Manifesto from the Dagstuhl Perspectives Workshop 13342

ICT for Bridging Biology and Medicine

Edited by

Jonas S. Almeida[1], Andreas Dress[2], Titus Kühne[3], and Laxmi Parida[4]

1 **University of Alabama at Birmingham, Dept. of Pathology, US**
jalmeida@uab.edu
2 **Wiss. Zentrum, infinity[3] GmbH, Gütersloh, Germany**
andreas.dress@infinity-3.de
3 **German Heart Institute Berlin, Germany**
titus.kuehne@dhzb.de
4 **IBM T J Watson Research Center, Yorktown Heights, NY, US**
parida@us.ibm.com

Abstract

The systems paradigm of modern medicine presents both, an opportunity and a challenge, for current Information and Communication Technology (ICT). The opportunity is to understand the spatio-temporal organisation and dynamics of the human body as an integrated whole, incorporating the biochemical, physiological, and environmental interactions that sustain life. Yet, to accomplish this, one has to meet the challenge of integrating, visualising, interpreting, and utilising an unprecedented amount of *in-silico*, *in-vitro* and *in-vivo* data related to health care in a systematic, transparent, comprehensible, and reproducible fashion. This challenge is substantially compounded by the critical need to align technical solutions with the increasingly social dimension of modern ICT and the wide range of stakeholders in modern health-care systems.

Unquestionably, advancing health-care related ICT has the potential of fundamentally revolutionising care-delivery systems, affecting all our lives both, personally and – in view of the enormous costs of health–care systems in modern societies – also financially.

Accordingly, to ponder the options of ICT for delivering the promise of systems approaches to medical care, medical researchers and physicians, biologists and mathematicians, computer scientists and information–systems experts from three continents and from both, industry and academia, met in Dagstuhl for a Dagstuhl Perspectives Workshop on *ICT Strategies for Bridging Biology and Medicine* from August 18 to 23, 2013, to thoroughly discuss this multidisciplinary topic and to derive and compile a comprehensive list of pertinent recommendations – rather than just to deliver a set package of sanitised powerpoint presentations on medical ICT. The recommendations in this manifesto reflect points of convergence that emerged during the intense discussions and analyses taking place the workshop. They also reflect a particular attention given to the identification of challenges for improving the effectiveness of ICT approaches to Systems Biomedicine.

Perspectives Workshop 18.–23. August, 2013 – www.dagstuhl.de/13342
1998 ACM Subject Classification I.2.1 Applications and Expert System: Medicine and science, H.3.5 Online Information Services: Web-based services, H.4 Information systems applications, J.3 Life and medical sciences: Medical information systems, etc.
Keywords and phrases Systems medicine, health-care related information systems, biomedical workflow engines, medical cloud, patient participation, ICT literacy
Digital Object Identifier 10.4230/DagMan.3.1.31

ICT for Bridging Biology and Medicine, *Dagstuhl Manifestos*, Vol. 3, Issue 1, pp. 31–50
Editors: J. S. Almeida, A. Dress,T. Kühne, and L. Parida

Dagstuhl Manifestos
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Executive Summary

"Water, water, everywhere, nor any drop to drink."

So goes Coleridge's *Rime of the Ancient Mariner*. Until recently, the same went for data: everywhere, but not of much use so far, neither for deriving new medical insights nor for improving medical care.

However, three key developments currently help to overcome this problem: the rapid adoption of electronic medical records [1], the dramatic advances in molecular biology [2], and, just as dramatic, the growing pervasiveness of social computing environments combined with a new attitude towards participatory health management [3, 4, 5].

The result is an exciting medley of initiatives devoted to supporting health-care related information flow ranging from patient-facing resources such as *PatientsLikeMe* [6] to initiatives such as *MD-Paedigree* [7] (EU's FP7) that provides a physician-centric sort of "*PatientsLikeMine*" analogue addressing treatment choices in paediatrics.

Managing the *creative deconstruction* [8] involved in advancing towards systems medicine requires fundamentally changing the use of ICT in both, health care and biomedical research. It requires in particular to take account of the new paradigm of *web-centric computing* which is a basic prerequisite for all these initiatives.

Reflecting these concerns, a Dagstuhl Perspectives Workshop on *ICT Strategies for Bridging Biology and Medicine* was held to discuss a wide range of fundamental and foundational issues. These ranged from architectural considerations to data-access policies including *Open/Linked Data, the Semantic Web, Pervasive Hardware Ecosystems, Medical Clouds, Patient-Participation Frameworks, "Health Care* 4.0*", Analytical Tools*, and *Medical Education*[1]. Clearly, the required changes can only be achieved by initiatives of a broader scale and scope than what can be accommodated within the existing academic organisations, and need to involve *all* stakeholders at every step of such initiatives. In response to these challenges, the discussions led to the following *theses and postulates*:

(i) An *open-data policy* for health-care related information systems is a fundamental and urgent imperative.
(ii) Following the *business-IT alignment* paradigm [9], health care should – on all levels – be supported by secure IT-platforms enabling clinical workflow engines that map health-care related processes while integrating pertinent data-analysis, visualisation, and engineering tools.
(iii) Such platforms should also take full advantage of advances provided by *cloud services*, *pervasive computing ecosystems*, and the *semantic web*.
(iv) The *participatory potential* of the Web should be exploited to advance new forms of partnership in the health-care environment.
(v) The acquisition of *ICT literacy* must become a required part of biomedical education.
(vi) Specifically in Germany, the Bundesnetzagentur should be encouraged to set up a Working Group *Medizinische Netze* to explore options for a *Medical Cloud* within the German health-care environment while for instance in China the new *National Health and Family Planning Commission* might be the correct organisation to approach.

[1] The raw notes collected during the Dagstuhl Perspectives Workshop can be found at http://bit.ly/dagict.

