The user input is initially converted into a sequence of pitch values (a *pitch vector*), with a pitch detection algorithm (PDA) [52–54]. The potential problems involved here include the frequency resolution and precision of the PDA (usually of minor significance in the QbSH task), octave errors (may occasionally become an issue) and the imprecision of the sung query itself. The pitch vectors, once obtained, are processed with methods which can be roughly divided into note-based and direct approaches.

5.2.1 Note-Based Approaches

In these methods the onset times and durations of individual notes are determined. In this way a compact representation, allowing for efficient melody searching with string matching algorithms [55], may be obtained. The methods proposed here include edit distance computation based on note insertion/deletion/replacement cost [56], transportation distances such as the Earth Mover Distance (EMD) [57, 58] and n-grams matching [59]. The note-based approaches rely on the quality of the note segmentation stage which generally makes them potentially imprecise or dependent on the user adhering to a requirement of singing every note on a given syllable (e.g. “ta” or “da”, cf. [56]).

5.2.2 Direct Matching

In this group of methods the note segmentation problem is deliberately ignored and the pitch vectors are directly compared on a per-frame basis. High matching precision may be usually obtained in this way but at the cost of increased computational complexity [60, 61]. Not only is the melody representation much longer than the sequence-of-notes form, but also variations of tempo in the sung query make the standard Euclidean distance between vectors inaccurate. The query has to be aligned with the template via non-linear scaling of the time domain, which is usually done with an algorithm known as Dynamic Time Warping (DTW). Proposed initially for isolated words recognition [62, 63] the DTW has been widely adopted in many other fields of artificial intelligence and signal processing, including audio and video stream monitoring, bio-medical signal inspection, financial data analysis, human motion and gesture recognition [64, 65]. The variants of the method include full sequence matching [63] and subsequence matching [65, 66]. Efficient indexing techniques allowing to significantly reduce the searching time in large databases were introduced in Keogh [64] and applied in the Musical Information Retrieval context in Zhu and Shasha [61]. The application of several variants of the DTW algorithm for the QbSH problem has been addressed in numerous works, including [60, 66–69].

Besides the computational complexity, an important issue in DTW-based direct matching approaches is to obtain a key-invariant representation. The melody is defined by a sequence of relative pitches so their absolute values are basically irrelevant. The user can sing a melody in any key so all the notes may be shifted with respect to the template by the same interval which may result in a large value of the DTW distance, even for a perfectly sung query.

There exist several approaches to deal with this issue. Many researchers use a simple method of subtracting the mean pitch from the whole sequence [67]. The problem occurs when the melody represented in a query is only a part of the template (or vice versa) and it may have totally different mean value even if both happen to be in the same key.

In a different approach the melody may be represented in the form of relative changes of consecutive pitches (differential/delta representation) [70]. This eliminates the problem but representing raw pitch vectors in this form often yields poor results. The pitch vectors representing the templates, which are often in MIDI format, consist mostly of zeroes with non-zero values only at the points of note transitions. On the other hand, the note transitions in a query may be spread over several frames which makes the true comparison impossible.

An effective alternative may be to repeat the matching procedure several times with different transpositions of the query pitch vector. The query may be transposed by e.g. all possible number of semitones within the octave [69] or from -5

to +5 semitones in half-of-the-semitone steps [70]. Various numbers of repetitions may be considered but in any way this is a brute-force approach which increases the computational complexity significantly. Another problem, which is still not solved, is that the transposition may appear within the query when the user fails to sing an interval (usually a greater one) precisely and continues in a different key.

In the remaining part of this chapter a new solution of this problem is presented [71], in which the query is "tuned in" to the template via gradual decrease of the pitch difference between the two. This "tune-following" procedure was intended to resemble the way in which humans follow the known melody irrespective of pitch inaccuracies and key changes.

5.2.3 Melody Matching

The input query pitch vector q_{raw} obtained from audio data is first preprocessed in order to obtain a smooth melody line without large jumps and unvoiced fragments. The preprocessing includes the following steps [71]:

1. The leading and trailing unvoiced fragments are removed.
2. The median of the remaining data is computed and all the pitch values distant from the median by more than a given threshold T_1 are marked as unvoiced. This may help in case of poor quality of the input data resulting from noise or from errors introduced by the pitch detection algorithm.
3. For the same reason the maximum jump between two consecutive frames can not exceed a threshold T_2.
4. Every unvoiced frame is set to the pitch value of the last voiced frame. In this way one continuous melody is obtained, without any breaks resulting from breathing or articulation. It should be noted that this operation also leads to rejecting some potentially useful information about the rhythm and beat.
5. Median filter is applied to smooth the pitch contour. This operation typically leads to enhancement of the overall melody matching results.

The smoothed query pitch vector q is then compared with all the templates from the database. For every template t the pitch difference d_{beg} between the beginnings of the query and the template is computed and then subtracted from all the elements of the query pitch vector:

$$\underset{j=1,2,\ldots,J}{\forall} q_j := q_j - d_{beg}, \tag{8}$$

where J is the length of q after preprocessing and:

$$d_{beg} = \frac{q_2 + q_3}{2} - \frac{t_2 + t_3}{2}. \tag{9}$$

This operation makes both sequences start in the same key. As an assumption is made that the users sing queries from the beginning (which is quite a typical case), this procedure enables to obtain good matching results with the standard DTW

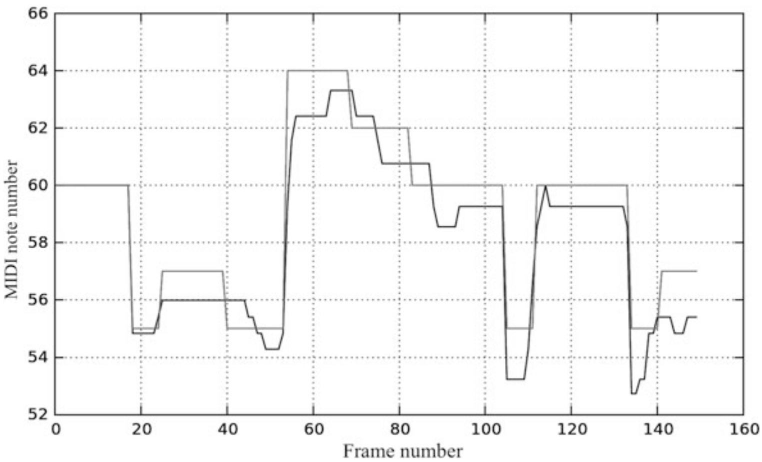


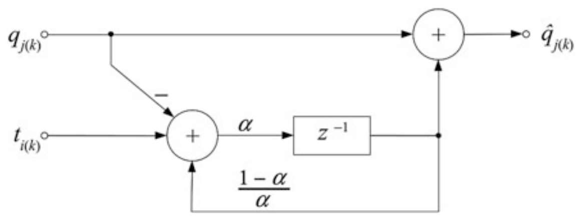
Fig. 2 Example of a transposition (*Old McDonald had a farm*). The first 17 samples of the template (*light*) and the median-filtered query (*dark*) are in tune. Most of the remaining part of the query is one–two semitones below the template [71]

algorithm. On the other hand, the queries often end in arbitrary positions with respect to the template sequences, e.g. due to pre-defined recording time, so using the arithmetic mean computed for *all* the values of q and t instead of d_{beg} in (8) usually yields poor results.

Having the beginning of the query shifted properly along the frequency axis, one have to deal with transpositions possibly occurring later in the course of the query (Fig. 2). For this purpose the standard DTW procedure is applied first to find the warping function aligning the query and the current template. The time-aligned query is then subjected to the procedure defined by the block diagram in Fig. 3. The resulting signal $\hat{q}$ is a version of q modified to follow the pitch values defined by the template t. This process is controlled by the parameter $\alpha \in (0, 1]$. The greater the value of α, the faster will the pitch of the query be aligned with the template.

The example with the same query and template sequences as in Fig. 2 is shown in Fig. 4. The enhancement introduced by the tune-following procedure is clearly visible. In most places the distance between the sequences decreased and two fragments in the second half of the query got tuned to the template exactly. It should be noted that both Figs. 2 and 4 present the aligned, i.e. time-warped version of the sequences.

Fig. 3 The block diagram of the tune follower [71]



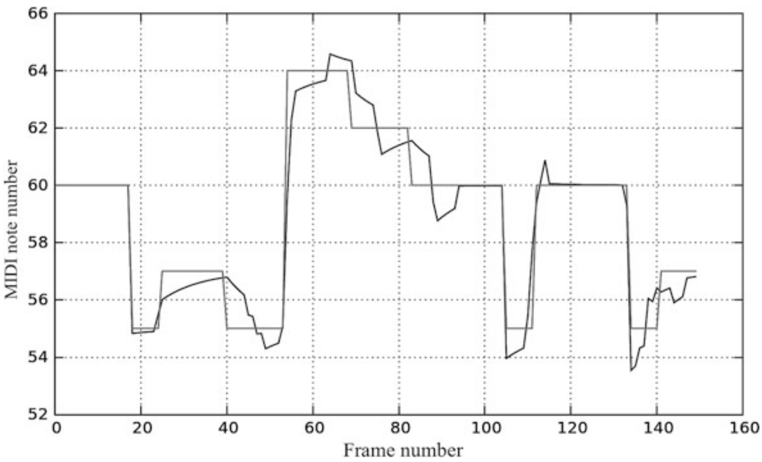


Fig. 4 The result of application of the tune follower ($\alpha = 0.1$) [71]

It should be noted, that although the query became more similar to the matching template, it can possibly also get closer to the non-matching ones. However, experimental validation [71] proved that the presented tune-following procedure indeed increases the recognition rate significantly. Although it is true that it generally makes most of the templates more similar to the query, it can be expected that this effect will be more significant in the case of the correct template than for all the non-matching ones.

This may result from the cumulation of the corrections for consecutive notes. For example, when the pitch of a note sung by a user is too low with respect to the correct template then it is gradually increased by the presented procedure until it reaches the right tune, provided that the note is long enough. If it is relatively short, it is at least partially corrected. In either case, if the note was sung too low, then it is probable that the pitch of the next note will also be too low in which case it will get corrected immediately or—at least—faster.

This type of correspondence between the signs of the pitch differences in consecutive notes cannot be generally expected when comparing a query with a non-matching template. Correcting one note may result in increasing the initial difference between the next note and the template. This may even result in making the query less similar to the template, although in case of long notes and infrequent pitch changes the tune follower will make the query closer to most of the templates.

6 Summary

In this chapter a survey of current techniques for music information retrieval, applicable to content-based recommender systems, has been presented. All the main, relevant research topics in the field have been covered along with current achievements, perspectives and known limitations.

The melody similarity problem has been discussed on the example of a Query by Humming user interface. It should be noted, however, that melody description and matching may be effectively applied also in broader context for music retrieval and recommendation systems, provided that reliable methods for automatic melody extraction from complex audio recordings are available. The latter problem—solved easily by a human listener—is still a difficult task for machines. In fact, despite the on-going progress in the areas of audio melody extraction and multiple F_0 estimation, this is still one of the open research issues in which science tries to keep up the pace with the amazing capabilities of the human mind.

7 Resource List

The sheer amount of algorithms and techniques that can be applied for tasks described in this chapter make it very difficult to conduct work without help of additional tools and already existing implementations. It is often unnecessary to prepare own software and tools for tasks that have been already thoroughly researched and give great results. We gathered a list of already available tools that can facilitate work with Music Information Retrieval related problems, as well as several valuable web resources where you can seek further information and help.

