International Journal of Economics and Business Modeling
ISSN:0976-531X & E-ISSN:0976-5352, Volume 3, Issue 3, 2012, pp.-192-199.
Available online at http://www.bioinfo.in/contents.php?id=37

A COMPUTATION SERVICE CENTERED BUSINESS MODEL FOR CLINICAL DIAGNOSTICS BASED ON IMAGE ANALYSIS OF MICROSCOPIC DATA

SCHMITT E.[1,2]*, FALKNER J.[3], KEPPER N.[1], WEISBECKER A.[3] AND HAUSMANN M.[1]

[1]Kirchhoff Institute of Physics, University of Heidelberg, Heidelberg, Germany.
[2]Institute of Numerical and Applied Mathematics, University of Göttingen, Göttingen, Germany.
[3]Fraunhofer Institute for Industrial Engineering IAO, Stuttgart, Germany.
*Corresponding Author: Email- eschmitt@kip.uni-heidelberg.de

Received: February 16, 2012; Accepted: September 06, 2012

Abstract- Parameters like the number of genetic elements or their position within the cell nucleus can give valuable hints for diagnosis or cancer treatment. They are determined by commercial or specially designed hybridization markers which are applied in a biochemical procedure to tissue or blood samples. Analysis of microscopic images reveals the necessary data, possibly after a statistical evaluation. We design a business model for one or more service providers in the biochemical, optical, or informatical branches. As this model is very flexible, it offers several possibilities for realization. We have implemented all experimental and computational steps of the model, some in several variants and simulated the whole diagnostic pipeline as a proof of principle, especially for the breast cancer gene Her2neu. We also point out the possible problems which arise in the context of medical data handling, computational algorithm licensing, monitoring, accounting, and billing. It is therefore still a challenge to organize the whole diagnostic pipeline within the framework of our business model.

Keywords- Workflow, Cancer Diagnostic, Fluorescence in situ Hybridization, Healthcare, Grid and Cloud Computing, COMBO-FISH, Cell Nucleus Architecture, Genomic Aberration

Citation: Schmitt E., et al (2012) A Computation Service Centered Business Model for Clinical Diagnostics Based on Image Analysis of Microscopic Data. International Journal of Economics and Business Modeling, ISSN:0976-531X & E-ISSN:0976-5352, Volume 3, Issue 3, pp.-192-199.

Introduction

Many modern methods in today's medical diagnostics are based on the interpretation of the status or the change of cellular parameters. One prominent example is the observation of nuclear parameters like the number of genetic elements or their nonregular duplication or repositioning within the cell nucleus [1,2]. The observation of such indicators is accomplished by hybridization techniques which bind specific fluorescing probes to specific DNA sequences [3]. The fluorescence signals are detected by 2D or 3D microscopic imaging [4]. The images are stored in appropriate file formats and the parameters under consideration are extracted by manual or automated image analysis [5]. The visual inspection of such images by the expert or statistical analysis of series of images then provides information which is utilized in the diagnostic process [6]. Nevertheless, maybe owing to the fact that hybridization targets and methods are specific to the diagnostic goals, the technique has not yet become a standard procedure in medicine [6]. It could therefore be helpful to design organization models for routine implementation.

It is one of the foremost interests of the patient and the medical doctor to establish a secure and safe pipeline that guarantees a fast and on the other hand cheap procedure to accomplish the diagnostic tasks. To analyze the process to be mapped into a business model [7], one has to be aware of the fact that there are completely different services engaged. The first step is concerned with the hybridization of appropriate probe material to the DNA of the cell material of the patient, which is a biochemical laboratorial procedure. In the second step, microscopic images are taken from the biological specimens, which is an opto-physical task. Then the images are evaluated by computer analysis using storage and computation power which has to be provided by informatical sources. The results of this last step are then transferred to the user, the medical doctor, who will decide on diagnosis and therapy based on the measured parameters in concert with his expert's knowledge and experience. It is therefore obvious, that experts and companies with completely different profiles will be involved in the design of a business model that realizes this pipeline of processes.

Presently, this method of diagnosis and monitoring of therapeutic success is being implemented in part in many clinics for a lot of different types of diseases. One of the foremost examples is cancer treatment. Yet, instead of using a unified business approach, the distinct steps of the diagnostic pipeline are performed by differ-

ent organizations, and in the end it is medical customers' duty to care for the overall supervision of the process. A pharmacological company sells specific marker sets which are hybridized to the tissue or blood samples by the clinic or a company receiving the sample by a shipping process. In principle, the same or a different company or the clinic will provide the microscopic images which are analyzed by experts in a usually semi-automated way. This seems to be a conceivable procedure for most standard diagnostic applications, the pathway being supervised by the clinic administration. But new hybridization methods, more elaborate and computation intensive image analysis techniques, and the demand for individualized medicine call for more efficient and professional processing of the diagnostic pipeline. We therefore design a business model that integrates the distinct steps and allows for adaptation to individual therapeutic cases. To define customers and providers and their relations, we can draw on the existing pathways which are extended to a combined business model. To this end, for the computational services we propose the use of distributed computing systems to cope with the constantly changing demands. Image analysis and extraction of the diagnostic parameters rely on this service, while microscopy is a prerequisite for it. Finally, it depends strongly on the individual application to find a way to achieve an agreement on the method of accounting mapping the activities of the different providers onto the final billing method.