Table of Contents

Executive Summary 32
Introduction 34
The need of adequate ICT architectures 36
An open-data policy for the health-care community 38
Patient participation 39
Creative deconstruction 40
The need of a medical cloud 41
Pervasive systems for medical ICT 42
Process modeling for “Healthcare 4.0” 43
Analytical tools 43
Medical education 44
Overview of the Program 44
Final Remarks 47
Participants 47
References 48

1 Introduction

It is a well-known and much deplored fact that, despite of great efforts, much of the overwhelming current progress in bio-medical research does not easily find its way to the bedside. While tools for generating ever better information regarding the individual makeup of a patient are becoming available, the sheer amount of data these tools produce soon becomes overwhelming and cannot easily be interpreted and processed for clinical application by medical practitioners, let alone assist patient engagement within the Health System. And extending medical education further and further has its limits as nobody can encompass and comprehend the onslaught of relevant new medical insights coming along year by year.

However, there is a growing international consensus that this can be mended using modern process-oriented and integrative ICT-platform solutions including specifically designed *cloud-based and service-oriented architectures*, thus establishing a context-aware computing service infrastructure (as anticipated in [10]) for supporting clinical workflows and connecting bioscience and health care.

The recent report of the National Academies of the USA on *Precision Medicine* [2] provides an excellent starting point for this discussion. It suggests the creation of a framework for a *Knowledge Network of Disease* (see Figure 1) which could integrate the rapidly expanding range of information on causes and expected progress of diseases. This solution has a number of key advantages over the current insular medical-institution-centric model as it would allow researchers, health-care providers, and the public to share and update their information. It asserts: *As well as improving health care, the new data network could also improve biomedical research by enabling scientists to access patient information through electronic health records, while still protecting patient rights. [...] Dramatic advances in research have generated a wealth of new data that could improve health outcomes. However, currently there is a disconnect between scientific advances in research and the incorporation of this information in the clinic. In addition, researchers don't have access to the wealth of clinical data on patients that is collected at the point of care. In order to harness the power of emerging disease data, systems are needed to collect and make the information widely accessible.* In a nutshell, a comprehensive ICT solution for Health Care must link and aggregate multiple data sources and provide advanced governance models enabling multiple, potentially disparate, or secondary data usage for operational and research tasks without ever compromising patient rights.

In Europe, the ITFoM initiative [11] states quite dramatically, that currently, a new spectre is haunting Europe and needs to be addressed, the spectre of "aging": *Europe ages. Its health budgets soar. Europeans suffer from ill-managed diseases. And all of this against a background of extremely successful life sciences with little societal impact, and revolutionary ICT that seems to affect all areas of society except for medicine. What will Europe do? Wait until the problems disappear? Wait until solutions originate in other countries? Or realise that it is best posed to provide a solution; a revolutionary ICT that revolutionises medicine.* And reinforcing ITFoM's expectation of a critical role for ICT in this domain, Google announced soon after the Dagstuhl workshop that it was starting a new company, Calico, with the broad goal of tackling *the challenge of aging and associated disease* (cf. [12]).

Such insights led to proposing a Dagstuhl Perspectives Workshop on *ICT Strategies for Bridging Biology and Medicine* that then took place in August 2013 and brought together a highly motivated group of leading researchers from academia, research institutions, and industry comprising physicians, medical researchers, biologists, mathematicians, computer scientists, and information–systems experts from 3 continents. Together, they exchanged ideas,

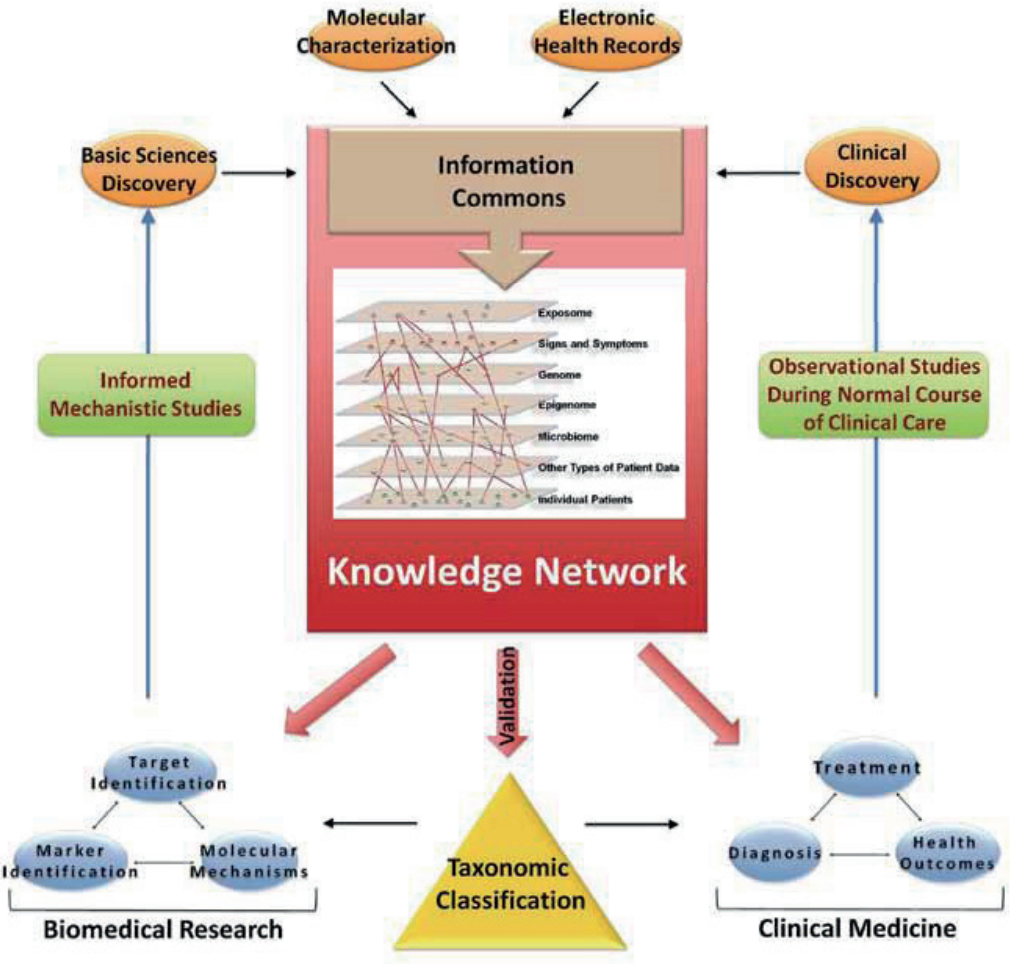

Figure 1 A diagram outlining the *Knowledge Network for Biomedical Research and a New Taxonomy of Disease* as envisaged in [2].