7.1 Software

7.1.1 Weka

Released under GNU General Public Licence, Weka is a Java free toolbox of machine learning algorithms. It can be used as a standalone application as well as a development library to use in own software. It contains various tools ranging from data pre-processing, classification, clustering, to as far as visualisation.

Available at http://www.cs.waikato.ac.nz/ml/weka/

7.1.2 Torch

Torch is a modular machine learning software library actively developed at Idiap Research Institute. It features routines facilitating for with MLPs, HMMs, Gaussian Mixtures and multitude of other learning algorithms. It can be also used as a base for creating own implementation. With very detailed documentation Torch is very well suited for rapid prototyping of various algorithms using machine learning algorithms.

Available at: http://www.torch.ch/

7.1.3 Marsyas

Unlike Torch and Weka, Marsyas is a software framework for rapid prototyping algorithms working with audio signals, with great emphasis on Music Information Retrieval. It also includes a wide collection of most published algorithms that can be used as building blocks for own experimentation. Marsyas is also well-suited for development of real time audio processing algorithms.

Available at: http://marsyasweb.appspot.com/

7.1.4 Praat

Praat is a software for speech analysis and synthesis, that can be used also for other Music Information Retrieval purposes due to a variety of flexible algorithms, such as pitch analysis, spectral analysis or segmentation and labelling. It also provides several features for data visualisation.

Available at: http://www.fon.hum.uva.nl/praat/

7.1.5 Audacity

Audacity is an open source audio editing software with several audio processing and analysis features. It also provides extensible plug-in framework to implement own features that can fit into the easy and quick to use graphical interface. Whereas it is not a software framework nor a tool specifically designed for Music Information Retrieval tasks, it can be very helpful for fast audio analysis and visualisation.

Available at: http://audacity.sourceforge.net/

7.2 Resources

7.2.1 EdaBoard

EdaBoard is one of the biggest and most respected internet forums for electronics. There is a very active digital signal processing forum on the site with extensive database of past discussions, as well as a community well versed in the matter of audio signal processing.

Available at: http://www.edaboard.com/forum40.html

7.2.2 comp.dsp

comp.dsp is one of the oldest discussion places for digital signal processing topics, including audio processing and Music Information Retrieval. It can be accessed through Usenet or Google Groups.

Available at: http://groups.google.com/group/comp.dsp

7.2.3 iMIRSEL

The website of International Music Information Retrieval Systems Evaluation Laboratory (IMIRSEL), who are behind MIREX evaluations, is a very valuable resource for seeking information about Music Information Retrieval and audio processing.

Available at: http://www.music-ir.org

References

1. Celma, O.: Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space. Springer (2010)
2. http://majorminer.org
3. Barrington, L., Oda, R., Lanckriet, G.: Smarter than genius? Human evaluation of music recommender systems. In: In Proceedings of the 10th International Conference on Music, Information Retrieval, pp. 31–36 (2009)
4. http://www.music-ir.org/mirex
5. http://www.ismir.net/
6. Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. Speech Audio Process. IEEE Trans. **10**(5), 293–302 (2002)
7. Stasiak, B., Yatsymirskyy, M.: Application of Fast Orthogonal Neural Network in content-based music genre classification. Pol. J. Environ. Stud. **17**(2A), 81–85 (2008)
8. Boersma, P., Weenink, D.: Praat, a system for doing phonetics by computer. Glot International **5**(9/10), 341–345 (2001)
9. Niewiadomy, D., Pelikant, A.: Digital speech signal parameterization by Mel Frequency Cepstral Coefficients and word boundaries. J. Appl. Comput. Sci. **15**(2), 71–81 (2007). http://edu.ics.p.lodz.pl/file.php/38/2-2007/niewiadomy.pdf
10. Hamel, Y.B.P., Lemieux, S., Eck, D.: Combining visual and acoustic features for music genre classification. In: Proceedings of ISMIR, 2011 (2011)
11. Seyerlehner, K., Widmer, G., Schedl, M., Knees, P.: Automatic music tag classification based on block-level features. In: Proceedings of the 7th Sound and Music Computing Conference (SMC 2010). Barcelona, Spain (2010)
12. Tardieu, D., Charbuillet, C., Cornu, F., Peeters, G.: MIREX-2011 Single-label and multi-label classification tasks: Ircamclassification2011 submission
13. http://www.last.fm/
14. Downie, J.S.: The music information retrieval evaluation exchange (2005–2007): A window into music information retrieval research. Acoust. Sci. Technol. **29**(4), 247–255 (2008)
15. Whitman, B., Flake, G., Lawrence, S.: Artist detection in music with Minnowmatch. In: Neural Networks for Signal Processing XI, 2001. Proceedings of the 2001 IEEE Signal Processing Society, Workshop, pp. 559–568 (2001)

16. Fujishima, T.: Realtime chord recognition of musical sound: A system using Common Lisp Music. In: Proceedings of the International Computer Music Conference, Beijing. International Computer Music Association (1999)
17. Lee, K.: Automatic chord recognition from audio using an HMM with supervised learning. In: Proceedings of ISMIR (2006)
18. Harte, C., Sandler, M., Gasser, M.: Detecting harmonic change in musical audio. In: Proceedings of Audio and Music Computing for Multimedia, Workshop (2006)
19. http://www.spotify.com/
20. http://www.grooveshark.com/
21. Seyerlehner, K., Schedl, M., Pohle, T., Knees, P.: MIREX-2010 Using block-level features for genre classification, tag classification and music similarity estimation
22. Bogdanov, D., Serra, J., Wack, N., Herrera, P.: MIREX-2010 Hybrid music similarity measure
23. Tzanetakis, G.: MIREX-2010 Marsyas submissions to MIREX 2010
24. Gkiokas, A., Katsouros, V., Carayannis, G.: MIREX-2010 Audio tempo extraction algorithm for MIREX 2010
25. Peeters, G., Papadopoulos, H.: Simultaneous beat and downbeat-tracking using a˜ probabilistic framework: Theory and large-scale evaluation. IEEE Trans. Audio Speech Lang. Process. **19**(6) (2011)
26. http://www.music-ir.org/mirex/abstracts/2011/KFRO1.pdf
27. Hainsworth, S., Malcolm, M.: Beat tracking with particle filtering algorithms. In: Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on. IEEE (2003)
28. Cemgil, A.T., Kappen, H.J.: Monte carlo methods for tempo tracking and rhythm quantization. J. Artif. Intell. Res. **18**, 45–81 (2003)
29. Scheirer, E.D.: Tempo and beat analysis of acoustic musical signals. J. Acoust. Soc. Am. **103**(1), 588–601 (1998)
30. Laroche, J.: Estimating tempo, swing and beat locations in audio recordings. In: Proceedings of the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'01), pp. 131–135 (2001)
31. Klapuri, A.P., Eronen, A.J., Astola, J.T.: Analysis of the meter of acoustic musical signals. In: IEEE Transactions on Speech and Audio Processing, pp. 342–355 (2004)
32. Peeters, G.: Template-based estimation of time-varying tempo. EURASIP J. Adv. Sig. Proc. (2007)
33. Schoenberg, A.: Theory of Harmony. University of California Press, California (1978)
34. Gomez, E.: Tonal description of music audio signals. PhD thesis, Universitat Pompeu Fabra (2006)
35. Krumhansl, C.: Cognitive foundations of musical pitch. Oxford University Press, USA (1990)
36. Temperley, D.: What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered, pp. 65–100. University of California, California (1999)
37. Ni, Y., Mcvicar, M., Santos-Rodriguez, R., Bie, T.D.: MIREX-2011 Harmony progression analyzer for MIREX 2011
38. Khadkevich, M., Omologo, M.: MIREX-2011 Audio chord detection
39. Cho, T., Bello, J.P.: MIREX-2011 A feature smoothing method for chord recognition using recurrence plots
40. Salomon, J., Gomez, E.: Melody extraction from polyphonic music: MIREX 2011
41. Eokhwan Jo, S.P.S., Yoo, C.D.: Melody extraction from polyphonic audio signal MIREX 2011
42. Liao, W.H., Su, A.W.Y., Yeh, C., Roebel, A.: Melody estimation for MIREX 2011
43. Zwicker, E.: Subdivision of the audible frequency range into critical bands. J. Acoust. Soc. Am. (33) (1961)
44. Vicent, E., Plumbley, M.: Predominant-f0 estimation using Bayesian harmonic waveform models. In: 2005 Music Information Retrieval Evaluation eXchange (MIREX) (2005)

45. Yeh, C., Roebel, A., Rodet, X.: Multiple fundamental frequency estimation and polyphony inference of polyphonic music signals. Trans. Audio Speech Lang. Proc. **18**(6), 1116–1126 (2010). doi: 10.1109/TASL.2009.2030006
46. Martin, B., Hanna, P., Robine, M., Ferraro, P.: Structural analysis of harmonic features using string matching techniques
47. Chen, R., Li, M.: Music structural segmentation by combining harmonic and timbral information. In: Klapuri, A., Leider, C. (eds.) ISMIR, pp. 477–482 (2011)
48. Peeters, G.: Sequence representation of music structure using higher-order similarity matrix and maximum-likelihood approach. In: Dixon, S., Bainbridge, D., Typke, R. (eds.) ISMIR, pp. 35–40. Austrian Computer Society (2007)
49. Hanna, P., Robine, M.: Query by tapping system based on alignment algorithm. In: ICASSP, pp. 1881–1884. IEEE (2009)
50. Smith, T.F., Waterman, M.S.: Identification of common molecular subsequences. J. Mol. Biol. **147**(1), 195–197 (1981)
51. Chen, C.T., Jang, J.S.R.: Query by tapping
52. Boersma, P.: Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In: Proceedings of the Institute of Phonetic Sciences, vol. 17 (1993)
53. Dziubiński, M., Kostek, B.: High accuracy and octave error immune pitch detection algorithms. Arch. Acoust. **29**(1) 1–21 (2004)
54. Gerhard, D.: Pitch Extraction and Fundamental Frequency: History and Current Techniques. Department of Computer Science, University of Regina, Regina (2003)
55. Ghias, A., Logan, J., Chamberlin, D., Smith, B.C.: Query by humming musical information retrieval in an audio database. In: Proceedings of the Third ACM International Conference on Multimedia. ACM (1995)
56. McNab, R.J., Smith, L.A., Witten, I.H., Henderson, C.L., Cunningham, S.J.: Towards the digital music library: Tune retrieval from acoustic input. In: Proceedings of the First ACM International Conference on Digital Libraries. ACM (1996)
57. Huang, S., Wang, L., Hu, S., Jiang, H., Xu, B.: Query by humming via multiscale transportation distance in random query occurrence context. In: Multimedia and Expo, 2008 IEEE International Conference on. IEEE (2008)
58. Typke, R., Wiering, F., Veltkamp, R.C.: Transportation distances and human perception of melodic similarity. Musicae Scientiae **11**(1), 153–181 (2007)
59. Uitdenbogerd, A., Zobel, J.: Melodic matching techniques for large databases. In: Proceedings of the Seventh ACM International Conference on Multimedia (Part 1). ACM (1999)
60. Jang, J.S., Lee, H.R.: Hierarchical filtering method for content-based music retrieval via acoustic input. In: Proceedings of the Ninth ACM International Conference on Multimedia. ACM (2001)
61. Zhu, Y., Shasha, D.: Warping indexes with envelope transforms for query by humming. In: Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data. ACM (2003)
62. Itakura, F.: Minimum prediction residual principle applied to speech recognition. Acoust. Speech Sig. Process. IEEE Trans. on **23**(1), 67–72 (1975)
63. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. Acoust. Speech Sig. Process. IEEE Trans. on **26**(1), 43–49 (1978)
64. Keogh, E.: Exact indexing of dynamic time warping. Knowl. Inf. Syst. **7**(3), 358–386 (2005)
65. Sakurai, Y., Faloutsos, C., Yamamuro, M.: Stream monitoring under the time warping distance. In: Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on. IEEE (2007)
66. Lijffijt, J., Papapetrou, P., Hollmen, J., Athitsos, V.: Benchmarking dynamic time warping for music retrieval. In: Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments. ACM (2010)