Basically, the business model will involve a biochemical component for the laboratory steps, a physical component for the imaging, and an informatical component for providing computation power and storage space for image analysis. We will discuss the different components and characterize different possibilities to realize such a business model. It will furthermore become clear that the problems concerned with data security, data storage and transmission that arise in the medical context are not trivial.

To have a practical example at hand, we concentrate on a special hybridization method [3], called FISH (Fluorescence In Situ Hybridization), which is well established in laboratories around the world. The method is used widely in breast cancer diagnosis [8,9]. Furthermore, it allows a variation called COMBO FISH (COMBinatorial Oligonucleotide FISH) [10] which can be used within the framework of the business model developed, but can also be applied for individuum focused diagnostic features thereby extending the range of the business model. In fact, depending on an additional computational feature involving genome combinatorics [11], it can be the first step into personalized medicine in breast cancer diagnostics [12,13]. The image acquisition procedures and the informatical requirements for image analysis for both applications of computational services, standard FISH and COMBO-FISH, make it necessary to consider grid or cloud computing facilities [14,15].

Besides formulation of the business model, we will focus on its implementation. Finally, we will investigate its limitations and problems which arise on the different levels from user to provider.

The Diagnostic Pipeline

The manuscript should be divided as: In many medical applications, numerical and geometrical properties of cell nuclei and genetic elements, which can be measured and characterized easily by hybridization experiments, are important parameters for decisions in routine medical diagnostics and therapy [16]. Genetic alterations like multiplication of genes or parts thereof or even the formation of new chromosomes as well as changes of the micro- and nano-architecture of chromosome regions or gene compaction give valuable hints for medical decisions [17]. Also, parameters describing the nucleus and the geometry of chromosomes like distances of or angles between genomic elements or form and roundness of the nucleus can change due to the influence of disease. In general, they are evaluated by statistical methods on the basis of a large number of cells objected to a hybridization experiment [see e.g. 18-22].

Therefore, fluorescence labeling techniques are amongst the most prominent standard tools in medicine on the laboratory level and are employed for diagnostics and for therapy monitoring. The standard method for labeling genetic elements on chromosomes still is Fluorescence In Situ Hybridization (FISH) [3,23]. Here, one DNA probe of several hundred or thousand bases binds to a single DNA strand. The probes necessary for such a diagnostic essay can nowadays be bought from several companies throughout the world. The diagnostic process then consists of three steps. First, the hybridization of the marker sequences to the nuclear DNA has to be performed by a laboratory technician, which is a biochemical procedure using a standardized protocol. Second, images of appropriate resolution have to be taken by microscopic methods which are appropriate for the parameter detection in question. In a third step, the images have to be evaluated by special image analysis programs and the results of the image analysis, usually after statistical evaluation, have to be interpreted in the medical context.

To achieve a more precise and DNA target focused labeling, a slightly different method has been and is being further developed, namely COMBO-FISH [10,24,25], which enables labeling of short single stranded as well as double stranded DNA sequences. For a diagnostically meaningful labeling, several oligomeres of 15 to 30 bases length are employed which colocalize exclusively at the genetic spot to be marked [25]. Using short oligonucleotides has several advantages compared to standard FISH, but to fully exploit these advantages, one has to solve an additional combinatorial problem, namely to find an appropriate set of short oligomeres which bind in the desired region but do not form clusters anywhere else in the genome [11].

The three different steps of the diagnostic process form independent categories within our business model. The first step consists of a standard laboratory procedure which has to be performed by a skilled technician. Usually, these procedures are under supervision of the clinic itself, but as there is a variety of different applications, e.g. to blood cells, bone marrow smeares, or fixated specimens, etc., the task of hybridization can also be transferred to an external laboratory which then becomes a provider in the business model.

In the case of application of COMBO-FISH, there is another extension of the first step possible: Either the COMBO-FISH probe is a standard probe ready at hand, or it has to be designed for the special purpose. In this case, a selection of an appropriate COMBO-FISH probe set is the prerequisite for the individual COMBO-FISH experiment. First, a starter set of sequences of admissible oligonucleotides is selected from the target sequence. For triplehelical COMBO-FISH [10,25], which uses only short subsequences containing only two of the four possible bases, one can restrict the

search on oligomers of this type. Based on a previous analysis of the whole human genome, all such subsequences have been extracted and stored in a special data base which is scanned for subsequences of lengths of 15 to 30 base pairs on the respective genetic element. The occurrence of all these sequences in the whole genome is then analyzed for undesirable colocalizations which are removed by excluding clustering sequences from the COMBO-FISH starter set. For double helical COMBO-FISH, any oligomer sequence can be used for labeling [26]. Consequently, the search has to be performed on the whole human genome data base [27]. In any case, the final probe set will then be used for labeling the desired genomic locus [25,28,29]. For standard applications in diagnostics, e.g. the prognostically important genes in breast cancer events, the respective set has been determined in advance and is used in a standard protocol optimized for routine application [30]. But for personalized medicine or research applications, the search is a first step in the diagnostic pipeline and therefore a separate part in the workflow of our business model as an informatical prerequisite within the first step, which is essentially a biochemical one.