(Reproduced with permission from "Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease", 2011, by the National Academy of Sciences, Courtesy of the National Academies Press, Washington, D.C.)

discussed key challenges, identified urgent tasks as well as long-term strategic opportunities, and explored how integrative ICT architectures can be used in medical care to bridge the widening gap between biomedical research and daily clinical routine. In particular, they discussed the pros and cons of, on the one hand, the more data-driven bottom-up strategy for web-based Big-Data exchange and annotation favoured by the *Semantic Web community* as represented by participants from W3C and from the *Digital Enterprise Research Institute* in Ireland, and on the other hand, by the more model-driven top-down strategy favoured by ITFoM.

In these discussions, it was a point of strong consensus that, as in many other areas that are undergoing a *data science revolution* [13], the development of ICT for the modern health-care environment depends critically on the availability of an *open-data infrastructure* sustaining an information society capable of engineering a knowledge economy [14]. Enabling the Web to form a global and secure space of linked data was therefore seen as a prerequisite for

- managing the creative deconstruction required to fundamentally change the use of ICT and advancing towards systems medicine by creating IT-platforms enabling clinical workflow engines mapping health-care related processes,
- integrating pertinent data-analysis, visualisation, and engineering tools,
- modelling and visualising the spatio-temporal organisation and dynamics of life and disease,
- and unfolding the potential of patient participation employing pervasive hardware ecosystems and medical clouds.

Of course, there are also many ethical, legal, societal, and political issues related to medical ICT that the workshop was not able to cover. Nor did we feel in a position – mainly in view of the strong heterogeneity of the group of participants – to devise a detailed *road map* of further progress going beyond the recommendations reported already above (i.e., our six "*theses and postulates*").

Instead, the workshop's most crucial insights have been detailed below, beginning with a discussion of critical role of adequate ICT architectures and an open-data policy, and continuing with sections on patient participation, personalised systems medicine, the need of a medical cloud, pervasive systems for medical ICT, process modeling for "Healthcare 4.0", analytical tools, and medical education.

2 The need of adequate ICT architectures

While the emergence of genomics and other *omics* technologies as well as many other data sources (imaging [15], sensoring [16], etc.) have specifically highlighted the critical need for adequate ICT frameworks, ICT plays a decisive role in a much wider range of processes involved in *acquiring*, *managing*, and *processing* all sorts of biomolecular and medical data.

Of course, ICT cannot make clinical decisions. It can only facilitate the access to medical information and support clinical workflow processes. The decisions still have to be made by the physicians (together with their patients) based on their clinical experience as well as information from all relevant sources – including systems-medicine data. ICT should help them to utilise the available data and their clinical experience most productively, to identify suitable experts, and to provide a platform for efficient communication between the concerned parties. That is, adequate ICT architectures must provide medical researchers and practitioners with means to easily and securely

- exchange their views regarding new biomedical developments,
- obtain, annotate, exchange, and interpret their patients' data in the light of the latest biomedical insights taking into account all that can be learned from the various relevant databases and bioinformatics tools the web provides,
- check data consistency and quality (cf. [17], see also [18]),
- specifically search for, collect, and quickly comprehend the most relevant insights exactly when needed and within the context of an efficient clinical workflow,
- derive pertinent predictions regarding their patients' fates on the basis of these data,
- and to discuss the implications of such insights for their clinical work with colleagues, patients, and health managers,
- while simultaneously making change and improvement a daily routine.

Adequate ICT architectures should also be open to involving other stakeholders in the medical environment like patients, nurses, researchers, managers, health-care insurers, or policy makers.

To this end, ICT must establish *network services* supported by associated *workflow engines* and *computing platforms* that enable

- systematic de-coupling of services,
- proper assignment and observance of rights and roles,
- task-dependent dynamic modification of rules, roles, and rights,
- activity monitoring and object management,
- assured security, e.g. by providing cryptographic protocols like *Transport Layer Security* (TLS) or *Secure Sockets Layer* (SSL), and
- mediation of conflicts[2].

However even today, meeting these demands is not as easy as it may sound. Indeed, the current state of the art is so dire that, for instance, the adoption of Electronic Health Records, spearheaded by the ONC, the Office of the National Coordinator for Health Information Technology [20], has at times been felt to decrease – rather than enhance – the effectiveness of care delivery. E.g., it is deplored for instance in [21] that *doctors become increasingly bound to documentation and communication products that are functionally decades behind those they use in their "civilian" life and that most EHR vendors, in order to protect their prices and market share, don't even embrace existing modular architectures with interfaces that allow extension of product capabilities, innovative uses of data, and interoperation with other software.*

And young medical professionals confessed in private talks that they are already dreaming of a *de-digitalisation* of medicine as they tend to see ICT in hospitals as a curse, not a blessing for medical care. So, there is an imminent danger that outdated , if not incompetent design of medically oriented ICT will actually be counter-productive, not enhancing the application of information technology in medicine, but stand in its way.

In particular, as *time* is still the most precious resource also in medical care, computer scientists need to find out how to design intelligent ICT so as to simplify the practitioners' professional life, not to overburden it with ever more additional requests, and to save, not to waste their time. Research aiming at improved means of data collection, communication, and

[2] See e.g. [19] for a discussion of divergent innovation priorities among the stakeholders in the medical environment.

visualisation along with more transparent user interfaces for health practitioners, patients, and other stakeholders as detailed for instance in [22] should, therefore, be a key priority.

Cautionary tales also abound in the realm of *biomedical informatics*. For example, the ICT associated with *The Cancer Genome Atlas* and the *International Cancer Genome Consortium* has, in spite of the tremendous value delivered by the data it contains, not proven to be as helpful as was expected. And, much more worrisome, the *Cancer Biomedical Informatics Grid* initiative was explicitly found to having directed hundreds of thousands of dollars to ICT developments that, in fact, did more harm than good [23, 24].