67. Jeon, W., Ma, C.: Efficient search of music pitch contours using wavelet transforms and segmented dynamic time warping. In: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE (2011)
68. Wang, L., Huang, S., Hu, S., Laing, J., Xu, B.: An effective and efficient method for query by humming system based on multi-similarity measurement fusion. In: Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on. IEEE (2008)
69. Yu, H.M., Tsai, W.H., Wang, H.M.: A query-by-singing system for retrieving karaoke music. Multimedia IEEE Trans. on. **10**(8), 1626–1637 (2008)
70. http://www.music-ir.org/mirex/abstracts/2011/jsslp1.pdf
71. Stasiak, B.: Follow that tune-dynamic time warping refinement for query by humming. In: Proceedings of Joint Conference on New Trends in Audio and Video Signal Processing: Algorithms, Architectures, Arrangements, and Applications (NTAV/SPA), pp. 109–114 (2012)

Composition Support of Presentation Slides Based on Transformation of Semantic Relationships into Layout Structure

Toyohide Watanabe and Koichi Hanaue

Abstract This paper presents a framework of supporting composition of presentation slides in presentation preparation.Traditional presentation tools do not always make it easy for a presenter to compose a deck of presentation slides that convey clear message and consistent story. This is because the tools do not allow a presenter to externalize the semantic connectivity among the concepts and messages. For a persuasive presentation, a presenter should clarify what to speak and how to speak in the form of a scenario. Therefore, we address this problem based on the concept of a presentation scenario. Our idea is to develop an interface for externalizing a presentation scenario and then to generate presentation slides based on the scenario. In our framework, a presentation scenario is represented as the network structure in which the slide components such as texts and images are connected with semantic relationships. In composing a presentation scenario, the view of the whole presentation story helps a presenter to examine the consistency among his/her concepts and messages. Therefore, we introduce the technique of zooming to our interface for scenario composition. In composing a presentation slide, the slide should reflect the semantic relationships between slide components. To achieve this, we take an approach of transforming the semantic relationships between the slide components into diagrammatic representation. We define typical patterns of diagrammatic representation as layout templates. The layout templates consists of the precondition and the layout structure. The appropriate templates are selected and applied to each topic in the scenario by comparing their preconditions with the semantic relationships. We present the implementation of our prototype system and the examples of the slides generated by our system.

T. Watanabe (✉) · K. Hanaue
Graduate School of Information Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
e-mail: watanabe@nagoya-u.jp

K. Hanaue
e-mail: hanaue@nagoya-u.jp

G. A. Tsihrintzis et al. (eds.), *Multimedia Services in Intelligent Environments*,
Smart Innovation, Systems and Technologies 25, DOI: 10.1007/978-3-319-00375-7_10,

1 Introduction

Presentation is one of the most important means for sharing and transferring knowledge. Today, presentations are shared in various forms through the services on the Web. For example, TED provides a video archive of past talks in TED conferences around the world [1]. Also, the live streaming service by USTREAM enables us to watch live presentations at a distant place [2]. In addition, the slides used in oral presentations are shared on the Web. Presentation slides play an important role to promote understanding of what presenters talk about. Most of the slides are prepared by using presentation tools such as Microsoft PowerPoint [3], Apple Keynote [4] and OpenOffice Impress [5]. SlideShare is one of the famous Web services for browsing, sharing presentation slides [6]. Prezi also provides a service for browsing and sharing presentation data [7]. Prezi expresses the topic transition during a presentation using the techniques of zooming and panning on a two-dimensional plane.

Although these tools enable us to make presentation slides easily with rich graphics and animations, they have some problems from the viewpoint of understandability about slides. First, presenters tend to focus on making slides rather than considering their messages and stories. Second, most presenters use the formats of slides set up by presentation tools, and thereby rely on the way of phrase headlines with bullet lists. The slides that have such formats tend to make the discussion point unclear [8]. As many criticisms such as Tufte [9] have pointed out, it is not easy to complement the semantic gaps among slides in one presentation or even components such as texts and images in one slide. One of the reasons is that the traditional presentation tools do not have the function for specifying semantic relationships among slides and components explicitly. Another reason is that the tools do not have a function for looking at the overview of a presentation story. This prevents presenters from comparing multiple slides and thereby checking if the presentation story is consistent during the process of preparation. A presenter has to focus on clarifying his/her message and presentation story so that his/her audience can understand them easily. Therefore, a presentation tool must allow him/her to externalize and edit them in order to make the process of presentation preparation effective.

In order to address this problem, we have prototyped a presentation authoring system that handles a scenario for presentations [10, 11]. Our prototype system has the characteristic functions as follows:

- Externalizing ideas of a presenter in the form of a slide component such as a text and an image,
- Organizing the externalized components using semantic relationships, and
- Generating slides automatically for each group of components.

A scenario represents the intention of a presenter on what to speak and how to convey it in a presentation. In the process of composing a scenario, representation of externalized intentions affects the cognitive process of a presenter [12]. Also, such representation should be edited and manipulated intuitively [13]. Therefore, spatial arrangement of idea fragments are effective to support scenario composition. In addition, it is necessary for a presenter to specify semantic relationships between ideas. Therefore, we represent a presentation scenario as a network structure in which a node and a link correspond to a slide component and a semantic relationship between components, respectively. By allowing a presenter to specify relationships between components, it becomes possible to capture and handle a presenter's intention on his/her presentation.

In our framework, a presenter organizes each topic by specifying semantic relationships between slide components. Here, a slide component is an object such as a text and an image that forms a part of a presentation slide. Then, slides are generated for each topic in the scenario based on the semantic relationships. Our idea is to apply appropriate layout templates to each topic. A layout template is a typical pattern of allocating slide components for diagrammatic representation. The template consists of the precondition for applying it to the semantic relationships and the allocation procedure of slide components. In our method, the appropriate layout template is selected by comparing the semantic relationships in a topic with the precondition of the template. Then, slides are generated according to the allocation procedure specified in the template. By achieving this, it becomes possible to generate slides with diagrammatic representation as prepared in SmartArts in PowerPoint.

The rest of this chapter is organized as follows. Section 2 presents existing work that is related to supporting the process of presentation preparation. Section 3 describes our framework for supporting presentation preparation based on the concept of a scenario. Section 4 gives the definition of slide components and semantic relationships among them. Section 5 describes the mechanism of generating presentation slides from semantic relationships. Section 6 shows our prototype system and the examples of generated slides. Section 7 explains the comparison of our system with the existing systems. Finally, Sect. 8 concludes this chapter and presents future directions.

2 Related Work

Until now, many systems have been developed for addressing the problems of the traditional presentation tools. We focus on three aspects of such systems: authoring support of semantic contents, user interfaces for presentation composition, and automatic generation of slides.

2.1 Authoring Support of Semantic Contents

A number of systems have been developed in order to support the process of externalizing and organizing pieces of ideas. Some systems are based on the concept of semantic authoring [14] and others are based on the technique of idea organization called KJ method [15, 16]. In KJ method, participants write their ideas on many pieces of paper, and then organize them through spatial arrangement and hierarchical grouping. Then, participants illustrate diagrams or passages about the ideas based on the result of the previous steps.

In order to support idea generation in KJ method, KJ Editor [17] and D-ABDUCTOR [18] have the functions for editing, manipulating and grouping idea fragments semi-automatically on a computer. VITE also provides a user function to intuitively arrange and manipulate structured data [19]. In VITE, the data set is represented as a set of small pieces of hypertext. In addition, a user can express relation between data items by visually configuring them on his/her own way. iMapping is a system for organizing and managing personal semantic knowledge [20]. iMapping is developed based on the approach of combining visualization techniques and semantic technologies. The user of the system is able to specify the semantic relation in his/her knowledge explicitly with a zooming interface. Although these systems are not used for generating slides automatically, their output can be used for supporting composition of presentation slides.

2.2 User Interfaces for Presentation Composition

Some systems enable a presenter to prepare presentation contents in order to perform a presentation flexibly by switching the sequence of slides during his/her presentation intuitively. Palette is a presentation tool that has a paper-based interface [21]. In Palette, a presenter can control a slide show by using index cards printed with slide contents. Customizable Presentation provides functions of grouping slides and specifying multiple paths on slide transition [22].

Some systems adopt zooming user interfaces (ZUIs) in order to facilitate a grasp of topic structure and stories by both presenters and audience. The basic approach is to add the functionality of zooming to the canvas on which presentation materials are composed and presented. Both a presenter and audience can focus on a specific topic and look over the whole story by zooming in and out. CounterPoint has a function of arranging slides prepared with PowerPoint on a two-dimensional plane and expressing topic transition by animations such as zooming and panning [23]. FLY [24] and Prezi [7] also provide ZUIs and specify views of presentation data by taking “snapshots” on a two-dimensional plane at arbitrary location and scale. These systems make it easy for a presenter to present the overview of a story for audience. On the other hand, the interrelationships among topics and components are not always clear because they cannot be specified explicitly.

From the viewpoint of reusing existing slides, Drucker et al. developed a system to support assembling slides for a new presentation from multiple versions of existing slide decks [25]. Their system provides an interactive interface that visualizes multiple presentations with their slides aligned based on the similarities among the slides. Outline Wizard is also a system to support composing presentation slides using the technique of outline-based search [26]. Outline Wizard has a repository of existing slides with its outline extracted from their hierarchical structure. In composing presentation, the slides are retrieved from the repository so as to match the the outline specified by a presenter. These systems make it easy to compose a presentation from existing slides. On the other hand, these systems do not support the process of considering a story in order to make it persuasive and consistent. That is because they do not handle the semantic relationships between contents or topics in a presentation.

2.3 Automatic Generation of Presentation Slides

Some systems generate slides automatically from existing resources such as paper manuscripts using techniques of text summarization. Miyamoto et al. proposed a method for generating slides from a research paper in TeX format [27]. In this method, sentences that have parallel relations are extracted for generating slides in an itemized format using conjunctive words and phrases. Shibata et al. proposed a method for generating slides by constructing discourse structure from a text segment [28]. In this method, a clause and a sentence are considered as a unit that forms a discourse structure. The discourse structure is constructed by detecting coherence relations between clauses and sentences using cue phrases, word/phrase chain, and similarity between sentences. The coherence relations are used for computing the depth of indentation in a slide for each clause and sentence. Although these methods generate slides based on semantic relationships, the slides follow the standard format of bullet-point lists in traditional presentation tools. Utiyama et al. proposed a method for generating slides from a text segment with annotations called the GDA tagset [29]. In this method, parse-tree bracketing (morphological and syntactic information of words and phrases), semantic relation, and coreference in the tagset are considered for generating slides. The semantic relation includes rhetorical relation such as cause, concession and elaboration. Since this method also generates slides in the format of bullet-point lists, it is difficult to allocate components such as texts and images so as to represent the discourse structure effectively on slides.

