Designing a digital patient avatar in the context of the MyHealthAvatar project initiative

Evaggelia Maniadi, Haridimos Kondylakis, Emmanouil G. Spanakis, Marios Spanakis, Manolis Tsiknakis, Kostas Marias and Feng Dong

Abstract— **The digital avatar is a vision for the digital representation of personal health status in body centric views. It is designed as an integrated facility that allows collection of, access to and sharing to life-long and consistent data. A number of Virtual Physiological Human (VPH) communities have started the movement to this direction by creating a digital patient road-map and by supporting data sharing infrastructures. As an innovative concept, the impact of digital patient and avatar to personalized medicine and treatment is yet to be clear. This requires a focused and concerted effort in addressing various questions regarding user perspective, use cases and scenarios. This paper presents use cases and future scenarios realizing the vision for the digital avatar as well as architectural consideration for the envisaged platform.**

I. Introduction

THE risks, developments and treatment of many major chronic diseases (neurological, cardiovascular, cancer, diabetes) are affected by a great number of individual health factors, genetic information, age, lifestyle and the environment. The dynamic and temporal nature of them could pose threats to the provision of high quality and integrated healthcare services. Therefore, gaining access to long-term and consistent data collection of these health factors could be particularly useful for supporting individualized prediction and treatment. Moreover, the collection of data across many individuals would lead to a huge and comprehensive population data resource, which would offer extremely valuable input to clinical research for new knowledge discovery.

The rapid progress of ICT offers great potential for addressing challenges in health information access, collection, sharing and analysis. The VPH initiative has led to the collection and integration of predictive models and heterogeneous data to interpret and predict the progress of diseases and the effectiveness of treatments, which have laid down the foundation for new knowledge discovery. However, access to these resources for clinically meaningful use remains a largely unresolved problem. A number of VPH communities have started the movement to this direction by creating a digital patient road-map and by supporting data sharing infrastructures. Among, the most relevant projects is Discipulus [2], which is a support action for producing a road-map for the future of digital patient, VPH Share [3], p-Medicine [4], ACGT [5] and Tumor [6], of all which have shown a high level of interest in data sharing, owing to the importance of data in VPH research.

The general concept of using digital avatar has been proposed by a number of systems. Typical examples include the 3D Avatar from IBM [7], HealthVault from Microsoft [8] and Google Health [9]. These systems have only shown limited impact on the current healthcare and after some time have stopped further development (i.e. like Google Health in 1st Jan, 2012[1]). Also, similar work on the display of a human avatar was described in Zygote Body [10], formally Google Body, which is a web application allows rendering and manipulation of 3D anatomical human bodies. However, it does not concern data collection for medical purposes.

MyHealthAvatar[2] is a research initiative through which the feasibility of an innovative representation of the health status of citizens for future healthcare will be studied. The goal is to create a digital avatar acting as a mediator between the end-users and health related data collections. It is envisaged as a personal container of heterogeneous sources of information (medical, environmental, lifestyle) all blended in a single framework, utilizing modern information and communication technologies, providing long-term and consistent health information along a person's a timeline [2]. In this paper we focus on the state of the architectural requirements and considerations for the implementation as well as real life key technologies for the envisaged service platform.

II. State of The Art Architectural Requirements

The amount of data in healthcare that need to be managed and the computational challenges introduced today are ground breaking and call for innovative system design strategies. These requirements have given rise to High

Manuscript received October July 30, 2013.This work was supported by the project MyHealthAvatar funded by the European Commission (EC) 7th Framework Programme, Grant agreement no: 600929.

Evaggelia Maniadi, is with the Institute of Computer Science (ICS), Foundation for Research and Technology - Hellas (FORTH), N. Plastira 100, Vassilika Vouton, PO BOX 70013 Heraklion, Crete, Greece (phone: +302810391479; fax: +302810391428; e-mail:maniadi@ics.forth.gr).