3 An open-data policy for the health-care community

As mentioned above, while ICT is being quickly commoditised and Big Data already pervade all ICT arenas of economic, medical, and research activities [25, 26], creating

- flexible open-data structures,
- linked open-data (LOD) web-technologies (as demanded e.g. by Sir Tim Berners-Lee in his celebrated call for *Raw Data Now* [27, 28, 29]),
- and transparent data-analysis tools delivering comprehensible, reproducible, and reliable results [22, 30, 31, 32]

is becoming a top priority in the context of medical ICT.

It is important to clarify that the call for an open-data policy is a call for open standards and interoperable systems. It is by no means a call that private data be forcefully made public. Yet, not adopting an open-data policy for public data was seen in this workshop as nothing short of unacceptable. More specifically, as emphasised for instance by William Hersh in his slide file *Translational Research: The Essential Role of Biomedical Informatics* [33], the required open-data policy for health-care related information systems includes:

- an open development of tools and infrastructure created by an open, participatory process,
- open access to resources that should be freely obtainable for the health-care community,
- and an open source code that is available to view, alter, and redistribute.

Health care is now seriously lagging behind other data-intensive activities in its ability to realise the benefits of data integration and secondary use of publicly funded data resources.

In conclusion, at a time when the web is evolving to become a global data space, the request of an open-data policy in health-care related information systems was therefore one of the most urgent and most consensual demands expressed by the workshop participants. More specifically, adopting the World Wide Web Consortium's *Resource Description Framework* (RDF) [35, 36] was seen as a particularly helpful strategy to achieve the goals of this policy.

Similar consensus was found with regard to

- the adoption of public-data dissemination policies – *with moratoria not longer than 6 months*
- and the request that any future publicly funded large-scale research project should be required to only use – or re-use – infrastructures that are scalable and can manage Big Data.

4 Patient participation

Clinical and molecular stratification of disease in patient populations [37], combined with the commoditisation of ICT, has empowered patients in unprecedented ways. The disruptive nature of this process, epitomised by initiatives such as PatientsLikeMe [6] or Leroy Hood's P4-medicine initiative [38, 39], is further deepening the integration of ICT into health-care activities and the health sciences. The control of this process is increasingly distributed and requires a new partnership in the health-care environment in order to advance and exploit – rather than to (hopelessly) curtail – the participatory nature of the system.

Such developments need to be pursued more aggressively if the health sciences – and health care in general – are to catch up with the fast-paced societal changes associated with the emergence of third-generation web and social computing technologies. And as ICT is increasingly provided to patients via new platforms such a mobile apps, this is in fact a development which, while inevitable, is clearly offering a promising route to make patient participation a key feature in improving outcomes and decreasing costs – a fact that is just starting to be recognised in academic and medical institutions.

However, adequate participatory systems do not yet exist, and just installing an open-data policy – though crucial as it is – will not be enough to bring such systems into existence. To encourage and support participation and cooperation, any open-data policy has to be augmented by an overall redesign of current health information systems constructing and maintaining systems that properly assign rights and roles, thus making participation and collaboration simple and secure while respecting the sharing preferences of the involved patients.

Yet, unlike the Open Data recommendation where even the technical details were clear, the best route towards patient participation is uncertain. For example, ownership of one's own data is seldom supported by current ICT in health-care related information systems.

Therefore, a new way has to be found. In consequence, a number of (not at all mutually exclusive) options were discussed by the participants:

- integrating all local and more or less isolated health-care related information systems within a shared, secure, and well-maintained cloud environment,
- developing software for merging data of different structure and from different sources while still guaranteeing data quality [17, 18], security [34], and privacy [40],
- and deploying systems that assign rules, roles, and rights in a dynamic way allowing for continuous modifications (see e.g. [41]) and for managing and resolving emerging conflicts as an integral part of the system.

The pursuit of each of these routes is fraught with practical difficulties. For example, managing change and dynamic modifications of rules, roles, and rights in the ICT landscape implies obtaining and maintaining patient consent for data sharing. However, obtaining informed consent is itself often a primitive process that often discourages the use of ICT tools – even by patients who overwhelming express a desire to share their data. Assuring the privacy and security of data, once shared, is even more complex as patients express great concern about subsequent misuse [34]. Patient access to health data is also impeded by complex and often inconsistent terminologies, and further complicated by the lack of effective mechanisms to validate agreed governance policies.

The lack of specific recommendation beyond the encouragement of patient participation is in stark contrast with the growing adoption of document-sharing platforms such as DropBox and Google Drive, or even Facebook. These user-facing tools have a mass appeal and find

an increasing use in health-care environments – often in contravention of explicit policies disadvising them sternly. Judging from the strong movement towards *client-facing ICT* in other data-intensive domains such as banking, transportation, and retail, the most effective recommendation here may well be that already the design of participatory medical ICT needs the participation of patients and patient organisations.

Nevertheless, there may be instances where the pursuit of patient participation leads to results that are the opposite of those intended. For example, special attention has to be given to the elderly as well as to the computer illiterates to avoid their discrimination and adding to health disparities that already exist.

5 Creative deconstruction

Managing the *creative deconstruction* [8] involved in advancing towards personalised systems medicine requires a fundamental change in the use of ICT in both, health care and biomedical research and modelling.

We are all different. Diseases are different. We react differently to the same therapy even if we have the "same" disease. Only an ICT-driven big-data medicine that takes advantage of the vast amount of personalised and research data available can help to take this individuality into account, providing better therapies at ultimately lower cost.

Indeed, the costs of sequencing a person's genome have dropped by close to a million fold in a bit more than a decade. Moreover, the needed information and communication technology is rapidly becoming inexpensive enough for routine clinical use. A modern tablet computer, while costing only a few hundreds of dollars, now has the power of a high-end computer of only a few years ago. And their power continues to grow by roughly three orders of magnitude every decade.

In consequence, we have a chance to change the lives of many by redesigning the mechanisms in health-care delivery to take full advantage of the enormous power of new computational resources and analytic tools as detailed below.

However, despite this exciting promise, many processes in medicine have hardly changed over long periods of time. If we bring our car to the garage for a service, the mechanic will be able to analyse megabytes of information, allowing a much more 'personalised' correction of problems than we are likely to receive from our doctors to whom typically only a minute fraction of the information is available compared to the information the mechanics have on our car. It is time

- to think ahead which effect these continuing changes of many orders of magnitude could have on health-care systems,
- to adress their enormous potential for health care and biomedical research and modelling,
- to consider radical changes as we face radical challenges in ageing societies, and
- to plan ahead for the future rather than to continue reacting after the fact with incremental adjustments – always too little too late – to these predictable developments once they have happened. Our lives as well as even the financial and social stability of our societies might be at stake.