In the second step of the diagnostic pipeline, images of the hybridized specimens are taken with different microscopic systems depending on the diagnostic questions to be answered, which determine the parameters to be detected. Different microscopic systems provide different optical resolution. The standard approach is using a confocal laser scanning microscope [31], whereas detailed information on the nanoscale is acquired by high resolution far-field microscopes [29]. For the special application of COMBO-FISH, we prefer high resolution fluorescence microscopy, e.g. Spatially Modulated Illumination (SMI) microscopy [18,19,28] for size measurements of nanotargets, or localization microscopy [32,33] like Spectral Precision Distance/Position Determination Microscopy (SPDM) for the elucidation of the nano-architecture of genome targets [18, 34,35]. All these imaging methods produce image stacks of some tens of image sections through hundreds up to thousand sections of one 3D nucleus, which are stored in different file formats representing 3D stacks or time sequence stacks (4D) of experimental preparations of one to several cells or parts of cells [36]. Due to the different data formats delivered by the microscope software the data stacks usually have to be transformed into the file formats used by the analysis programs applied in local, cluster, grid or cloud computation [31].

The images are analyzed in the third step which is a purely computational task with possibly high throughput requirements for the analysis and the needs for large storage capacities for the image files. In a first segmentation step, eventually using threshold functions, cells and their parts like nuclei and the labeled genomic areas are identified [37,38]. The interesting numerical and geometrical parameters and the exact location of labeled genetic elements within a single nucleus are determined [39]. In addition, their center and border distance can be computed and other geometrical relations like size, shape, structure etc. can be extracted [20,21]. This includes also more complex measurements like angles between three marked regions or global parameters as volume, surface and roundness of labeled target sites, nuclei or cells [40]. With appropriate microscopic setups mentioned above (SMI, SPDM), microlocal parameters as gene volume and compaction or nano-architecture can be elucidated [18,19].

Medical diagnosis and therapeutical decisions very often depend on the evaluation of a large number of specimens like blood cells or fixated cells. In this case, the measured parameters for the appropriate ensembles are evaluated statistically [41]. The methods can range from simple calculations of mean and variance values up to elaborate statistical tests [42]. We include this evaluation into the third step within our business model, but one should keep in mind that this can also be defined as a separate fourth step in more complex applications.

Computational Services

Computer Programs and Platforms

Financial contributions to the work being reported should be clearly acknowledged, as should any potential conflict of interest. In our diagostic pipeline, the last step obviously depends heavily on the use of computers. In addition, the first step, when employing individual probe set design for COMBO-FISH, also makes use of a program running on data allowing to compose different oligomers in such a way that they form a single cluster at the desired genetic spot [10,11]. On the other hand, the second step, namely the acquisition of microscopic image stacks, does not involve any computer action except for the fact, that it produces files as output which have to be further processed. We therefore concentrate on two types of computer programs to be used in our business model:

a) A program for the design of COMBO-FISH probe sets for unique labeling of a specified genomic location, and

b) A program system for image analysis and statistical evaluation of the extracted parameters.

There are several major problems for the implementation of the two types of program systems. First of all, the image files appear in a lot of different formats depending on the specific microscope companies and the software available on the side of the users who have different requirements and prerequisites. During the development of our programs we dealt with different file formats and different programming languages. It turned out that one has to provide programs for every possible microscope image format to transform the files to a standard format that can be evaluated by the image analysis software available, because raw image data from the microscope cannot be used directly for analysis. In our applications, we decided to provide data transfer to kde, tiff, and png image files for those microscopes which were used for COMBO-FISH analysis. Also files from clinics which applied standard FISH were transformed to tiff format.

The second problem concerns the image analysis software itself. Though in our first developments we used Matlab programs [43], we encountered massive problems concerning licensing on different computer clusters, grids, or clouds. Therefore, the main algorithms have been redesigned and implemented in C and C++ language code and run on different platforms including an 8 processor cluster node within the German D-Grid [44]. The algorithms for the segmentation of the single images belonging to one stack were manually adapted to the special type of image specific for the respective application in order to run automatically on the rest of the data. From the segmented 2D images, the 3D structure is comput-

ed and numerical and statistical parameters are extracted. They are the basis of statistical evaluations which are returned to the user for his diagnostical decision. We also used open source software for image analysis [38] which seems to be another possibility to avoid licensing problems with commercial software.

As for the algorithms for COMBO-FISH probe set design in the first step of the diagnostic pipeline, the problems are comparatively small. The human genome data, namely the sequence and the annotation, are retrieved and downloaded from NIH, Bethesda, USA, from the NCBI data base [27]. The search and analysis programs are written in C and run on essentially any platform with an appropriate compiler. For double helical COMBO-FISH, the BLAST algorithm [45] which is open source software can be used. The probe sets can be generated automatically or with user interference to guide the selection of oligomers.

Applied Grid Technology Components and Data Management

We have implemented all computational algorithms for image evaluation, parameter extraction, statistical analysis, and also for genome analysis for COMBO-FISH probe set design on different kinds of computers. The programs were tested in several case studies [44] within the research project Services@MediGRID on German Grid devices. To provide the computation services, we have concentrated on grid and cloud computing. As for the MediGRID environment, we relied on Globus Toolkit 4.0.x [46] as the basic middleware to distribute jobs specified by the model user. The data comprising programs and image files were implemented and stored locally. Workflows submitted to MediGRID were scheduled and coordinated to resources by the Generic Workflow Execution Service (GWES) [47]. For uploading and maintaining service datasets, a user interface in XML language is provided. The language used for the description of resources and services is the D-Grid Resource Description Language (D-GRDL).

For data sensitive applications, authentication via PKI based personal certificates [48,49] is necessary. For less privacy-relevant services (e.g. computation of COMBO-FISH sets for a specific genomic location) a guest account with e-mail verification in the registration process is available in combination with a robot-certificate used by the service client portlet. All grid jobs are therefore executed using PKI and identifying a person legally responsible for the actions of the job.

