

# eInfrastructure

## Bioimaging informatics software and support

(This project description corresponds to 2-pager sketch nr 253)

Host institution:

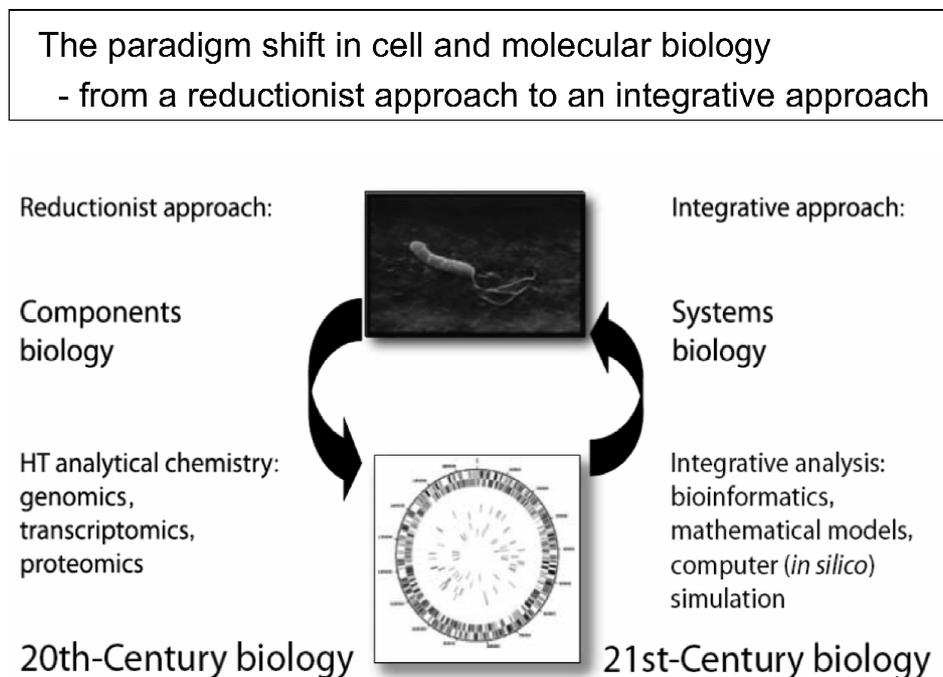
University of Bergen, Faculty of Medicine and Odontology, Department of Biomedicine

Scientific contact person:

Professor Arvid Lundervold (Arvid.Lundervold@biomed.uib.no/55586353)

### 1. Vision and scientific goals

During the past decades, progress in imaging technology and molecular biology has started a revolution in our understanding of molecular and cellular processes in biological systems. In combination, a new interdisciplinary field of research is emerging: image-based systems biology. *Systems biology* focuses on analysis and modelling of complex biological and chemical interactions in cells and tissue, rather than breaking down the systems into isolated pieces (cf. Fig.1). This integrative approach can thus give new biological understanding - insight that traditional “reductionism” can not provide.



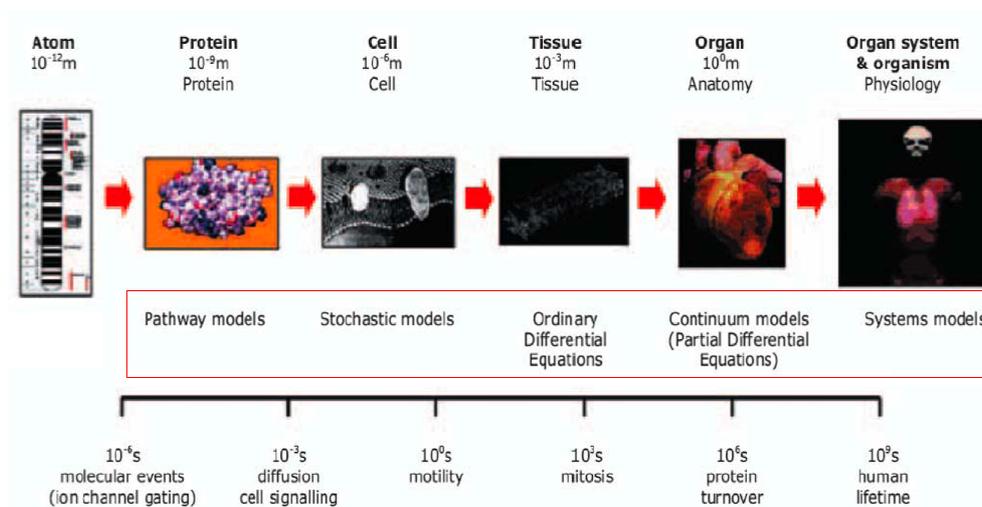
From: B. O. Palsson, *Systems biology: Properties of reconstructed networks*, Cambridge University Press, 2006.