The workshop participants also agreed that the required changes can only be achieved by initiatives of a broader scale and scope than what can be accommodated within existing academic organisations while the fact that most health care is delivered in community

practices, not in academic institutions, was seen as a decisive motivation for always involving medical practitioners and other stakeholders in the medical environment at every step in all such initiatives.

Depending on local conditions, this could be pursued by initiating pertinent calls for tenders, setting up appropriate working groups, or establishing suitable new institutes, cross-institute initiatives, or national or international branches of ICT or health-care related organisations.

Such organisations do exist already in the US: the *Office of the National Coordinator for Health Information Technology* (ONC) is already mandated *to support the adoption of health information technology and the promotion of nationwide health information exchange to improve health care* [20] (though currently challenged to take the next step: fitting EHRs into a dynamic, state-of-the-art, rapidly evolving information infrastructure rather than jamming all health care processes and workflows into constrained EHR operating environments [21]). And the National Institutes of Health's autonomous *Big Data to Knowledge* (BD2K) initiative [26] might also be a good model to follow.

In Germany, a first step could be to contact the *Gesundheitsforschungsrat* proposing to set up a "Arbeitsgruppe Medizin" within the framework of its *Bundesnetzagentur*, involving also the newly founded *German Centers of Health Science* (Deutsche Zentren der Gesundheitsforschung DZG, cf. [42]).

And in China, authorities on many different levels could get involved with such activities, from the federal level like e.g. the new *National Health and Family Planning Commission* to provincial and prefectural level down to the level of counties and townships.

Such initiatives could also offer a fast route to incorporating the advances of modern medicine in daily health care in third-world countries.

6 The need of a medical cloud

Remarkably, the new paradigm of *cloud-computing* seems to be able to address many of the requests regarding e.g. the performance of adequate ICT architectures, the introduction of open data policies, and the support of patient participation. A broadband-supported interoperable public cloud infrastructure distributed between a multitude of nodes connected by an appropriately designed enterprise service bus [47], could offer "abstract" services (on request and according to well-defined protocols)

- to rapidly and securely store, access, annotate, process, interpret, exchange, and discuss immense amounts of data stretching from family-health histories to health-event logs, medical image data, and genomic information,
- to inspect medical and life-science databases,
- to enable effortless, yet well-protected data management and data communication to support e.g. joint expert online-discussions regarding specific medical implications in the light of newest medical insights,
- to introduce transparent governance models that do not rely on "security by obscurity" and do not compromise interoperability or involve unscalable formalisms,
- and, thus, to escape the current EHR trap deplored in [21] and to regain control of the configuration of EHR infrastructures that most medical professionals are currently only allowed to manage,

The reliance on cloud services for biomedical *research* is now a tried and tested possibility, as illustrated by the research-driven Europe-wide Elixir network [43].

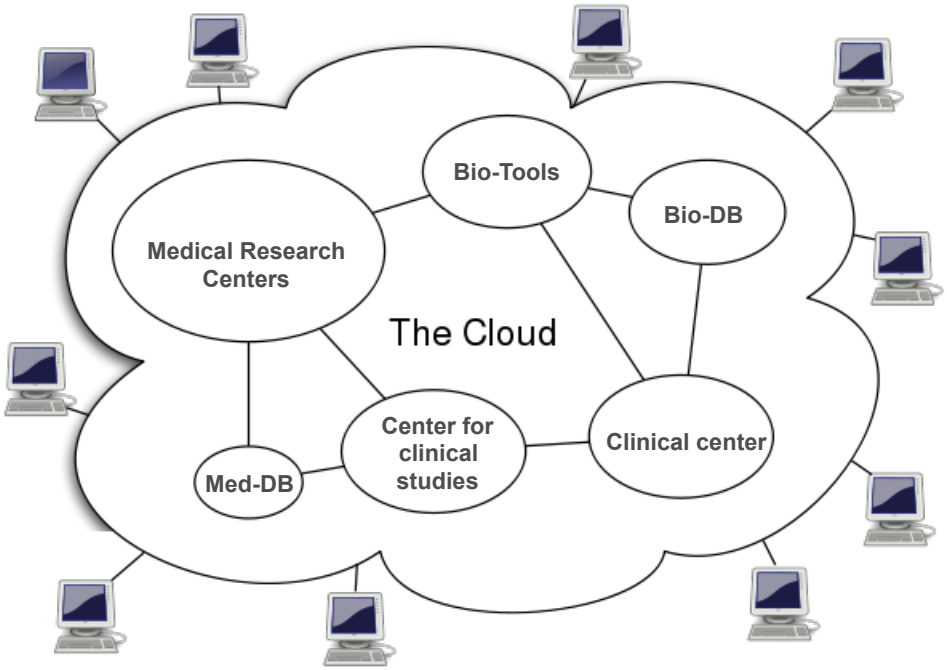

Figure 2 A cartoon of a medical cloud connecting stakeholders in the health-care environment, medical and bio-data bases, and analytical tools (as described e.g. in [44, 45, 46]).

There are legitimate concerns about assuring privacy, security and trust in health data stored in the cloud. They can and must be addressed and they should be object of continued research and advancement.

Nevertheless, the time to initiate the development of cloud services in support of *health-care delivery* is now.

7 Pervasive systems for medical ICT

Pervasive computational systems are among the most recent, promising, and disruptive developments in medical ICT. They range from the usage of standard laptop computers, tablets, terminals, and phones to more specific home-based web appliances (such as wireless weight-scale and blood-pressure measuring devices) and webApp ecosystems (browser hosted software applications – a topic of the next *Conference on Semantics in Healthcare and Life Sciences* CSHALS 2014). There is a mounting profusion of web-connected sensing devices in varying stages of development. Some are ubiquitous in their distribution, some are wearable mobile devices, and still others are intended for use only in the home.