The Business Model

Business Concept Ingredients

The diagnostic pipeline contains three ingredients: A biochemical process of hybridization, a microscopic process of image acquisition, and an informatical process of parameter extraction via computationally intensive image analysis. In addition, for the method of double or triple helical COMBO-FISH, a preparation step of configuration of a probe set for the hybridization by combinatorial analysis of genome data may be necessary. These processes are essentially independent and can by performed by different service providers. In fact, present diagnostic systems, usually based on standard FISH, use different business models depending on their users' capabilities. If biochemical equipment and laboratory technicians are available, hybridization procedures can be performed within the respective institution, whereas in other cases the specimens are sent to a laboratory for further processing. If clinics are equipped with appropriate microscopic systems and personnel, they can care for image acquisition themselves, otherwise they have to search for assistance by a specialized organization. Image analysis is usually performed with standardized software, be it by the clinic itself or by a commercial provider. These existing pathways realizing the diagnostic pipeline already show the complexity of the network of the business partners involved. In fact, to keep the procedures for diagnostics as simple as possible, most diagnostic kits provided by the biochemical industry are designed to be easy to handle in a standardized protocol. Still, the client, usually a clinic, has to decide whether to do the biochemical procedure and image acquisition and analysis on their own or to leave it to one or several specialized laboratories.

We want to include these existing pathways into our business model, but in view of the development of personalized medicine and individualized diagnostic methods, for which COMBO-FISH can be considered as one example, we will formulate it as a rather flexible concept. The bedside practitioner may be interested in the numerical, geometrical and statistical results of the hybridization experiment, which is specific for his diagnostic decision. This may imply a spectrum of methods, ranging from fully automated standard applications to individual special microlocal investigations, e.g. for gene compaction. The latter is one example for which COMBO-FISH probe sets for marker genes have to be predesigned.

The subsequent image analysis may require different approaches including manual processing on an interactive basis as well as completely automated runs on the computer. In clinical applications, parameters, like the amplification of gene numbers in cell nuclei, are extracted from the hybridization image of the standard probe manually or automatically. In more complex applications, a whole process chain from a manually optimized image acquisition to the statistical evaluation of the data may be necessary. In this regard, the computational domain of the business model is by far more complex than the biochemical and microscopic parts which are performed in rather standardized ways. This will also influence the possible choice of computational platforms like local computers, local clusters, clouds, or even computing grids, and define the relations between client and provider.

Realizing the Business Concept

In our business model, the client is a clinic that will provide a hybridized specimen which has to be analyzed by microscopic imaging and image analysis and statistical evaluation. The service provider is a company which performs these tasks, be it on their own or by subcontracting. For the design of a probe set for triple helical COMBO-FISH, it may be necessary for the client before performing the hybridization procedure to contact the provider who in this case may be the same company or another. In addition, the client may be subcontracting the hybridization of the specimen to the same or another provider.

On a general level, we can visualize [Fig-1] the work flow in our business model in four steps: The client takes a specimen from the patient and performs the hybridization procedure (step 1) which is then sent to the provider for microscopic imaging (step 2), image analysis and statistical evaluation of parameters (step 3). The results are returned to the client (step 4) together with an accounting

record, and the client pays the bill. As already mentioned, there may be an additional step for the task of triple helical COMBO-FISH probe design which would have to be integrated in [Fig-1] as an optional step on the customer level.

In our schematic description showing the different processes, we have introduced three layers to display the different types of interactions of customers and service providers. The first layer shows the customer, which may be scientific institutions or industrial laboratories, but most prominently clinics in our application. The second layer, which we call genome analysis provider here and which includes imaging laboratories, performs a labor transparent to and on the order of the customer by taking microscopic images and analyzing them. In fact, the computer power intensive computational steps are transferred to the third layer, performed by the grid service provider.

This type of business model contains all steps for an instantaneous image analysis and subsequent possible diagnostics, which will in general be performed by the customer himself, but may of course also be deligated to a service provider. In so far, this is a general description of the business model. But as already indicated, the roles of customer and provider may be distributed differently. This will be shown in two variations of the business model, which on the other hand still stick to the same general scheme.

In fact, as depicted in [Fig-2], the services of microscopy and / or imaging may be performed by the customer himself, which depends on his capabilities. In this case, the possible workflow scenarios represent more complex interactions.

In practice, clinicians are interested in a subsequent interpretation of their image data and in storing the images and results for further use in research projects or diagnostic improvements. Therefore, an additional step during image processing can be necessary, namely an online interpretation of the resulting genomic architecture and an iteration of the image analysis procedure. Usually, this will be done by the customer himself after he has received and inspected the image analysis results. This step again depends on the hardware and software he has at his disposition, which may be limited, especially in situations in which peaks of computational power are required. Therefore, in [Fig-3], we have introduced the possibility to visualize the imaging results and to retransfer the image analysis procedure to the grid by processing the results via a grid app client.

It should be noted that these three business model variants cover the most relevant use cases which appear in modern cancer diagnosis based on fluorescence microscopy and computer image analysis of (COMBO-)FISH labeled cell specimens. If performed in a standardized way, they can cope for a fast and stable pipelining in diagnostic processes. In the next subsection, we will investigate the role of the provider in relation to the customer and his responsibilities.