Haridimos Kondylakis, E mmanouil G. Spanakis, Marios Spanakis, Kostas Marias are with Institute of Computer Science (ICS), Foundation for Research and Technology - Hellas (FORTH), (e-mail: {kondylak, spanakis, marspan, kmarias} @ics.forth.gr).

Manolis Tsiknakis is with the Technological Educational Institute of Crete and the Institute of Computer Science (ICS), Foundation for Research and Technology - Hellas (FORTH) (e-mail: tsiknaki@staff.teicrete.gr).

Feng Dong is with the Department of Computer Science and Technology, University of Bedfordshire, Luton, UK (e-mail: Feng.Dong@beds.ac.uk).

[1] http://www.google.com/intl/en-US/health/faq.html
[2] http://www.myhealthavatar.eu/

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

Performance Computing (HPC). There are multiple instantiations of HPC, including cluster computing; grid computing and cloud computing [17]. The cloud platform is currently the preferred solution for offering on-demand computing and data storage resources and today many such systems exist but are available but face performance restrictions and inherent legal and ethical implications. In contrast, private clouds [18] become very useful for biomedical scientific research in order to have a full control on data access, reliability and storage management. A cloud platform includes virtualization and software stack for the infrastructure, optimization of resources and allocation, access to external sources, support of scalable data repository and multiple cloud storages (via data federation), data reliability and integrity, and security over the network and infrastructure.

From the perspective of web service, an emerging trend is the use of simple REST web services [19] that present a small entry barrier and a transition from the SOAP and WSDL Web Services technologies. These architectures are semantic-web friendly since they share common basic infrastructure and interaction protocols. Architecture linking to external sources for data exchange concerns standardization and interoperability issues include EHR, HL7, OpenEHR and more.

MyHealthAvatar will take advantage of a range of ICT advances, such as simulation models, semantics, visual analytics technology, web-based technologies. It can be seen as a multimodal user interface allowing clinicians and citizens to:

- collect, store and manage individual citizen information
- exchange and share information
- access external resources
- support the presentation of clinical information
- assist clinical analysis and decision making
- discover new knowledge (biomedical and clinical research)

MyHealthAvatar proposes a solution for access, collection and sharing of long term and consistent personal health status data through an integrated digital representation in *in-silico* environment, which will help to deliver clinical analysis, prediction, prevention and treatment tailored to the individual citizen.

III. Key Technologies

Based on different perspectives, four diverse key technologies enabling a digital avatar for patient empowerment are presented: a) semantic interoperability and linked data, b) use of mobile apps and social networks for extracting useful information, c) multi-scale visualization and d) use of digital avatar interface for *in silico* clinical trials. Digital avatar will use innovative technologies to support data collection, search and reasoning, with minimal user effort. For example, mobile apps and social networks will be used to collect useful and valuable data from patients. Moreover, ontologies will support the annotation and transformation of multi-scale medical and social data and models in a unified way. This will facilitate data search and reasoning and the integration of data from a range of different resources (e.g. clinical, laboratory, social).

Digital avatar involves a significant amount of multi-scale biomedical data. These data will be a 3D+time dataset of which multiple instances at different scales will have to be displayed together. Information will be on very different spatial and temporal scales going from the molecule up to body level, in different forms and of heterogeneous dimensionality. Such large scale data can lead to over-plotting, which significantly prevent the capability of human vision in identifying data patterns. One of the key challenges is the effective, interactive visualization of biomedical data. Moreover, digital avatar will serve as a tool for clustering population into similar cohorts of patients and/or healthy groups which on a next step can be the "virtual participants" in respective *in-silico* clinical trials. Furthermore, it will be beneficial for the continuous development of *in-silico* approaches as they are applied in many pharmaceutical organizations. Finally, it can initiate the development of approaches for generating *in-silico* observational studies.

Detailed descriptions of the key technologies for enhancing the digital avatar vision are given in the rest of the paper.