This is an area where research towards a systematic approach to their deployment and integration is only taking its first steps. Early results suggest that this is one of the most important innovations for the future health care systems. It will have an impact that spans from novel opportunities to deliver care at the patients' home, to laboratory medicine that is not only integrated, but in fact involved in the choice of therapy. The transformative nature of these opportunities suggests that the work towards the development of pervasive ICT for health care systems needs to be specifically targeted by research programs.

In a recent survey regarding the *Grand challenges in interfacing engineering with life sciences and medicine* [48], pervasiveness and pre-emptiveness were described as the missing "P's" in the engineering of P4 Medicine. Accordingly, also the development of *Biomedical Systems Engineering Programs* was proposed to be of critical strategic value.

8 Process modeling for "Healthcare 4.0"

The intensive monitorisation, formal description, and exhaustive representation of clinical workflows and other medical processes are the collective blind spot of present-day health sciences. This is not so much a problem that is specific to medicine as it is a reflection of an early stage in the development of industrial processes capable of coping with complex systems. In Germany, the development of high-tech "smart processes" was dubbed *Industrie 4.0* [49] and described as a new (the fourth!) industrial revolution. Here, a plethora of manufacturing processes reacts to specifications characterised by strong customisation of products under the conditions of highly flexibilised (mass-)production. The required automation technology is improved by the introduction of methods of self-optimisation, self-configuration, self-diagnosis, cognition, and intelligent support of workers in their increasingly complex tasks.

This is also seen as an apt description of what is desirable for the future of ICT in the medical environment. Following the business-IT alignment paradigm [9] mentioned already, health care should – on all levels – be supported by secure IT-platforms enabling clinical workflow engines that map health-care related processes while integrating pertinent data-analysis, visualisation, and engineering tool. The key advantage of this novel approach to process specification is that it promises to remove the intrusive nature of current health-related ICT that often disrupts, rather than assists, the development of integrated and efficient care-delivery practices.

However, given the early stage of development of this new system, it still needs to be explicitly demonstrated, e.g. by comparative effectiveness research, that the envisaged new clinical-workflow platforms can be used easily and deliver improved clinical outcome and cost effectiveness.

9 Analytical tools

The key role played by data processing, modelling, visualisation, and disease- and patient-stratification tools in systems medicine is at the core of the personalisation of health-care delivery. The data-analysis pipeline is also where advances in medicine find their way to preventive medicine. Advances in mechanistic modelling exploiting the spatio-temporal organisation and the dynamic nature of life and disease for generating virtual-patient models (see e.g. [11]), in data mining (see e.g. [44, 45]), and in machine learning, as illustrated by its iconic use of IBM's Watson [46], are among the most promising developments in biomedicine.

On the other hand, this remains an area hampered by serious availability and reproducibility issues [30, 31, 32]. Interestingly, these are issues that reflect the lack of integration of present-day computational statistics environments rather than an inherent limitation of discriminant-analysis techniques (including machine learning). In other words, the main

barrier to a wider availability of advanced analytical tools is to be found in the ICT infrastructure, not the analytical methods per se.

Accordingly, the recommendation regarding analytical tools is that the data-processing environments be selected on the basis of their ability to take full advantage of advances towards open data (Section 3), pervasive systems (Section 7), and cloud services (Section 6).

10 Medical education

The last half-day session of the Perspectives Workshop was entirely dedicated to education and workforce issues. The ongoing advances in medicine, particularly the increasing sophistication of ICT and the growing role of patient participation, necessitate that the acquisition of appropriate ICT literacy must be supported in biomedical education, and incorporated as a required part into the teaching of medicine, nursing, and other health professions.

While no specific route has been suggested, it was noted that, in the US, Clinical Informatics is emerging as a board-certified medical subspecialty within Pathology [50]. Public Health and Genetics were also considered.

Also, an argument was made that ICT topics are best introduced at an early stage of medical training, particularly since they are already being added to primary and secondary education [51]. Medical education and workforce training in health-related ICT will have to be a core mandate of any initiative claiming leadership in this field. Ultimately, the goal is to equip and empower physicians, nurses, and other care providers to assist patients with the assimilation of what will otherwise be an overflow of information about themselves. Given the time it takes to realise investments in education and workforce training, the integration of ICT topics in medical programs is an *urgent* matter.

The acquisition of ICT literacy – including an education in the utilisation of "high-dimensional" data – is an important point not only in medical and biomedical education, but also for the public in general (comprising, of course, also *future* patients and medical practitioners, nurses, policy makers, etc.) and requires a more quantitative approach to the teaching of Science and Technology at the K-12 education.

11 Overview of the Program

During the workshop, the following topics were addressed in regular sessions, complemented with impromptu evening presentations and discussions. Rapports for each of them were produced and later collected at http://bit.ly/dagict:

1. **Information and Communication Technology for Bridging Biology and Precision Medicine – Chances and Challenges**
   - From web appliances to web services
   - From home medicine to cloud-based health information systems
   - From laboratory systems to reference genomic atlases
   - Architectures and APIs for user-governed ICT

   Coordinator: Jonas S. Almeida
   Contributors: Hans Lehrach, Wolfgang Maass, ...
   Rapporteur: Mark Braunstein

2. **Big and Heterogenous Data**
   - What can Big Data tell us?
   - Genomic Medicine
   - The integration and modeling of heterogenous multi-omic and imaging data from disease and its implications for diagnosis and therapy
   - Fusing bioimaging data with clinical and molecular information (for enhancing a systems view of disease)

   Coordinator: Joel Saltz
   Contributors: Klaus Maisinger, Stefan Decker, Scott Kahn, ...
   Rapporteur: Alex Pothen

3. **How can ICT help us learning more about disease mechanisms?**
   - Architectures and APIs for user-governed ICT
   - Medical Clouds as platforms for annotation, exchange, and joint interpretation of healthcare data by medical experts (and patients?)
   - Statistics, machine learning etc.
   - Electronic health records

   Coordinator: Bernhard Balkenhol
   Contributors: Eric Neumann, Eric Gordon Prud'hommeaux, ...
   Rapporteur: David Gilbert