Role of Customer and Provider

The main A major difference between the three variants of the business model concerns the role of the respective customers and service providers as well as their tasks. Though, at first sight, this might only be an organizational difference, it can constitute a severe problem for the provider, especially in the second and third example. In the first model [Fig-1], one can think of a single provider, e.g. the laboratory selling the hybridization probes or an optical laboratory, which would then care for microscopic imaging and image analysis on own devices or grid or cloud computing facilities, possibly by subcontracts. As this provider would also transmit the results, the whole process is in his hands. He can plan the schedule of his actions and has complete control of accounting and billing. This is especially the case if he is already a well established enterprise on the market, e.g. in the biomedical business. It is therefore well conceivable that this business model may be attractive for such companies which would render them the opportunity to extend their business range and to attract new customers.

On the other hand, in model variants two [Fig-2] and three [Fig-3], planning may be much less reliable for a company, especially if they only concentrate on grid computing services. The same, of course, applies to a company in business model one [Fig-1] which does not have an alternative biochemical or optical branch. In fact, the whole life cycle management of the grid application services or any other means of high throughput computing service necessary for efficient image analysis includes deployment, upgrading, development, and further services which are rather cost intensive. In addition, integration of software and implementation of algorithms may be complicated within different grid resources or infra structures connecting grids, clouds, and clusters. For the same reason, platform services like accounting, billing, monitoring, brokering, and workflow management may be complicated. All this can only be successfully performed, if the economic risk is minimized, e.g. by a reliable business foundation on standard economic procedures in this branch like production of hybridization probes or microscopic imaging. In this case, the company already has a fixed clientel of customers and they can offer them a new integrated solution for the whole diagnostic pipeline.

The complete process pipeline from sampling of the specimen to statistical evaluation of the microscopic images is appealing in two ways. On the one hand side, it offers the possibility for the respective company to sell the whole diagnostic kit in one package inclusive of all handling procedures, and on the other hand side, for the clinic a standardized diagnostic pipeline offers more security for treatment. On the contrary, for non standard or rare diagnostic procedures, the business models two or three might be more adequate.

Discussion and Outlook

We have designed a business model for the diagnostic pipeline from sampling of a tissue specimen to the parameter evaluation of the microscopic hybridization images, which may be a basis of the diagnostic decisions. The business model includes all steps from biochemical laboratory work via microscopic imaging to image analysis and parameter detection as well as statistical evaluation. In addition, for new COMBO-FISH applications, a combinatorial search for a special hybridization set can be performed in advance. The respective algorithms have been programmed and implemented on a cluster within the German D-Grid infrastructure [44]. Within this system, the performance of the diagnostic pipeline has been simulated for several genetic elements, especially the Her2neu gene (ERBB2) which plays a prominent role in breast cancer diagnostics. Tissue and blood samples were provided by clinics in Freiburg im Breisgau and Jena. The images were taken by microscopic

equipment for instance in Freiburg, Jena, Tel Hashomere (Israel), and Heidelberg. Image analysis was performed by specially developed software on stand alone computers and on the beforehand mentioned cluster, but also commercial and open source software was used. For statistical parameter evaluation and combinatorial search for COMBO-FISH hybridization sets, special programs were developed in C programming language. In so far, we have covered a whole lot of different implementation possibilities for the various steps in our three business model variants. We can conclude that in principle the business model is feasible and can be realized.