A. *Semantic Interoperability and Linked Data*

Data collection and semantic interoperability is one of the keys to the success of the proposed avatar. There are huge amounts of heterogeneous, disparate and multilevel data that should be collected and linked to the aforementioned avatar. These data could be personal information, lifestyle, social data, health status, clinical and medical information, laboratory values, predictive models, medical data etc. Moreover the input of these data needs to be as much effortless for the patients as possible.

The amount, diversity, and heterogeneity of that information have led to the adoption of data integration systems in order to manage it and further process it. However, the integration of these disparate data sources raises several semantic heterogeneity problems. By accepting an ontology as a point of common reference naming conflicts are eliminated and semantic conflicts are reduced. Ontologies are used to identify and resolve heterogeneity problems, usually at schema level, as a means for establishing an explicit formal vocabulary to share. During the last years, ontologies have been used as global schemata in database integration [11], obtaining promising results, for example in the fields of biomedicine and bioinformatics [5].

As a result, the approach to semantic interoperability and data integration is an ontology-based one. The ontologies selected/generated will be used to annotate and/or transform multi-scale data and models that will be made then available in a unified way. This will facilitate registration, discovery, comparison and interoperability between models, and the integration of data from clinical, laboratory, social and modelling domains. However, the idea is not to "re-invent the wheel" from scratch but to exploit and reuse past approaches to describe the semantics of clinical domain, social activity and models.

Then techniques from data translation, query translation and information linkage will be combined to achieve better results. For example, for data which do not change so often data translation will be employed, whereas for rapidly changing data, query translation will be used. Moreover, information linkage to several web-sites and web-databases will allow the augmentation of the available information. To this direction, different sources of information will be identified and mapped to the ontology used in order to enable data integration.

Being able to access the mass amount of information another set of techniques will be used to augment patient experience. So instead of presenting big results with large sets of information, we will construct summaries that will be presented to the users. Those summaries will include anchors to complete ontological descriptions and will also function as indices in the knowledge space, being able to be reused by other tools and by other semantic-based decision support systems.

Finally, we also aspire to account for evolving models by advancing our prior work [11-12]. Since the digital avatar is indented to be a lifelong companion of a citizen, the core ontology might have to change to depict the new knowledge that is acquired. Ontology evolution is also a subject that, with the exception of the aforementioned works, remains largely untouched in the prominent approaches to achieving semantic integration.

B. Mobile Lifestyle and Social Data

It is very beneficial to use information extraction and mobile applications to allow for data collection to the data repository dynamically along a timeline with minimal user effort. These data are useful to doctors in diagnosis and clinical decision making. Both smart phones and social media are supported by the latest technology and they have a huge number of participants and customers. It is expected that connecting the avatars to these facilities will significantly increase the level of participation from public.

In particular, mobile applications will monitor user's "health-status", "lifestyle" and "wellness" and upload data to the avatar system for close monitoring of health conditions and prevention of many diseases. The system then will be able to analyze user's lifestyle and medical data. Special "alerts" will be applied to support end users with feedback supporting and assisting their daily activities and well-being.

A social media service will be used to allow the interconnection of end users. This social media service, accessible by smart phones, will be used in a dual mode allowing the users to insert information about themselves (like they do in common social media technologies) but also will be a mean of supporting personalized services to them from the system in the form of alerts and guidance (i.e. post therapy monitoring of user's behaviors after orthopedics operation, cancer patients reaction to treatment, etc.). The user will be able to take advantage of mobile digital technology using 3D visualization models the project will deploy.

C. Multiscale Visualization

Various tests are used for the diagnosis, monitoring and treatment of various diseases. These may include patient history, physical examination, laboratory tests, computed tomography (CT), magnetic resonance imaging (MRI), electroencephalography, electrocardiogram etc.