4. **Virtualisation, Computation, and AI in Medicine**
   - The status of structured (pathway, model etc) databases
   - The virtual oncology, diabetes, ... patient in medical practice
   - Mechanistic models
   - The vision of ITFoM
   - How can we extract information from the scientific literature as well as from 'low-grade information' in the web (text mining, the semantic web in healthcare, search strategies in semantic webs)?
   - Virtualisation in drug development

   Coordinator: Hans Lehrach
   Contributors: Laxmi Parida, Pietro Lio', Joel Saltz, ...
   Rapporteur: Andrea Splendiani

5. **Molecular Systems Medicine and ICT I**
   - Assessing emergent properties of chronic diseases and disease mechanisms:
     – The dynamics of disease progression, implications for disease mechanisms
     – Parallel toponome decoding and genome sequencing
     – Cancer and the immune system
   - Family genome and tumor genome sequencing – the use of Next Generation Sequencing and its implications for therapy and disease stratification

   Coordinator: Peter Walden
   Contributors: Walter Schubert, Robert Burk, Markus Löffler, ...
   Rapporteur: Eric Gordon Prud'hommeaux

6. **Molecular Systems Medicine and ICT II**
   - A systems approach to diagnostics (including the use of proteins, mRNAs and miRNAs from blood and from tissue)

- Biomedical engineering and systems-optimisation strategies in medical care
- Assays for wellness
- Precision medicine and guidelines for evidence-based medicine: complementary or incompatible?

Coordinator: Ina Koch
Contributors: Anja Hennemuth, Helena F. Deus, Susana Vinga, ...
Rapporteur: Laxmi Parida

7. **The Stratification of Disease into Discrete Subtypes and its Implications for Science, 'Medical Ontology', Diagnosis, and Therapy**

   Coordinator: Markus Löffler
   Contributors: Mark Braunstein, Eric Prud'hommeaux, Peter Walden, ...
   Rapporteur: Susana Vinga

8. **Does the Potential of ICT in Medical Care Require New Forms and Levels of Medical Training, Medical Data Collection, Storage, and Reproducibility, Clinical Logistics, Clinical Trials, and Patient Participation?**

   Coordinator: Susana Vinga
   Contributors: David Gilbert, Jochen Dreß, ...
   Rapporteur: Ina Koch

Figure 3 The *word cloud* for the combined reports of all regular sessions and evening presentations, reflecting the central focus on open data, nucleating discussions about medical information systems, and disease modeling.

12 Final Remarks

We suggest that our recommendations be discussed, modified, and further improved by the Dagstuhl Scientific Directorate and appropriate committees of the GI to be then submitted to the pertinent authorities in governments and professional societies.

We propose in particular that research funding organizations are specifically encouraged to initiate specially funded programs for supporting technology and research projects pursuing these goals.

13 Participants

- Jonas Almeida
University of Alabama at Birmingham, USA
- Bernhard Balkenhol
infinity^3 GmbH – Gütersloh, DE
- Mark Braunstein
Georgia Tech – Atlanta, US
- Robert Burk
Yeshiva Univ. – New York, US
- Stefan Decker
National University of Ireland – Galway, IE
- Helena F. Deus
Foundation Medicine, Inc. – Cambridge, USA
- Andreas Dress
Shanghai Institutes for Biological Sciences, CN & infinity^3, DE
- Jochen Dreß
DIMDI – Köln, DE
- David Gilbert
Brunel University, GB
- Anja Hennemuth
Fraunhofer MEVIS – Bremen, DE
- Scott Kahn
Illumina – San Diego, USA
- Ina Koch
Goethe-Universität Frankfurt am Main, DE
- Titus Kühne
Deutsches Herzzentrum Berlin, DE
- Hans Lehrach
MPI für Molekulare Genetik – Berlin, DE
- Pietro Lio´
University of Cambridge, GB
- Markus Loeffler
Universität Leipzig, DE
- Wolfgang Maaß
Universität des Saarlandes, DE
- Klaus Maisinger
Illumina – United Kingdom, GB
- Eric Neumann
Foundation Medicine, Inc. – Cambridge, USA
- Laxmi Parida
IBM TJ Watson Res. Center – Yorktown Heights, USA
- Alex Pothen
Purdue University, USA
- Eric Prudhommeaux
MIT, USA
- Joel Saltz
Emory University, USA
- Walter Schubert
Universität Magdeburg, DE
- Andrea Splendiani
DERI – Galway, IE
- Marc Van Regenmortel
IREBS – Illkirch, FR
- Susana Vinga
Technical Univ. – Lisboa, PT
- Peter Walden
Charité – Berlin, DE
- Zhenbing Zeng
East China Normal University – Shanghai, CN

Acknowledgement. We thank all participants and, in particular, Mark Braunstein and Peter Walden for many valuable comments and suggestions regarding this text. We also thank the infinity[3] GmbH in Gütersloh for its generous support and very helpful scientific and technical advice regarding the potential of modern process-oriented and integrative ICT platforms.

References

1 Tracy D. Gunter and Nicolas P. Terry *The Emergence of National Electronic Health Record Architectures in the United States and Australia: Models, Costs, and Questions.* J Med Internet Res 7: 1 (2005). DOI: 10.2196/jmir.7.1.e3

2 Susan Desmond-Hellmann et al. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease.* National Research Council (US), Committee on A Framework for Developing a New Taxonomy of Disease. The National Academies Press, Washington D.C., USA (2011).

3 Wikipedia. http://en.wikipedia.org/wiki/Social_computing

4 Barbara A. Israel et al. *Community-based participatory research: policy recommendations for promoting a partnership approach in health research.* Education for Health, 14:2 (2001): 182–197.

5 Melanie Swan. *Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking.* Int. J. Environ. Res. Public Health 6 (2009): 492–525. DOI: 10.3390/ijerph6020492

6 Ben Heywood et al. http://www.patientslikeme.com/

7 Bruno Dallapiccola et al. http://www.md-paedigree.eu/. The European Commission.

8 Eric J. Topol. *The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care.* Basic Books, New York, NY, USA (2012).

9 Wim van Grembergen and Steven De Haes.*Enterprise Governance of IT: Achieving Strategic Alignment and Value.* Springer, New York Heidelberg Dordrecht London (2009).