Yasumura, et al. proposed a method for supporting composition of slides in XHTML format from a research paper manuscript written in TeX format [30]. Their system takes a paper manuscript as an input and computes the number of slides for each section in the paper, then selects the appropriate layout template for each topic such as a section and a subsection. The layout templates are applied to specific type of a topic and specific combination of components. This system extracts the components such as texts, images and mathematical expressions by computing their importance. On the other hand, it does not extract the relationships among the components. Therefore, the expressiveness of semantic relationships in a slide is limited.

In summary, existing works on generating slides do not prepare the layout templates of slides in which the semantic relationships among slide components are embedded explicitly. The layout templates in these works are prepared based on the types of components to be included in slides (e.g., the template for a bullet list and the template with a graph). Therefore, the generated slides do not always convey the interrelationships among concepts in slides to the audience. In order to address this problem, we define a template with its inherent semantic construction. By considering semantic relationships as preconditions for applying a template, we attempt to transform the semantic relationships among slide components into the slide with an appropriate layout.

3 Framework

3.1 Process of Presentation Preparation

We consider the process of presentation preparation composed of three phases as illustrated in Fig. 1. First, a presenter externalizes ideas for a presentation and organizes them to construct a consistent story. Second, a presenter groups the ideas to form a set of topics and composes a blueprint for presentation slides. Finally, a presenter makes a sequence of slides by allocating slide components on slide sheets for each topic.

Based on the process described above, we take the following approaches to support the process of presentation preparation:

- Providing a user interface with which a presenter externalizes and organizes his/her ideas intuitively (first and second phases),
- Generating presentation slides automatically from the structured ideas externalized by a presenter (third phase).

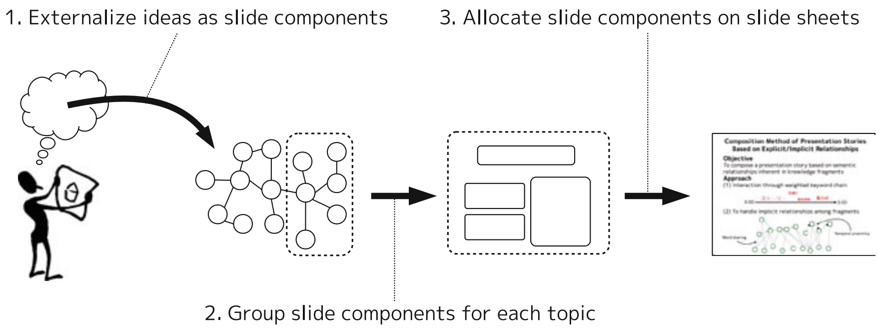


Fig. 1 Process of presentation preparation

3.2 Representation of Presentation Scenario

Traditional presentation tools are not suitable for a presenter to organize ideas and to construct a story in the process of preparation. This is because the thought of a presenter on what and how to speak is not always clear in an early stage of presentation preparation. Since a presentation scenario is composed through a process of structuring presenter's idea fragments, it is quite difficult for him/her to express his/her thought in the form of presentation slides. Also, it is more important for a presenter to consider semantic relation among topics or among items within each topic than to consider the appearance of presentation slides and method for visualizing his/her thought. However, a presenter cannot specify such relationships on traditional presentation tools. This makes it difficult for a presenter to confirm if topics or items are related to each other appropriately.

In order to solve the problem in an early stage of presentation preparation, we introduce a concept of a presentation scenario. A presentation scenario is the externalized representation of a presenter's intentions on his/her presentation. Figure 2 illustrates the representation of a presentation scenario in our framework. This representation has the following characteristics:

- A scenario consists of semantic blocks and is free from physical constraints.
- The semantic blocks are partially ordered, but not necessarily linearly ordered.
- In each topic, the concepts to be mentioned are associated with each other using semantic relationships.

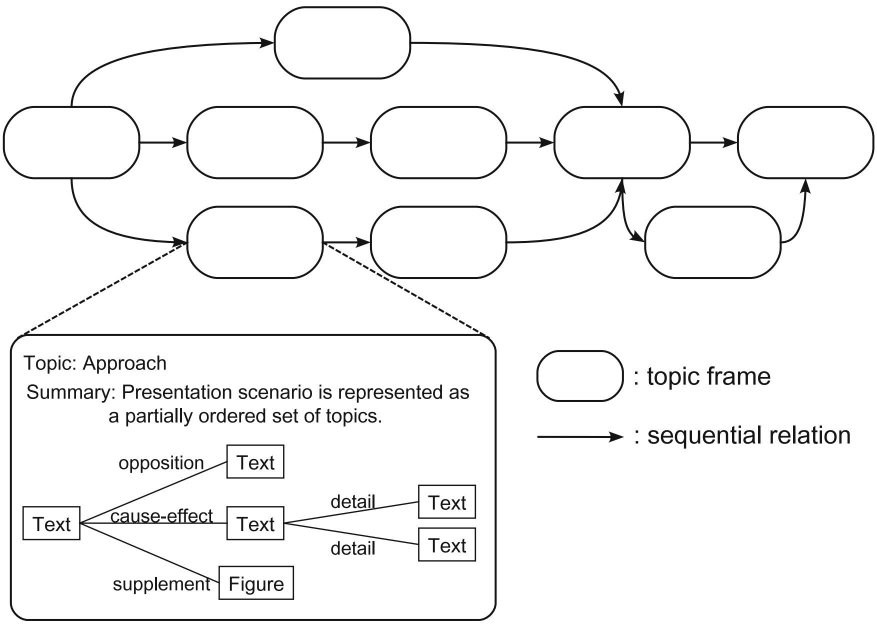


Fig. 2 Representation of presentation scenario

We call a semantic block in a scenario a *topic frame*. A topic frame represents a topic such as a chapter and a section in documents. Topic frames are free from the physical constraints such as the sizes of slides or displays. In each topic frame, slide components such as texts and images are organized. We assume that each text component contains a clause or a sentence written in a natural language. In the organization process, the components are connected by semantic relationships such as "cause and effect" or "an example of." Also, a presenter can specify a relationship between topics by determining the component shared by them. Semantic relationships among topics and components are regarded as implicit knowledge on how to proceed a presentation. Therefore, a presenter is able to externalize such implicit knowledge in composing a scenario.

3.3 Generating Presentation Slides Using Layout Templates

We take an approach of generating slides from the topic frames described above. In order to reflect the semantic relationships in generated slides, we attempt to arrange slide components diagrammatically on a slide. Presentation slides must include visual components such as images, graphs, diagrams and tables in order to promote understanding of what a presenter talks about. That is because the slides with these visual components convey the message of a presenter intuitively and the relationships among his/her ideas effectively. In addition, the layout of components on a slide must not be too complex [31]. Therefore, the diagrammatic representation of generated slides should be as complex as that of SmartArts in PowerPoint.

In order to reflect the structure of a topic frame in the layout on slides, we define a layout template using the semantic relationships among slide components. Figure 3 illustrates our approach. A layout template of a presentation slide is a characteristic pattern of allocating slide components. We attempt to find an appropriate template for a given topic frame. To achieve this, we define the precondition for applying a layout template as the logical structure of slides. Generally, document structure is defined as the combination of a logical structure and a geometric structure [32]. A logical structure represents semantic relationships among document elements such as the title, texts, subtexts, figures, tables and footnotes. The elements have their own physical properties such as size, length, location, font style and size in texts. The semantic relationships are assigned among two or more elements: is-a, part-of, grouping, etc. A geometric structure is related to the physical positions of individual elements from a location-oriented point of view on a sheet: neighboring, left-right, upper-lower, overlap, etc. Thus, it is important to transform the elements from a logical structure to a geometric structure. The transformation is dependent on the semantic relationships among the elements in the logical structure and the physical properties of elements, and is performed interpretatively from the upper elements in the logical structure, which is specified commonly by a tree structure.

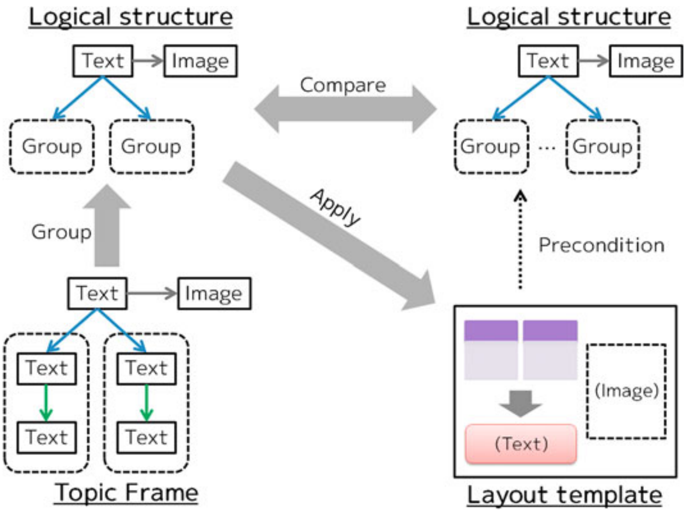


Fig. 3 Transformation from a topic frame into a slide

In our method, layout templates are defined as a pair of a logical structure and a geometric structure. The layout templates are used for transforming the logical structure into the geometric structure. To do this, we first derive the logical structure of slides for each topic frame. Then, we find appropriate templates by comparing the logical structure derived from a topic frame and those defined as preconditions in templates. Finally, we generate slides by allocating slide components according to the templates. In our current method, the slides are generated for each topic. This means that the relationships among topics in the scenario are not considered in generating slides. Of course, there is a case in which it is necessary to generate a slide that summarizes multiple topics. However, we focus on the generation mechanism of a slide from one topic and leave the mechanism considering relationships among topics for our future work.

4 Composition of Presentation Scenario

4.1 Data Structure of Presentation Scenario

Our presentation scenario model is illustrated in Fig. 4. In our model, a presentation scenario is represented as a three-layered structure: a skeleton layer, a topic frame layer, and a component layer. In Fig. 4, circles in a skeleton layer and a topic frame layer denotes topic frames. Rectangles in a topic frame layer and a component layer denote slide components. The colored rectangle represents the component that corresponds to the most important point in each topic. In our model, a presentation scenario is defined as follows:

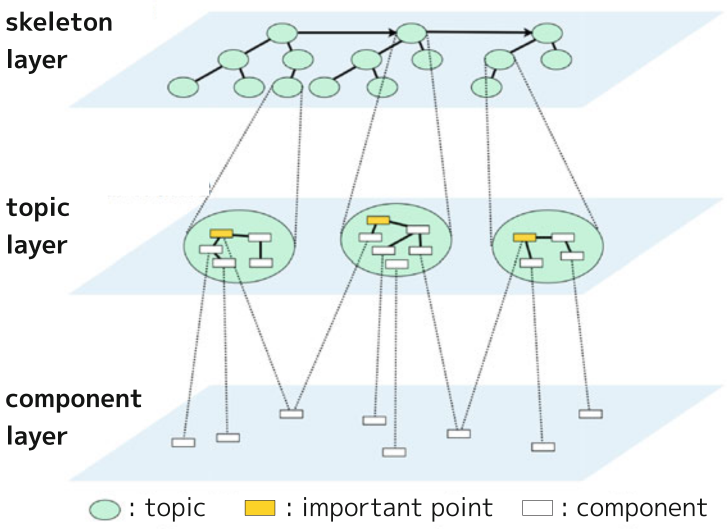


Fig. 4 Three-layered model of presentation scenario

$$\text{Scenario} = (\text{Skeleton},\ \text{Topics},\ \text{Components}).$$

A skeleton is a graph structure of topic frames that constitute a presentation story. The data structure in a skeleton layer is specified as follows:

$$\begin{aligned} \text{Skeleton} &= (\text{V}_S,\ \text{E}_S),\\ \text{V}_S &= \{f | f \in \ \text{Topics}\},\\ \text{E}_S &= \{(f_{src}, f_{dst}, type) | f_{src}, f_{dst} \in \ \text{V}_S\}. \end{aligned}$$

In these expressions, f_{src} and f_{dst} are references to topic frames in a topic frame layer. *Type* represents a type of relationship between topic frames.