Specifically, patient history helps a doctor to assess an individual's past and current health situation. It also helps evaluating any medical problems, developing a plan of treatment, and monitoring the patient's health over time. This may include information about age, sex, history of current illness, past medical history, memory loss events etc. Physical examination enables the doctor to assess the overall physical condition of the patient. The physical exam includes, for example, an examination of vital signs (temperature, blood pressure, pulse), height and weight, etc. Furthermore, there are hundreds of laboratory tests available to help a doctor make a diagnosis. The most common are blood tests and urinalysis. Blood tests can be used to look for the presence of a specific gene that has been identified as a risk factor for a disease.

Considering the case of chronic disease, the above multi-level medical data will exhibit a strong dynamic and temporal nature. Interactive multi-scale visualization is necessary for supporting data reasoning and search. This will offer a useful input to doctors and will help them to carry out personalized healthcare. A first step, targeting multi-scale visualization is the use of different markers on the avatar, indicating the existence and the location of available datasets on different levels, from molecule to body level. Using these markers, navigation on different levels (from body to molecule level and reverse) is supported. These temporal multi-scale data can be represented by "animating" the visualization over the time. Each frame displays the value of each parameter at a given time point. On the other hand, for the spatial multi-scale data, a user can configure multiple views of the same dataset. The user can move from one scale to other by clicking on the visual markers, which show the presence of lower scale data.

D. In-silico clinical trials

In-silico clinical trials (ISCTs) try to enhance prediction of the effects of medicines in "virtual populations" [13, 14]. ISCTs developed through 3 distinctive and same time interacting steps (Figure 1) [15]. The first step is to characterize the underlying disease which requires a detailed understanding of its pathophysiology in molecular and clinical level as well as to identify possible drug targets and estimation of some basic pharmacokinetic properties (PK) for drug-candidates which are designed for these targets. The second step intends to gather data regarding clinical efficacy and safety in the specific subpopulations of patients that the new drug is aimed to be administered. In this step, specific parameters and covariates regarding pharmacodynamic action (PD), possible pharmacogenomic patterns and development of virtual patient profiles are needed. In the final third step, as long as disease model and drug PK/PD properties have been described, in-silico clinical trials can be designed and execute through simulations in order to predict and evaluate the possible results [15].

Digital avatar will serve as a tool for clustering population into similar cohorts of patients and/or healthy groups which on a next step can be the "virtual participants" in respective *in-silico* clinical trials. Following patient's approval parts of their data from avatar profiles could be shared among different research disciplines (pharmacological, bio-medical, biological, computational etc) for evaluation of laboratory observations. In addition avatar can initiate strategies of generating data in special cases that clinical trials cannot be performed (e.g. pregnancy). Furthermore, digital avatar will be beneficial for the continuous development of *in-silico* approaches as they are applied in many pharmaceutical organizations (such as development of novel cancer drugs). It is believed that the application of simulating and predicting models regarding the PK/PD properties of novel drugs will cover the 20% of research and development (R&D) expenditure in the upcoming years [16].

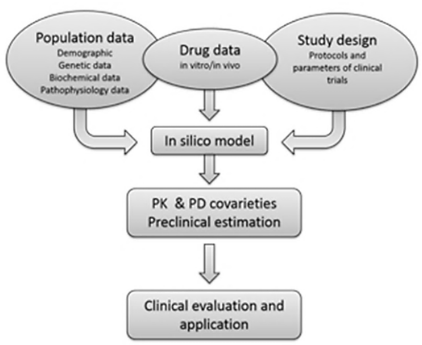

Figure 1: ISCT development process

Apart of the advantages on developing and generating results for simulating interventional studies, digital avatar as a tool for monitoring health status, can initiate the development of approaches for generating *in-silico* observational studies (i.e. cohort, case control, descriptive or cross-sectional studies). These studies try to provide information for the general population regarding risk factors for a disease susceptibility and evaluation of treatment effectiveness. The possible accessibility in integrated population data – such as those that will be in a digital avatar – could be beneficial regarding the results obtained and presented from research teams that are focusing in observational and pharmacovigilance studies.