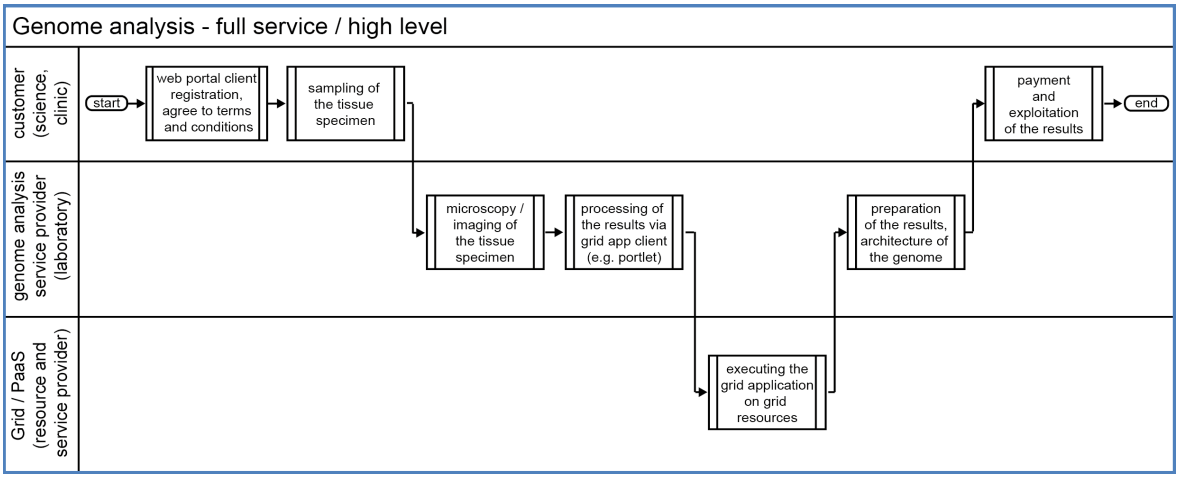

Fig. 1- Blueprint of full service genomic image analysis

PaaS = Platform as a Service

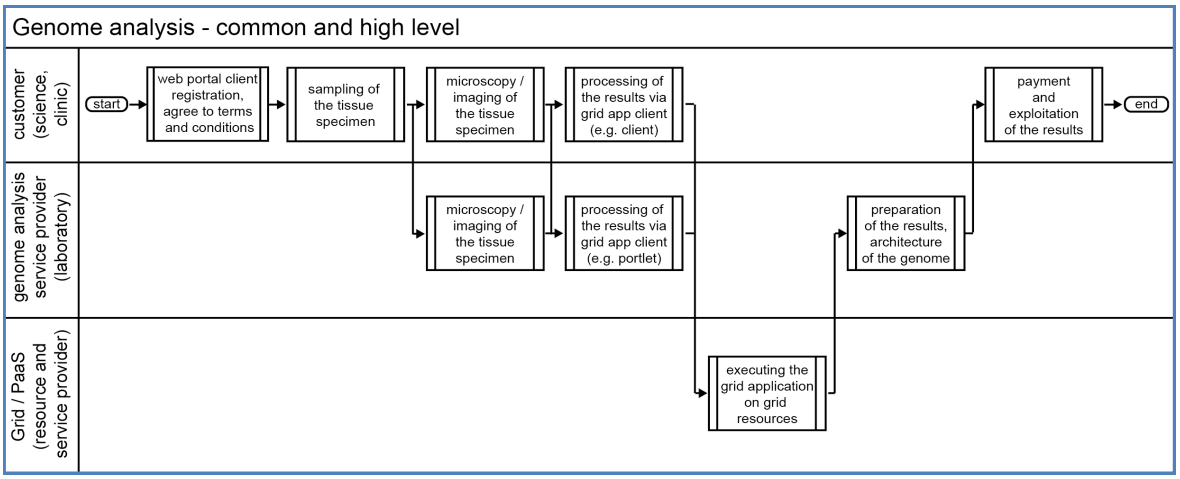

Fig. 2- Blueprint of full service genomic image analysis with possible customer interference

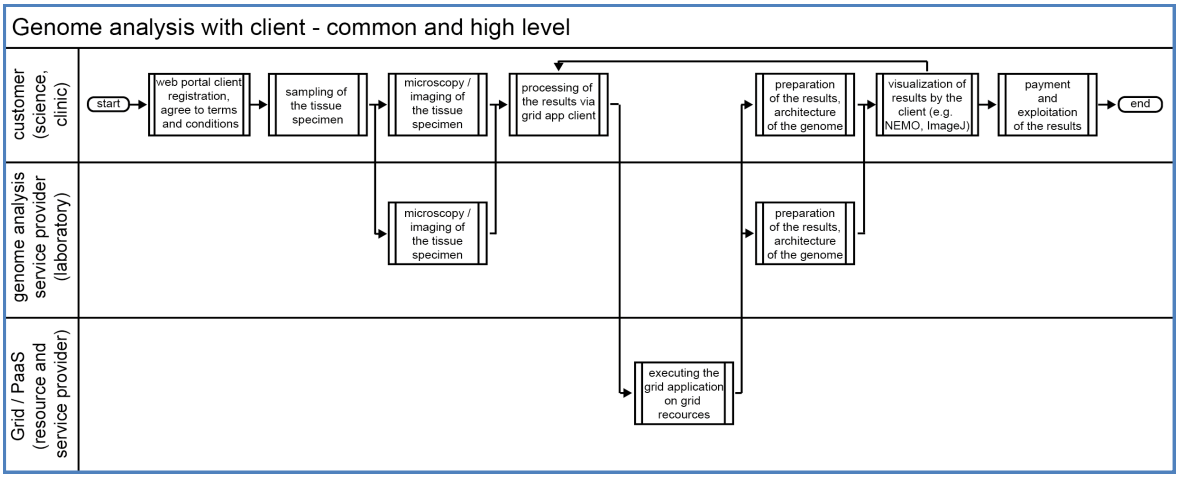

Fig. 3- Service Blueprint for genomic image analysis including an iterative client application by the customer for controlled visualization

On the other hand, we encountered several problems which show, however, that such a realization will carry some risk and specific problems still have to be solved. A very special problem concerning transmission, temporary storage, and processing of clinical data is data security. On local computers and clusters, this may be a minor problem, but on grids it may become a legal matter. In our case, data can be anonymized and transmitted as pure image files without any relation to personal data. This may constitute a solution to most data security problems. Things become different, if such data have to be stored permanently, e.g. for scientific use in the future. In any case, the provider of services concerned with such data has to care for secure procedures. In addition, all transactions within the grid have to obey the usual security standards. In the case of the German D-Grid, we have indicated possible ways of handling in section 3.2.

Other problems concern the software itself. If commercial software is used, which is a very comfortable way in many image processing applications, licensing may become complicated. In fact, also in connection and cooperation with other German grid initiatives, we were up to now not able to reach acceptable standards of licensing for even very common commercial software for the grid community. It would therefore be the responsibility of the provider to settle suitable agreements with software companies.

The last complex of issues that needs additional research and development of methods is monitoring, accounting, and billing. It depends very much on the precise implementation of the respective operational systems and their specificities, how monitoring can be organized. Even the term monitoring is controversely discussed in the community. Though it originally only referred to the supervision of the process workflow itself, it has become clear during the development of the business model here and others as well, that additional detail information has to be gathered to enhance the accounting process. In so far, the two terms of monitoring and accounting are closely related and it will be the providers duty to set up a transparent system of information retrieval by monitoring the workflow and computing performance to gain the necessary data for accounting. Ultimately, it is billing which makes his business run. Here, the providers skill is challenged to make accounting and billing profitable for himself and attractive for the customer. It will therefore be important to evaluate the model's behavior [50-52] within the healthcare market.