In a topic frame layer, topic frames are stored. A topic frame is composed of its name, and the slide components and semantic relationships among them. The data structure in a topic frame layer is defined as follows:

$$\begin{aligned} \text{Topics} &= \{f\},\\ f &= (id, name,\ \text{C},\ \text{L}),\\ \text{C} &= \{(c, role) | c \in \ \text{Components}\},\\ \text{L} &= \{(c_{src}, c_{dst}, type) | c_{src}, c_{dst} \in \ \text{C}\},\\ role &\in \{title, point, normal\}. \end{aligned}$$

In each topic frame, *id* represents its identifier. C is a set of references to slide components in a component layer. The *role* of a component represents its role in a topic frame. The component whose *role* is *title* corresponds to the title of a topic.

The component whose *role* is *point* represents the important point in a topic. This attribute is reflected in the logical structure of slides. L is a set of links between components included in the topic frame f. *Type* represents a type of semantic relationship between components. We define the types of semantic relationship as shown in Table 1 by referring to the Rhetorical Structure Theory [33]. The types of relationships are categorized from two viewpoints. One is whether the relationship has direction. The components with such relationship have constraints on sequence, dependency, etc. If *type* in the above definition is directed, the relationship has a direction from component c_{src} to component c_{dst}. The other is whether the relationship is hierarchical. The components related by such relationship have difference in their importance in presentation. If *type* in the above definition is hierarchical, component c_{src} is considered to be more important than c_{dst}.

In a component layer, slide components are stored. The data structure in this layer is defined as follows:

$$\begin{aligned}\text{Components} &= \{c\},\\ c &= (id, content, type, size),\\ type &\in \{text, image\}.\end{aligned}$$

In these expressions, *id* represents an identifier of a component. *Content* and *type* represent the component itself and its type specified by *id*, respectively. If the *type* of a component is *text*, its *content* is a character string. If the *type* of a component is *image*, its *content* represents a corresponding image. In our definition, *type* takes one of the only two values above. This is because visual components such as an image, a table, a graph and a diagram are similar in the sense that they occupy a larger region than a text component. Another reason is that we focus on the relationships among components rather than the types of the components. The *size* of a component represents its horizontal length and vertical length. The value of *size* is assigned to the components whose *type* is *image*.

Table 1 Types of relationships between slide components

Type	Directed	Hierarchical
caused-by	yes	yes
assumed-by	yes	yes
opposes	yes	no
compared-with	no	no
exemplified-by	yes	yes
detailed-by	yes	yes
specialized-by	yes	yes
supplemented-by	yes	yes
illustrated-by	yes	yes
followed-by	yes	no
related-to	no	no

4.2 Operations for Scenario Composition

The composition process of presentation scenario is divided into two phases: topic frame construction and skeleton construction.

A topic frame is constructed by the following operations:

- Create a new topic frame,
- Add/Remove components to/from topic frames,
- Assign roles to components,
- Connect two components within a topic frame.

The operation of connecting components is realized by specifying the type of semantic relationship between them. Since each topic must have a message to be emphasized, we introduce a constraint that the role *point* must be assigned to at least one of the slide components for each topic frame.

The skeleton of a presentation scenario is constructed by connecting topic frames. In this phase, topic frames are connected to construct a partially ordered set of topics. In our framework, two topic frames are connected if they share one or more components. That is because the topics that are related to each other should share some concepts.

As we discussed in Sect. 3, a topic frame represents a semantic block such as a chapter, a section, etc. In books or research papers, there are two types of relations between chapters or sections: sequential relation (predecessor-successor) and hierarchical relation (section-subsection). Therefore, the type of relationships between topic frames is either sequential or hierarchical.

In order to realize the operations described above, we visualize the structure of a presentation scenario in the form of a network. In the level of a skeleton, a topic frame is represented as a node and the relationships between topics are represented as a link. In the level of a topic frame, a slide component is represented as a node and the semantic relationships between them are represented as a link. In both levels, whether a link is directed is determined according to the type of a relationship.

5 Automatic Generation of Presentation Slides

5.1 Processing Flow

From the presentation scenario externalized by a presenter, presentation slides are generated based on the semantic relationships between slide components. As we explained in Sect. 3.3, we currently focus on the generation mechanism of slides from one topic frame without considering the relationships among topic frames.

The basic mechanism of generating slides is the mapping from a logical structure to a geometric structure based on the semantic relationships. The processing flow consists of two steps: grouping and allocation. First, for each topic frame in a scenario, a logical structure is derived by grouping slide components according to the semantic relationships. The logical structure represents inclusive relationships and sequential relationships among groups of slide components. Next, appropriate templates are selected by comparing the logical structure with that of the templates. Finally, the components are allocated according to the geometric structure defined in the selected template. If part of the components cannot be allocated on a slide sheet, other templates are applied to remaining part of the logical structure. These steps are repeated until all the components in a scenario are allocated on slide sheets.

5.2 Grouping Slide Components

The topic frame constructed by a presenter is taken as an input for generating presentation slides. In our method, slides are generated by finding and applying appropriate templates for diagrammatic representation. In order to apply templates to the topic frames, the logical structure of a presentation slide is constructed. This is achieved by grouping the components based on the semantic relationships among them.

Given a topic frame $f = (id_f, name_f,\ \mathrm{C}_f,\ \mathrm{L}_f)$, the logical structure LS_f is defined as follows:

$$LS_f = \{(c, S_c) | c \in \ \mathrm{C}_f\},$$
$$S_c = \{(r, c', S_{c'}) | (c, c', r) \in \ \mathrm{L}_f \wedge (c', S_{c'}) \in LS_f\}.$$

Here, S_c is a set of triples $(r, c', S_{c'})$. r is a type of relation and $S_{c'}$ is a set of components that are related to c by the relationship r. $S_{c'}$ is a subgroup of S_c. Computing S_c means grouping components related to c according to the types of relationship. Figure 5 illustrates an example of a logical structure constructed from a topic frame. The rectangles with dotted lines represent groups in a logical structure.

As described in Sect. 4.2, each topic frame has at least one component whose role is *point*. Therefore, we start the process of grouping from such components and group slide components recursively according to the breadth-first policy of visiting nodes in a network. During this process, the relationships between groups of components are determined in order to derive a logical structure. Based on the categories of the semantic relationships in Table 1, we introduce three types of relationships between groups as follows:

Hierarchical relationship If $(c, c', r) \in R$ and r is a hierarchical relationship, $S_{c'}$ becomes the subset of S_c. Namely, $S_{c'}$ is included in S_c

Directed relationship If $(c, c', r) \in R$ and r is a directed relationship, $S_{c'}$ becomes the sibling of S_c. Namely, a directed relation exists from S_c to $S_{c'}$ and these sets are

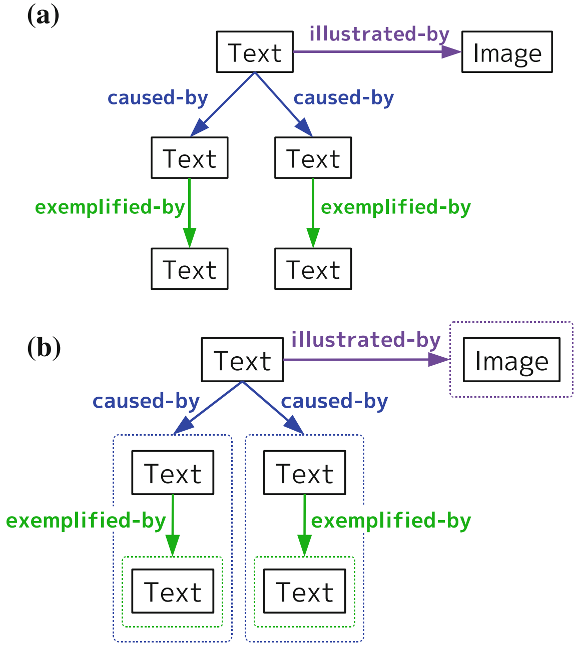


Fig. 5 An example of constructing logical structure. **a** Input topic frame. **b** Constructed logical structure

included in the same set S_{c_0}. Here, c_0 satisfies $(c_0, c, r_0) \in R$ and r_0 is a hierarchical relationship.

Parallel relationship If $(c, c', r) \in R$ but r is not directed or hierarchical, $S_{c'}$ is considered to be independent of S_c. Namely, S_c and $S_{c'}$ are disjoint with each other and included in the same set S_{c_0}. Here, c_0 satisfies $(c_0, c, r_0) \in R$ and r_0 is a hierarchical relationship.

5.3 Allocating Slide Components on Slides

A layout template is a characteristic pattern of allocating slide components that is practically used for visualizing a semantic structure within a topic. We define a layout template as a pair of a logical structure and a geometric structure as illustrated in Fig. 3. A logical structure corresponds to the preconditions for applying a template. A geometric structure contains relative positions among slide components and additional shape objects such as lines and arrows. In our method, a layout template is selected by comparing its precondition with the logical structure constructed from a topic frame.

We take the logical structure LS_f and the region R_0 of a blank slide as inputs for allocating slide components on a slide sheet. The process of allocating slide components starts from the component c_0 whose *role* in a topic frame f is *point*. First, the layout templates are selected by comparing their preconditions represented as logical structures with the partial logical structure (c_0, S_{c_0}) in LS_f. If a template t can be applied to the partial logical structure, the mapping from a subgroup S_{c_0} to a subgroup in t is computed. Next, c_0 is allocated according to the template t and the given region R_0. If c_0 can be allocated, the remaining region R is computed by removing the region occupied by c_0 from R_0. Otherwise, c_0 cannot be allocated on a slide because there is no longer enough space on a slide sheet. In this case, c_0 is marked as "un-allocated". After allocating the region for c_0, the regions for the components in the subgroup S_{c_0} are computed according to the mapping from S_{c_0} to a subgroup in t. Then, the process of allocation continues recursively for each subgroup in S_{c_0} and the subregion R.

The process described above is repeated until all the components in C_f are allocated on a slide or none of the remaining components can be allocated on the remaining regions. A slide is generated from the components that can be allocated at this step. After that, the allocation process is resumed to generate a slide for the remaining part of the logical structure marked as "un-allocated." This means that more than one slides can be generated from one topic frame.