IV. Conclusion

The ageing of the population, healthcare and associated social welfare costs are growing exponentially and they will soon become unsustainable unless we change the way in which people are supported. In many cases, there is a need to shift medical care from institutions to the home environment. To this end, ICT tools are being proposed and studied to reform the traditional ways in which medical data are recorded, tested and analyzed, without in any way reducing its quality. Digital avatar vision makes it possible to set up new interactions between doctors and patients, supporting individualized medicine and treatment. As an innovative concept, more effort is required to focus in addressing questions regarding user perspectives, use cases, scenarios, the power and limitations of modern technologies, the clinical acceptability of digital avatar in terms of supporting personalized medicine and potential legal and ethical aspects of digital avatars.

Acknowledgment

This work has been supported by the eHealthMonitor (http://www.ehealthmonitor.eu) and MyHealthAvatar (http://www.myhealthavatar.eu/) projects and has been partly funded by the European Commission under contract FP7-287509 and FP7-600929 respectively.

References

[1] Migration Information Source, Available: http://www.migrationinformation.org/feature/display.cfm?ID=402
[2] DISCIPULUS, Roadmap for the Digital Patient, Available: http://www.digital-patient.net/
[3] VPH-SHARE, Available: http://www.vph-share.eu/
[4] P-MEDICINE: From data sharing and integration via VPH models to personalized medicine, Available: http://www.p-medicine.eu/
[5] ACGT: Advancing Clinico-Genomic Trials on Cancer, Available: www.eu-acgt.org
[6] Transatlantic TUmour MOdel Repositories, Available: http://tumor-project.eu/index.html
[7] http://www-03.ibm.com/press/us/en/pressrelease/22375.wss
[8] HealthVault, Available: https://www.healthvault.com/gr/en
[9] Google Health, Available: http://en.wikipedia.org/wiki/Google_Health
[10] Zygote Body, , Available: http://www.zygotebody.com/
[11] H. Kondylakis and D. Plexousakis."Exelixis: Evolving Ontology-Based Data Integration System", in Proc. SIGMOD Conference, Athens, 2011, pp. 1283-1286.
[12] H. Kondylakis and D. Plexousakis. "Ontology Evolution without Tears", Journal of Web Semantics, vol 19, 2013, pp.42-58, Elsevier.
[13] M. Jamei, G. L. Dickinson, and A. Rostami-Hodjegan, "A framework for assessing inter-individual variability in pharmacokinetics using virtual human populations and integrating general knowledge of physical chemistry, biology, anatomy, physiology and genetics: A tale of 'bottom-up' vs 'top-down' recognition of covariates", Drug Metab Pharmacokinet, vol. 24, 2009, pp. 53-75.
[14] A. Rostami-Hodjegan and G. Tucker, " "In silico" simulations to assess the "in vivo" consequences of "in vitro" metabolic dru-drug interactions", Drug Discovery Today: Technologies, vol. 1, 2004, pp. 441-448.
[15] S. Michelson, A. Sehgal, and C. Friedrich, "In silico prediction of clinical efficacy," Curr Opin Biotechnol, vol. 17, 2006, pp. 666-70.
[16] H. van de Waterbeemd and E. Gifford, "ADMET in silico modelling: towards prediction paradise?" Nat Rev Drug Discov, vol. 2, 2003, pp. 192-204.
[17] P. Mell and T. Grance. The NIST Definition of Cloud Computing. National Institute of Science and Technology. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
[18] Amazon Virtual Private Cloud http://aws.amazon.com/vpc/2011.
[19] R.T. Fielding and R.N. Taylor. Principled design of the modern Web architecture, ACM Transactions on Internet Technology (TOIT), 2(2), 115 – 150, 2002.