In our business model, we have simulated a whole diagnostic pipeline, especially for cancer diagnosis and therapy control. Though the issue of business model itself is controversely discussed in literature [53-55], we are sure that this is a suitable concept for further developments of the kind in healthcare and medical research. Further aspects are the integration of research projects into the business model on the basis of long term image data storage in large distributed data bases, development of adapted image analysis software and statistical methods, and the extension to further diagnostic applications. Grid, cloud, and cluster computing are well suited resources to build on.

Acknowledgement

The authors thank the German Federal Ministry of Education and Research (BMBF) for financial support within the joint projects Services@MediGRID (FKZ 01IG07015A-G) and COMBO-FISH (FKZ 13N11163). Furthermore, the authors are indepted to the European Center of Feasibility Studies, Strasburg, especially to Dr. Boris M.B. Fertig for continuous support in any bottomless pit case and to Dr. Dora F. German for a stimulating breakup to preposterous environments.

References

[1] Speicher M.R., Ballard S.G., and Ward D.C. (1996) *Nature Genetics*, 12, 368-375.

[2] Janz S., Potter M. and Rabkin C.S. (2003) *Genes, Chromosomes & Cancer*, 36, 211-223.

[3] Bridger J.M. and Volpi E.V. (2010) *Fluorescence in Situ Hybridization (FISH): Protocols and Applications*.

[4] Frigault M.M., Lacoste J., Swift J.L. and Brown C.M. (2009) *Journal of Cell Science*, 122, 753-767.

[5] Pitas I. and Lyroudia K. (2007) *Cytometry Part A*, 71A, 439-450.

[6] Ventura R.A., Martin-Subero J.I., Jones M., McParland J., Gesk S., Mason D.Y. and Siebert R. (2006) *Journal of Molecular Diagnostics,* 8(2), 141-151.

[7] Kurtzman G. (2005) *Biotechnology Healthcare*, 50-55.

[8] Smerage J.B. and Hayes D.F. (2006) *British Journal of Cancer*, 94, 8-12.

[9] Camara O., Rengsberger M., Egbe A., Koch A., Gajda M., Hammer U., Jörke C., Rabenstein C., Untch M. and Pachmann K. (2007) *Annals of Oncology*, 18, 1484-1492.

[10] Hausmann M., Winkler R., Hildenbrand G., Finsterle J., Weisel A., Rapp A., Schmitt E., Janz S. and Cremer C. (2003) *BioTechniques*, 35, 564-577.

[11] Schmitt E., Wagner J. and Hausmann M. (2012) *Journal of Computational Science*, 3(5), 328-334.

[12] Ferrara J. (2007) *McGill Journal of Medicine*, 10(1), 59-61.

[13] Eppinger E. and Kamprath M. (2011) *R&D Management Conference*, Linköping, Sweden.

[14] Hoffa C., Mehta G., Freeman T., Deelman E., Keahey K., Berriman B. and Good J. (2008) *IEEE Fourth International Conference on eScience*, 640-645.

[15] Rings T., Caryer G., Gallop J., Grabowski J., Kovacikova T., Schulz S. and Stokes-Rees I. (2009) *J. Grid Computing*, 7, 375 -393.

[16] Pachmann K., Zhao S., Schenk T., Kantarjian H., El-Naggar A.K., Siciliano M.J., Guo J.Q., Arlinghaus R.B. and Andre M. (2001) *British Journal of Haematology*, 112, 749-759.

[17] Geyer C.E., Forster J., Lindquist D., Chan S., Romieu C.G., Pienkowski T., Jagiello-Gruszfeld A., Crown J., Chan A., Kaufman B., Skarlos D., Campone M., Davidson N., Berger M., Oliva C., Rubin S.D., Stein S and Cameron D. (2006) *N. Engl. J. Med.*, 355, 2733-43.

[18] Grossmann C., Schwarz-Finsterle J., Schmitt E., Birk U., Hildenbrand G., Cremer C., Trakhtenbrot L. and Hausmann M. (2010) *Microscopy: Science, Technology, Applications and Education, Formatex Research Center,* Badajoz, 688-695.

[19]Hildenbrand G., Rapp A., Spöri U., Wagner C., Cremer C. and Hausmann M. (2005) *Biophysical Journal*, 88(6), 4312-4318.

[20]Wiech T., Stein S., Lachenmaier V., Schmitt E., Schwarz-Finsterle J., Wiech E., Hildenbrand G., Werner M. and Hausmann M. (2009) *European Biophysics Journal*, 38(6), 793-806.

[21]Timme S., Schmitt E., Stein S., Schwarz-Finsterle, Wagner J., Walch A., Werner M., Hausmann M. and Wiech T. (2011) *Anal Cell Pathol (Amst).*, 34(1-2), 21-33.

[22]Cremer T. and Cremer C. (2001) *Nature Reviews Genetics*, 2 (4), 292-301.

[23]Liehr, Thomas (2009) *Fluorescence In Situ Hybridization (FISH) - Application Guide*, Springer, Berlin, 452.

[24]Schwarz-Finsterle J., Stein S., Großmann C., Schmitt E., Trakhtenbrot L., Rechavi G., Amariglio N., Cremer C. and Hausmann M. (2007) *Journal of Biochemical and Biophysical Methods*, 70(3), 397-406.

[25]Schwarz-Finsterle J., Stein S., Grossmann C., Schmitt E., Trakhtenbrot L., Rechavi G., Amariglio N., Cremer C. and Hausmann M. (2005) *Cell Biology International*, 29(12), 1038-1046.