Figure 6 illustrates an example of allocating components in the topic frame of Fig. 5a. In generating slides, layout templates are applied to logical structure of Fig. 5b. In this example, an image is included in the logical structure. First, the template for slides that contain images is selected. The region on a slide is divided vertically into two subregions and the image is allocated on the subregion on the right of a slide. Second, the remaining five texts are allocated on the subregion (G0) on the left of a slide. Here, the template for illustrating cause and effect relationship is selected because the two subgroups are related to the parent text by the relationship "caused-by". Then, the region G0 is divided into subregions G1, G2 and G3. Next, the shape that represents an arrow is added on a slide. The parent text is allocated on G1 and the two subgroups are allocated on G2 and G3. Third, the template for representing a hierarchy is applied to each of the two subgroups because both of them have the relationship "detailed-by". Finally, the template for expressing emphasis on a component is applied to the parent text. Thus, the final presentation slide is generated from the logical structure.

6 Prototype System

This section presents our prototype system and examples of slides generated from topic frames. We developed the system as a desktop application using Java. The system provides two interfaces for users. One is for composition of a presentation scenario by editing slide components and specifying semantic relationships among them. In implementing this interface, we used Piccolo [34], a Java library for developing a zoomable user interface. The other interface is for browsing generated slides. This interface allows a presenter to choose the most preferable slide from the generated candidate slides. The chosen slide is exported in the format of

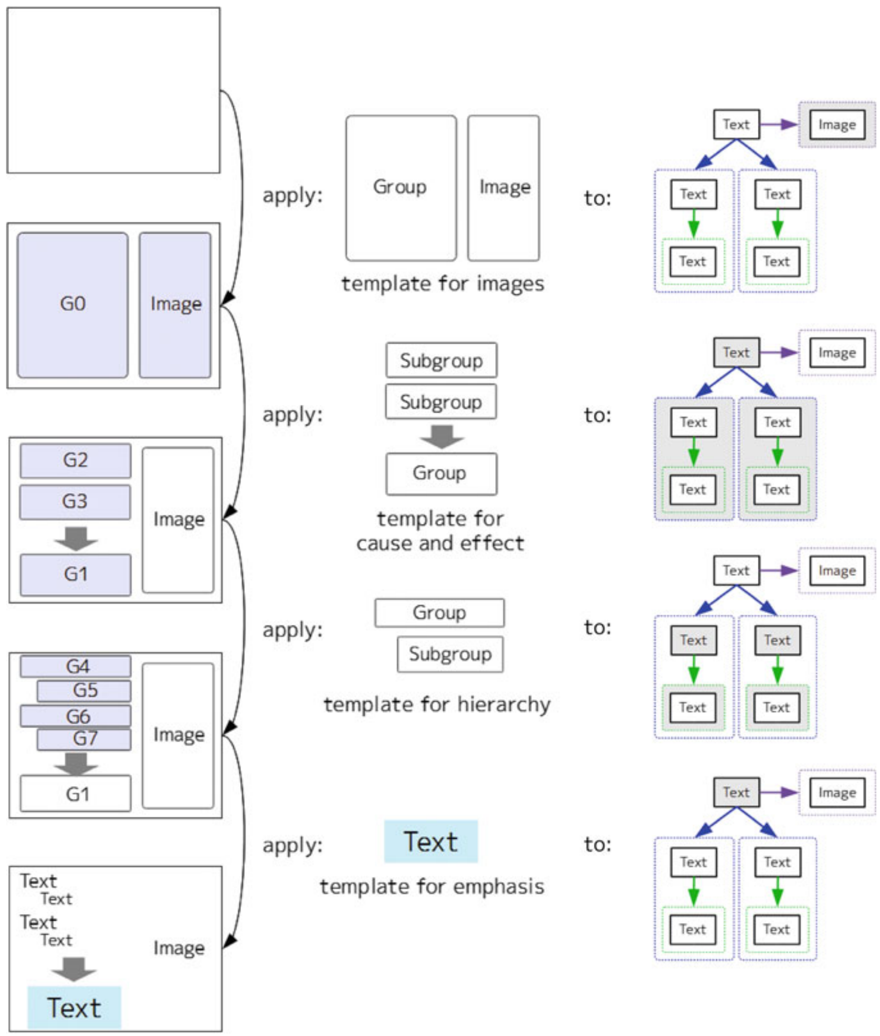


Fig. 6 An example of applying templates

PowerPoint presentation (*.ppt). We used Apache POI [35] for generating slides in the format of Microsoft PowerPoint.

6.1 Scenario Composition Interface

In composing a presentation scenario, a presenter should consider the consistency both between topics and within one topic. Therefore, it is effective to offer a function to change the working view of a presentation scenario easily and intuitively. To achieve this, we introduced the technique of zooming and panning in a scenario composition interface. Figure 7 shows the screenshots of the interface. With this interface, a presenter can change his/her working view by zooming and panning. In the local view shown in Fig. 7a, a presenter is able to see the detail of a specific topic frame. In the global view shown in Fig. 7b, a presenter is able to see the whole presentation. In these figures, the area inside a circle corresponds to a topic frame. Semantic relationships between components are represented as a link with caption. For example, a text component "PowerPoint has made it easy to make presentation slides." in Fig. 7a is connected to a text component "It is easy to make presentation slides with colorful graphics and animations." with a link that represents the relation "detailed-by". In Fig. 7b, the topic frame "Objective" (one at the upper left side) is connected to the two topic frames "Background", and "Related Work" (ones at the lower left side) with hierarchical relation. On the other hand, the topic frame "Objective" is connected to the topic frame "Approach" (one at the upper right) with sequential relation.

In addition to zooming functions, our interface provides a presenter with the operational functions as follows:

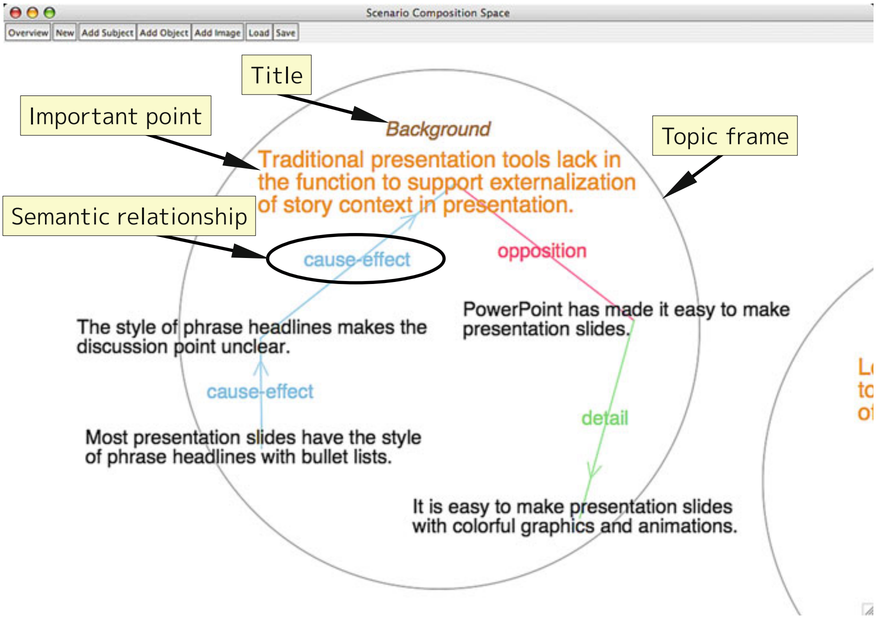
- Add a topic frame. This operation is done by specifying the name and the important point of a topic using a dialog window.
- Add a component. This operation is done by specifying a component using a dialog window.
- Connect a component to another one in a topic frame. This is done by dragging and dropping a component near the one that a presenter wants to connect to. After that, the type of relationship can be specified using a dialog window.
- Connect a topic frame to another one. This is done by dragging and dropping a component in one topic frame into another topic frame that a presenter wants to connect. After that, the type of relationship between the two topic frames and the type of relationship between the components are specified using a dialog window.

Since our interface provides the functions described above, it is possible for presenters to compose scenarios through intuitive manipulation on topic frames and slide components. Also, presenters can specify the semantic relationships between components, while they cannot do it using traditional presentation tools. Moreover, zooming and panning functions enable presenters to look at multiple topic frames at the same time. Therefore, our interface helps presenters to consider the consistency of their story of presentation scenarios.

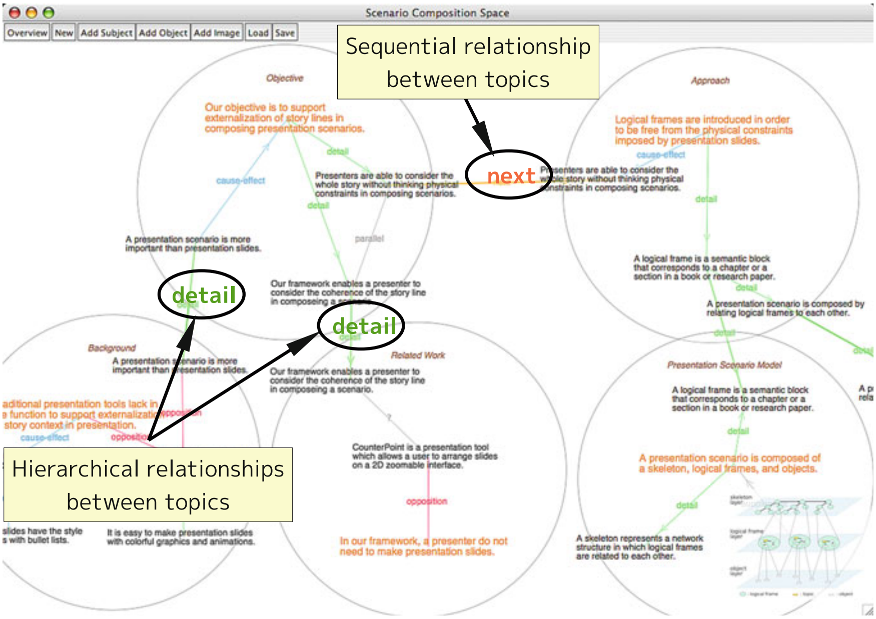
In Sect. 3, we have argued that the representation of a presentation scenario must be suitable for an early process of presentation preparation. Especially, the representation must have the following characteristics:

- A set of partially ordered topics, and
- Semantic relationships between items specified by a presenter.

Our interface has the functions for a presenter to compose a presentation scenario with the characteristics described above. Specifically, our system has the following functions to support composition of a presentation scenario by a presenter:



(a)



(b)

Fig. 7 Screenshots of the scenario composition interface. **a** Local view of a topic. **b** Global view of topics

- Zooming and panning for switching a working view smoothly,
- Specifying sequential/hierarchical relation between topics, and
- Specifying semantic relationships between slide components in each topic.

The first function enables a presenter to switch his/her working view between overview and detail. This makes it easy for a presenter to consider a presentation story and confirm if topics are connected with each other appropriately. The second function is realized by sharing a text component with the topics to be connected. This makes it easy for a presenter to understand how topics are related to each other. The third function is realized by moving a component near to the one to be connected and then selecting a type of relationship. Since moving components and topic frames is achieved by dragging them, a presenter is able to specify relationships between topics or components within a topic intuitively in our interface.