10 Jason I. Hong and James A. Landay. *An Infrastructure Approach to Context-Aware Computing.* Human–Computer Interaction 16:2-4 (2001): 287-303. Special Issue: Context-Aware Computing. DOI: 10.1207/S15327051HCI16234_11, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.27.2757&rep=rep1&type=pdf

11 Hans Lehrach et al. www.itfom.eu Max Planck Institute for Molecular Genetics, Berlin, Germany (2012).

12 http://googlepress.blogspot.com/2013/09/calico-announcement.html

13 Chris Wiggins. *The Data Science Revolution.* http://www.mathaware.org/mam/2012/pdfs/DataScienceRevolution.pdf

14 Keith Smith. *What is the 'Knowledge Economy'? Knowledge Intensity and Distributed Knowledge Bases.* Discussion Papers from United Nations University, Institute for New Technologies, No. 6. (2002).

15 Wikipedia. http://en.wikipedia.org/wiki/Medical_imaging

16 Charu C. Aggarwal (Ed.). *Managing and Mining Sensor Data.* ISBN: 978-1-4614-6308-5 (Print), 978-1-4614-6309-2 (Online), DOI: 10.1007/978-1-4614-6309-2, Springer New York Heidelberg Dordrecht London (2013).

17 Anonymous. *Data Quality: High-impact Strategies – What You Need to Know: Definitions, Adoptions, Impact, Benefits, Maturity, Vendors.* Emereo Publishing (2013).

18 Wikipedia. http://en.wikipedia.org/wiki/Data_quality

19 Mattijs S. Lambooij and Marjan J. Hummel. *Differentiating innovation priorities among stakeholder in hospital care.* BMC Medical Informatics and Decision Making 13:91 (2013). DOI: 10.1186/1472-6947-13-91

20 http://www.healthit.gov/newsroom/about-onc

21 Kenneth D. Mandl and Isaac S. Kohane. *Escaping the EHR Trap – The Future of Health IT*. N Engl J Med 366 (2012): 2240–2242.

22 Alexander W. Zaranek et al. *Transparent Informatics: A Foundation for Precision Medicine*. Arvados Open Source Bioinformatics Platform for Genomics. https://arvados.org/. (2013).

23 NIH Board of Scientific Advisors' Ad Hoc Working Group. *An Assesment of the impact of National Cancer Institute's Cancer Biomedical Informatics grid (caBIG)*. http://deainfo.nci.nih.gov/advisory/bsa/bsa0311/caBIGfinalReport.pdf

24 Wikipedia. http://en.wikipedia.org/wiki/CaBIG

25 James Manyika et al. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, Washington D.C., USA (2011).

26 Vivien Bonazzi et al. *NIH BigData to Knowledge (BD2K)*, NIH, Bethesda, Maryland, USA (2013). http://bd2k.nih.gov/

27 http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html

28 Jim Hendler and Tim Berners-Lee. *From the Semantic Web to social machines: A research challenge for AI on the World Wide Web*. Artificial Intelligence 174.2 (2010): 156–161.

29 http://www.w3.org/DesignIssues/LinkedData.html

30 Carole Goble.http://www.slideshare.net/carolegoble/ismb2013-keynotecleangoble

31 Roger D. Peng. textslReproducible research in computational science. Science, 334.6060 (2011): 1226–1227.

32 Bill Howe. *Virtual appliances, cloud computing, and reproducible research*. Computing in Science & Engineering 14.4 (2012): 36–41.

33 William Hersh. *Translational Research: The Essential Role of Biomedical Informatics*. Available via "http://skynet.ohsu.edu/~hersh/pumch-translational.pdf".

34 Bradley Malin et al. *Technical and Policy Approaches to Balancing Patient Privacy and Data Sharing in Clinical and Translational Research*, NIH, Bethesda, Maryland, USA (2013). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2836827/

35 Frank Manola and Eric Miller. *RDF Primer*. Published at http://www.w3.org/TR/2004/REC-rdf-primer-20040210.

36 David Booth et al. *Yosemite Manifesto on RDF as a Universal Healthcare Exchange Language*. http://goo.gl/mBUrZ. Semantic Technology and Business Conference, San Francisco, USA (2013).

37 National Research Council (US). Committee on A Framework for Developing a New Taxonomy of Disease. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*. National Academies Press, Washington D.C., USA, 2011.

38 Leroy Hood and Stephen H Friend. *Predictive, personalised, preventive, participatory (P4) cancer medicine*. Nature Reviews Clinical Oncology 8.3 (2011): 184–187.

39 Leroy Hood et al. *Revolutionizing medicine in the 21st century through systems approaches* https://www.systemsbiology.org/blog-topics/p4-medicine. Biotechnology Journal 7.8 (2012) :992–1001. DOI: 10.1002/biot.201100306

40 Wikipedia. http://en.wikipedia.org/wiki/Medical_privacy

41 Wikipedia. https://en.wikipedia.org/wiki/s3db

42 http://www.bmbf.de/de/gesundheitszentren.php. Bundesministerium für Bildung und Forschung (2013).

43 Niklas Blomberg et al. http://www.elixir-europe.org/. EMBL-European Bioinformatics Institute (2013).

44 http://xbtransmed.com/

45 http://www.broadinstitute.org/gatk/

46 Eliza Strickland and Eddie Guy. *Watson goes to med school.* [2013 Tech To Watch]. Spectrum, IEEE 50.1 (2013): 42–45.

47 Dave Chappell. *Enterprise Service Bus.* O'Reilly Media (2004).

48 Bin He et al. *Grand challenges in interfacing engineering with life sciences and medicine.* IEEE Trans Biomed Eng. 60:3 (2013): 589–98. DOI: 10.1109/TBME.2013.2244886

49 http://www.bmbf.de/en/19955.php. Bundesministerium für Bildung und Forschung (2012).

50 Peter J. Tonellato et al. *A National Agenda for the Future of Pathology in Personalised Medicine – Report of the Proceedings of a Meeting at the Banbury Conference Center on Genome-Era Pathology, Precision Diagnostics, and Preemptive Care: A Stakeholder Summit.* American Journal of Clinical Pathology 135.5 (2011): 668–672.

51 Valerie Barr and Chris Stephenson. *Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community?* ACM Inroads 2.1 (2011): 48–54.