[26]Schmitt E., Schwarz-Finsterle J., Stein S., Boxler C., Müller P., Mokhir A., Krämer R., Cremer C. and Hausmann M. (2010) *Fluorescence in Situ Hybridization (FISH): Protocols and Applications*, Humana Press, New York, 2010, 185-202.

[27]IHGSC (2004) *Nature*, 431(7011), 931-945.

[28]Hausmann M., Hildenbrand G., Schwarz-Finsterle J., Birk U., Schneider H., Cremer C. and Schmitt E. (2005) *Biophotonics International*, 12(10), 34-37.

[29]Schmitt E., Stein S. and Hausmann M. (2009) *Studies in Health Technology and Informatics*, 147, 251-256.

[30]Zeller D., Kepper N., Hausmann M. and Schmitt E. (2013) *IFMBE Proceedings*, 38, 82-85 *[Article in Press]*.

[31]Schmitt E., Müller P., Stein S., Schwarz-Finsterle J. and Hausmann M. (2010) *Studies in Health Technology and Informatics*, 159, 272-276.

[32]Lemmer P., Gunke M., Baddeley D., Kaufmann R., Urich A., Weiland Y., Reymann J., Müller P., Hausmann M. and Cremer C. (2008) *Applied Physics B: Lasers and Optics*, 93(1), 1-12.

[33]Lemmer P., Gunkel M., Weiland Y., Müller P., Baddeley D., Kaufmann R., Urich A., Eipel H., Amberger R. and Hausmann M. (2009) *Journal of Microscopy*, 235(2), 163-171.

[34]Bohn M., Diesinger P., Kaufmann R., Weiland Y., Müller P., Gunkel M., von Ketteler A., Lemmer P., Hausmann M., Heermann D.W. and Cremer C. (2010*) Biophysical Journal*, 99 (5), 1358-1367.

[35]Müller P., Schmitt E., Jacob A., Hoheisel J., Kaufmann R., Cremer C. and Hausmann M. (2010) *International Journal of Molecular Sciences*, 10(11), 4094-4105.

[36]Cremer C., Kaufmann R., Gunkel M., Pres S., Weiland Y., Müller P., Ruckelshausen T., Lemmer P., Geiger F., Degenhard S., Wege C., Lemmermann N.A.W., Holtappels R., Strickfaden H. and Hausmann M. (2011) *Biotechnology*, 6, 1037-1051.

[37]Ferreira T. and Rasband W. (2011) *Image J. User Guide*.

[38]Iannuccelli E., Mompart F., Gellin J., Lahbib-Mansais Y., Yerle M. and Boudier T. (2010) *Bioinformatics*, 26(5), 696-697.

[39]Monajembashi S., Rapp A., Schmitt E., Dittmar H., Greulich K.O. and Hausmann M. (2005) *Biophysical Journal*, 88(3), 2309-2322.

[40]Jhunjhunwala S., van Zelm M.C., Peak M.M., Cutchin S., Riblet R., van Dongen J.J.M., Grosveld F.G., Knoch T.A. and Murre C. (2008) *Cell*, 133(2), 265-279.

[41]O'Rourke N., Hatcher L. and Stepanski E. (2005) *A Step-by-Step Approach to Using SAS for Univariate and Multivariate Statistics*, 2nd ed., SAS Press and John Wiley Sons Inc., Cary, NC, USA.

[42]Rice J.A. (1988) *Mathematical Statistics and Data Analysis*, Pacific Grove, CA, USA, 595.

[43]Attaway S. (2012) *MATLAB: A Practical Introduction to Programming and Problem Solving*, 2nd ed., Elsevier Science.

[44]Dickmann F., Falkner J., Gunia W., Hampe J., Hausmann M., Herrmann A., Kepper N., Knoch T.A., Lauterbach S., Lippert J., Peter K., Schmitt E., Schwardmann U., Solodenko J., Sommerfeld D., Steinke T., Weisbecker A. and Sax U. (2012) *Journal of Computational Science*, 3(5), 280-297.

[45]Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. (1990) *J. Mol. Biol.*, 215(3), 403-410.

[46]Foster I. (2006) *Journal of Computer Science and Technology*, 21(4), 513-520.

[47]Hoheisel A. (2006) *The Cracow Grid Workshop*, 18(10), 1101-1113.

[48]Krefting D., Bart J., Beronov K., Dzhimova O., Falkner J., Hartung M., Hoheisel A., Knoch T.A., Lingner T., Mohammed Y., Peter K., Rahm E., Sax U., Sommerfeld D., Steinke T., Tolxdorff T., Vossberg M., Viezens F. and Weisbecker A. (2009) *Future Generation Computer Systems*, 25(3), 326-336.

[49]Chakrabarti A. (2007) *Grid Computing Security*, Springer, Berlin.

[50]Huelsbeck D.P., Merchant K.A. and Sandino T. (2011) *The Accounting Review*, 86(5), 1631-1654.

[51]de Cesare S., Lycett M. and Paul R.J. (2003) *Communications of the Association for Information Systems*, 12, 223-241.

[52]George G., Bock A.J. (2011) *Entrepreneurship Theory and Practice*, 35(1), 83-111.

[53]Normann R. (1977) *Management for Growth*, John Wiley and Sons, Chichester.

[54]Osterwalder A., Pigneur Y. and Tucci C.L. (2005) *Communications of Association for Information Systems*, 15, 1-25.

[55]Seppänen M. and Mäkinen S. (2007) *International Journal of Management Concepts and Philosophy*, 2(4), 389-404.