6.2 *Examples of Generated Slides*

In order to examine the feasibility of our method, we conducted a case study using our prototype system. The purpose of the case study is to confirm that our method makes it possible to generate slides from a given topic frame. We asked several graduate students to prepare for presentations for introducing the overview of their own research. The students were specializing in information science. First, they were asked to make presentation slides using PowerPoint. Then, they were asked to compose the scenario of the same topic using our system. We investigated how the system reflected the semantic relationships in the generated slides. From the viewpoint of the usability, it is necessary to evaluate the easiness of the operations in composing a scenario. We leave how to evaluate the utility of our system as our future work.

We prepared the templates listed in Fig. 8 for generating slides. For each template in this figure, the logical structure shown on the left is the precondition for applying the template. The descriptions of the templates are as follows:

Template for title A text component whose *role* is *title* is located on the top of the given region.
Template for an important text A text component whose *role* is *point* is located at the center of the given region with its background colored for emphasizing it.
Template for an image An image component is located on the bottom of the given region.
Template for sequence Groups that are related by "followed-by" relationship are arranged vertically with directed arrows on the given region.
Template for hierarchy Subgroups of a group that are related by hierarchical relationships are horizontally indented on the given region. Also, the subgroups are arranged vertically.

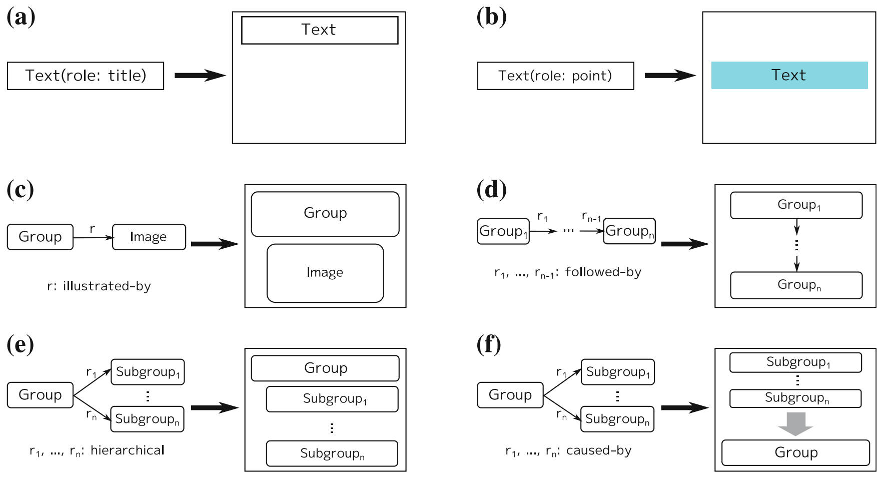


Fig. 8 Templates and preconditions defined in our prototype system. **a** Template for title. **b** Template for an important text. **c** Template for an image. **d** Template for sequence. **e** Template hierarchy. **f** Template for cause and effect

Template for cause and effect Groups that are related by "caused-by" relationship are arranged using an arrow between two groups on the given region.

Figure 9 presents an example. Figure 9a is the topic frame prepared by one participant. The uppermost text component with the largest font size in Fig. 9 is the title of a topic. The links between slide components are semantic relationships. In this example, the topic frame consists of one image and seven texts including the title. All the texts are related by the semantic relationship "detailed-by". Figure 9b is a generated slide. The correspondence of a slide component in these figures and its translation in English is shown in Table 2. The slide reflects the hierarchical relationships between text components in the style similar to a bullet-point list. Also, the image is allocated on the bottom of the slide. In addition, the generated slide looks quite similar to the original one (Fig. 9c). From this result, we can see that the generated slides are as expressive as those generated using existing methods such as [27–29].

Figure 10 presents another example. Figure 10a shows the topic frame prepared by another participant. The text component on the top of the figure is the title and the one on the bottom of the figure is an important text (a component whose *type*, *role* is *point*, *text*, respectively). The important text is related to one text by the relationship "caused-by" and the other texts are related by the relationship "detailed-by". Figure 10b is a generated slide. The correspondence of a slide component in these figures and its translation in English is shown in Table 3. In this slide, an important text is highlighted according to the template for emphasis (Fig. 8b). Also, the relationship of cause and effect is expressed by using an arrow between text components. This diagrammatic representation is adopted because

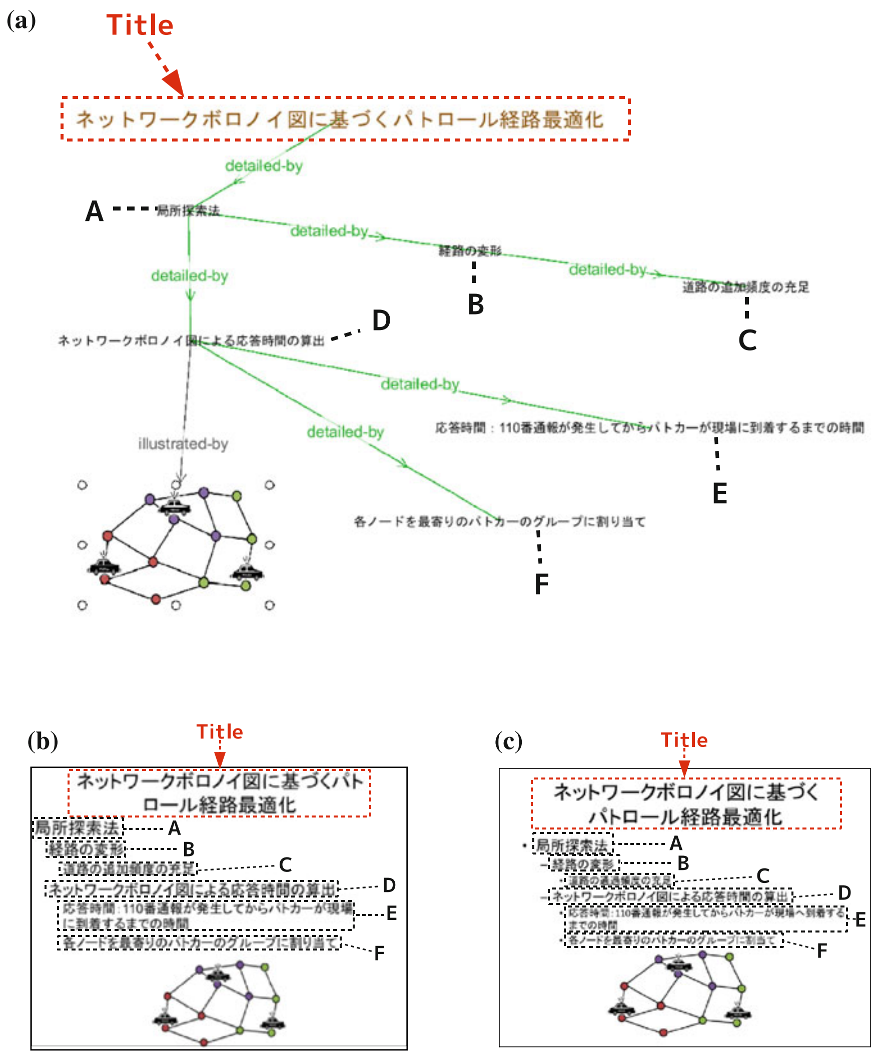


Fig. 9 An example of a generated slide from semantic relationships. **a** Topic frame composed by a participant. **b** Generated slide. **c** Original slide

the logical structure matched the precondition of the template for cause and effect (Fig. 8f). Although the layout of the generated slide is not so balanced as the original slide (Fig. 10c), it has almost the same expressiveness as the original one. From this result, we can see that our prototype system generates slides that are more expressive than those generated using existing methods.

From the examples described above, we confirmed that our prototype system generates presentation slides that are as expressive as those generated by existing methods as shown in Fig. 9. In addition, our prototype system also generates slides

Table 2 Correspondence between a component in Fig. 9 and its translation in English

Component	Translation in English
Title	Optimization of patrol routes based on network Voronoi diagram
A	Local search method
B	Modification of patrol routes
C	To satisfy the frequency of passing each road segment
D	Estimation of response time using network Voronoi diagram
E	Response time: the time it takes a police car to arrive at the scene after an emergency call
F	To assign nodes in a road network to a group of police cars

with the diagrammatic template as shown in Fig. 10. While the existing methods generate slides with limited format such as bullet-point lists, our method can generate the slides with more effective layout. Specifically, our prototype system generated the slides with highlights for expressing emphasis and diagrammatic representation for expressing cause and effect.

Although we confirmed that our method generates slides with diagrammatic representation, we have two main problems. The first problem is that the physical features of the slide components are not considered in the precondition of a layout template. Currently, our prototype system checks if a layout template can be applied by referring to its logical structure only. However, the physical feature such as the number of components and their sizes should be considered in deciding the layout of a slide. A possible solution to this problem is to introduce the physical constraints to the precondition of a layout template. The second problem is that the templates in our method does not cover enough diagrammatic representation. For example, our prototype system cannot generate slides with the following representation:

- Tabular layout,
- Mutual dependency like Venn diagram,
- Circular layout.

A possible solution to this problem has two steps. First, templates have to be increased in order to cover the variety of SmartArts prepared in PowerPoint. Second, a mechanism to attach an attribute to groups of components that indicates relationships such as “complement”, “independent-of” and “circular”.

7 Comparison with Existing Systems

7.1 Systems Handling Semantic Contents

The existing systems based on the KJ method [17, 18] provides the function for associating fragments of ideas with each other by grouping and linking them.

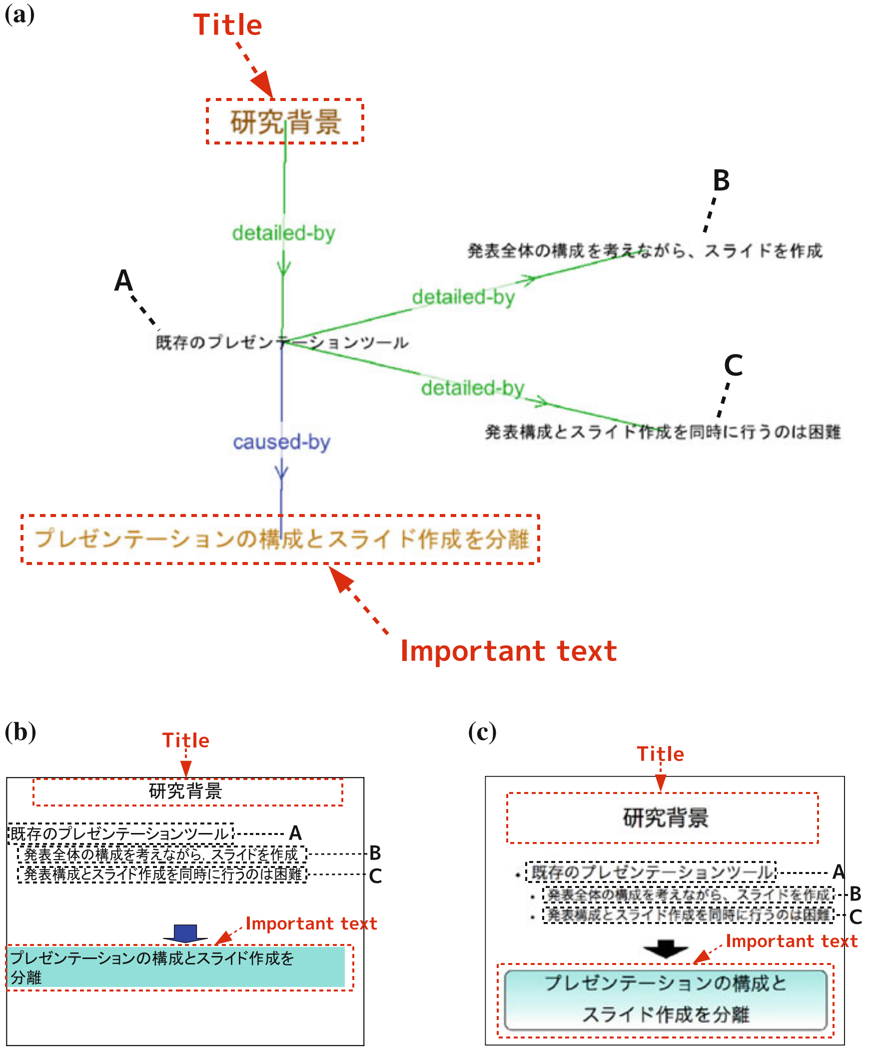


Fig. 10 An example of a generated slide with an important text emphasized and a causal relationship represented using an arrow. **a** Topic frame composed by a participant. **b** Generated slide. **c** Original slide

Although the systems enables a user to specify a link between groups of fragments, the user cannot specify how they are related with each other. On the other hand, our system enables a user to specify the type of a semantic relationship between fragments. This is especially effective in supporting presentation preparation because a presenter must clarify the connectivity between his/her concepts and messages in order to make a presentation story persuasive.

Table 3 Correspondence between a component in Fig. 10 and its translation in English

Component	Translation in English
Title	Research Background
Important text	Composing a story and making slides should be separated.
A	Conventional presentation tools
B	Presenters have to consider a whole story while making slides.
C	It is difficult for a presenter to compose a story and make slides at the same time.

7.2 Systems Supporting Slide Composition

Many of the existing slide generation systems apply the techniques of text summarization to resources such as paper manuscripts. The method proposed by Shibata et al. generates slides based on the coherence relations between clauses and sentences as follows [28]:

Coherence relations between two clauses in a sentence list, consrast, additive, adversative,
Coherence relations between two sentences list, contrast, topic-chaining, topic-dominant changing, elaboration, reason, cause, example.

These relations are detected and used to determine which of the sentences and clauses should be included in slides. Although their method considers various types of the relations among clauses and sentences, the output format of a slide is limited to the bullet-point list as prepared in the traditional presentation tools. On the other hand, our system can generate a slide in the format other than the bullet-point list from a topic, even if the topic contains only text components.

From the viewpoint of the usage, existing systems for automatic slide generation are effective when a presenter prepares for a presentation in an academic conference. Also, they are effective for the preparation in reading a specific book by turns. On the other hand, our system generates a presentation slide from the pieces of slide components accumulated by a presenter. Since our prototype system enables a presenter to organize ideas for a presentation, our method is expected to be effective for the presentation in group meetings and lightening talks.

The system for supporting slide composition developed by Yasumura et al. [30] prepares the templates as follows:

- Bullet-point lists,
- Figures (a figure itself, a figure with description, a figure with a bullet-point list),
- Tables (a table itself, a figure with description, a figure with a bullet-point list),
- Mathematical expressions (an expression itself, an expression with description),
- Templates for specific topics (background and objective, proposed method, experiment, summary).

While their system makes it possible to compose a slide based on one template, our system generates slides by selecting and applying appropriate layout templates recursively from the important component in a topic. This means that our system can generate a slide so that different layout templates will be applied to the partial blocks of a slide. Therefore, our system provides the diversity of diagrammatic representation in generating slides compared with the existing systems.

8 Conclusion

In this chapter, we presented our prototype system for supporting presentation preparation based on the concept of a scenario. The interface for composing a presentation scenario has two features. One is to allow a presenter to externalize the semantic relationships between slide components. The other is to allow a presenter to change the working view smoothly through the functions of zooming and panning. It is expected that these features make it easy for a presenter to compose a consistent story of his/her presentation. Also, we have developed a function of generating presentation slides automatically from a presentation scenario. This function is based on the mechanism of transformation from the logical structure to the geometric structure. In order to generate slides with diagrammatic representation effectively, our system constructs the logical structure from a topic frame and then find appropriate layout templates by comparing the logical structures. In the case study, we confirmed that some of the slides were successfully generated by our system and had more expressiveness than the slides generated by existing methods. Although this concept of a logical structure and a geometric structure with a view to interpreting document organization has been successfully investigated as research fields of document image understanding, in our case this transformation takes an important role to generate the presentation slides.

One of our future works is to evaluate the effectiveness of our framework through the case study. The effectiveness must be evaluated from the viewpoints of usability, availability for thinking support, and possibility of reuse. For example, we have to confirm how much effort of users our system reduces for making slides. Also, it is necessary to extend the definition of layout templates and to evaluate the effectiveness of our method for generating slides. In order to extend the definition, we will increase the number of layout templates and introduce physical features for their precondition. Specifically, we have to confirm whether the generated slides cover the expressiveness of SmartArt in PowerPoint. In addition, we have to devise a mechanism to generate slides from multiple topics. Since our method assumes that the topic frame as the input corresponds to only one topic such as background, objective and conclusion, we have to develop a mechanism of generating slides from the semantic relationships corresponding to more than one topics. Furthermore, it is important to support the process of presentation preparation from the viewpoint of what content should be included in a presentation. That is because the content varies according to the situations such as the spoken language, the background knowledge of audience and time constraints, even though the subject of a presentation were the same. Therefore, we will consider how to support the preparation, selection and reuse of scenarios and slide components as well as slides.

References

1. TED: Ideas worth spreading, http://www.ted.com/ [accessed 18/6/2012]
2. USTREAM, http://www.ustream.tv/ [accessed 18/6/2012]
3. Microsoft Inc.: PowerPoint, http://office.microsoft.com/powerpoint/ [accessed 18/6/2012]
4. Apple Inc.: Keynote, http://www.apple.com/iwork/keynote/ [accessed 18/6/2012]
5. OpenOffice.org: OpenOffice Impress, http://www.openoffice.org/ [accessed 18/6/2012]
6. SlideShare, http://www.slideshare.net/ [accessed 12/6/2012]
7. Prezi, http://prezi.com/ [accessed 18/6/2012]
8. Alley, M., Schreiber, M., Muffo, J.: Pilot testing of a new design for presentation slides to teach science and engineering, Proceedings of ASEE/IEEE FIE2005, (2005) S3G-7–12.
9. Tufte, E.R.: The Cognitive Style of PowerPoint. Graphics Press, Cheshire (2004)
10. Hanaue, K., Watanabe, T.: Externalization support of key phrase channel in presentation preparation. International Journal of Intelligent Decision Technologies **3–2**, 85–92 (2009)
11. Watanabe, T., Ishiguro, Y., Hanaue, K.: Automatic composition of presentation slides, based on semantic relationships among slide components, Proceedings of the 4th International Conference on Intelligent Interactive Multimedia Systems and Services (KES/IIMSS 2011: SIST 11. Piraeus, Greece, 261–270 (2011)
12. Norman, D.: Things That Make Us Smart. Addison-Wesley, New York (1993)
13. Yamamoto, Y., Nakakoji, K., Takada, S.: Hands on representation in a two-dimensional space for early stages of design**13–6**, 375–384 (2000)
14. Hasida, K.: Semantic authoring and semantic computing. New Frontiers in Artificial Intelligence, LNCS **3609**, 137–149 (2007)
15. Kawakita, J.: A Method for Idea Organization. Chuko-Sinsho, Japan (1967). [in Japanese]
16. Spool, J.M.,: The KJ-technique: A group process for establishing priorities, User Interface Engineering, http://www.uie.com/articles/kj_technique/ [accessed 24/9/2012]
17. Oiwa, H., Takeda, N., Kawai, K., Shiomi, A.: KJ editor: a card-handling tool for creative work support. Knowledge-Based Systems **10–1**, 43–50 (1997)
18. Sugiyama, K., Misue, K.: A generic compound graph visualizer/manipulator: D-ABDUCTOR. LNCS **1027**, 500–503 (1995)
19. Hsieh, H., Shipman, F.M.: VITE: A Visual Interface Supporting the Direct Manipulation of Structured Data Using Two-Way Mappings. Proceedings of the 5th International Conference on Intelligent User Interfaces (IUI '00), 141–148 (2000)
20. Haller, H., Abecker, A.: iMapping - A zooming user interface approach for personal and semantic knowledge management, Proceedings of the 21st ACM Conference on Hypertext and Hypermedia (HT '10). Tronto, Ontario, Canada (2010)
21. Nelson, L., Ichimura, S., Pedersen, E.R., Adams, L.: Palette: A paper interface for giving presentations. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99), 354–361 (1999)
22. Moscovich, T., Scholz, K., Hughes, J.F., Salesin, D.H.: Customizable presentations, Technical Report CS-04-16. Brown University, Computer Science Department (2004)
23. Good, L., Bederson, B.: Zoomable user interfaces as a medium for slide show presentations. Information Visualization **1–1**, 35–49 (2002)

24. Lichtschlag, L., Karrer, T., Borchers, J.: Fly: A tool to author planar presentations, Proceedings of the CHI 2009 Conference on Human Factors in Computing Systems. MA, USA, Boston (2009)
25. Drucker, S., Petschnigg, G., Agrawala, M.: Comparing and managing multiple versions of slide presentations, Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST '06), pp. 47–56. NY, USA, New York (2006).
26. Bergman, L., Lu, J., Konuru, R., MacNaught, J., Yeh, D.: Outline wizard: presentation composition and search, Proceedings of the 15th International Conference on Intelligent User Interfaces (IUI '10), pp. 209–218. NY, USA, New York (2010).
27. Miyamoto, M., Sakai, H., Masuyama, S.: Research on automatic generation of presentation slides from a LaTeX manuscript of a paper. Journal of Japan Society for Fuzzy Theory and Intelligent Informatics **18–5**, 752–760 (2006). [in Japanese]
28. Shibata, T., Kurohashi, S.: Automatic slide generation based on discourse structure analysis, Proceedings of the 2nd International Joint Conference on Natural Language Processing (IJCNLP-05), pp. 754–766. Jeju Island, Korea (2005).
29. Utiyama, M., Hasida, K.: Automatic slide presentation from semantically annotated documents. Proceedings of the ACL Workshop of Coreference and its Applications, Stroudsburg, PA, USA (1999)
30. Yasumura, Y., Takeichi, M., Nitta, K.: A support system for making presentation slides. Transactions of the Japanese Society for Artificial Intelligence **18–4**, 212–220 (2003). [in Japanese]
31. Reynolds, G.: Prezentation Zen: Simple ideas on presentation design and delivery. New Riders Press, Upper Saddle River (2007)
32. Watanabe, T.: Document analysis and recognition. IEICE Transactions on Information and Systems **E82–D–3**, 601–610 (1999)
33. Mann, W., Thompson, S.: Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. Text **8–3**, 243–281 (1988)
34. Bederson, B.B., Grosjean, J., Meyer, J.: Toolkit design for interactive structured graphics. IEEE Transactions on Software Engineering **30–8**, 534–546 (2004)
35. Apache POI Project, http://poi.apache.org/ [accessed 18/6/2012]