**Figure 1. Components biology versus systems biology.**

In this context *bioimaging informatics*, enabling structure and function of molecules, cells and tissue to be studied in time and space, constitutes a distinct scientific and technological discipline. It is a rapidly evolving branch of information technology acting in the field of molecular cell biology and physiology with strong links to applied mathematics, statistics and computer science. Bioimaging informatics, including high-throughput image processing, represents a valuable tool in systems biology by providing in situ information from complex systems at a wide range of spatial and temporal scales (cf. Fig.2) and enables the quantification, modelling and visualization of underlying biological processes from huge collections of recorded data.

## Tools towards (image-based) Systems Biology

The large span of spatial and temporal scales in biological systems



From: C. Dollery and R. Kitney, Systems biology: A vision for engineering and medicine. Tech. report, The Academy of Medical Sciences and The Royal Academy of Engineering, London, UK, Feb. 2007.

**Figure 2. Recordings, analysis and modeling of biological structures and processes at different scales.**

New and powerful bioimaging devices produce an increasing amount of image data both in 2-D, 3-D, 3-D+time (time-lapse imaging) and even multispectral 3-D+time (i.e. 5-D) data, and high-throughput imaging systems can produce several thousand images and gigabytes of data in less than one hour. This development of highly efficient image acquisition systems such as confocal spinning-disk microscopy and small-animal high-field MRI is in great contrast to the typical situation for storage/retrieval and analysis of such data. In connection with a biological experiment involving image acquisitions the researcher typically has to provide him or herself a routine for (long-term) *storage* of the image files and possibly also for the *metadata* that accompanies these files in order to track image acquisition parameters and experimental conditions. Moreover, the important part of *analysing* the image data typically involves time-consuming and operator-dependent spatial measurements, delineations of regions of interest (ROIs), manual counting of objects (e.g. cells), and visual integration of information from the same biological sample examined with different imaging devices such as confocal microscopy and MRI (i.e. multimodal imaging). In repeated measurements or in time-lapse imaging, important biological questions are typically related to the concept of *change* - in size, in

shape, in number, or in position – and these questions can not be answered without proper image analysis tools for e.g. segmentation and object tracking.

The above description should illustrate the difference between ‘bioimaging’ (in the meaning of image producing devices for biomedical applications) and ‘bioimaging informatics’, where the last term addresses the important issue of storage and retrieval of image files as well as image processing, image analysis and visualization.

**In biomedical imaging of today it seems that the ‘informatics’ part (storage/retrieval, analysis and visualization) is a bottleneck that partly limits the impact of digital imaging devices in molecular and cell biology research. The goal of present infrastructure application is to remedy this situation both regionally i.e. at the Molecular Imaging Center (MIC-Bergen), and also nationally e.g. through the FUGE-funded Norwegian Molecular Imaging Consortium (NorMIC) and thus make better use of the advanced image producing equipment that has been invested, partly by FUGE, over the years. To achieve this goal, a cross-disciplinary force must come into action. Therefore, both biologists/medical doctors, computer scientists and mathematicians at UoB will take part in this effort, putting their expertise into the infrastructure build-up according to the project plan and the requested resources that are described below.**

In addition, the NRC/FUGE has clearly pointed out the importance of addressing this in connection with the FUGE board decision to fund the establishment of the **NormIC consortium**. The grant letter to NorMIC (with regard to imaging data) stated that *“There must be developed a strategy for how to...increase the focus on bioinformatics and data handling..”*. This statement reflects the situation where development and availability of general bioinformatics tools, services and expertise at our universities and research institutions are more oriented towards e.g. DNA and protein sequencing, gene finding, genome assembly, and protein-protein interactions and are not so much of help in image analysis, advanced visualization or image-based modeling. **A long-term goal will therefore be to bridge cultural and methodological gaps between researchers in Bergen and Norway working within the seemingly separate fields of bioinformatics, computational biology, systems biology, and biomedical image processing and advanced visualization. This could be achieved by gradually integrating infrastructure and facilitate exchange of tools, data, skills, and ideas.**

As a first step, this **e-infrastructure application** aims to address the imbalance between image production on one side and image storage /retrieval, analysis and visualization on the other. It represents an answer the guidelines and recommendations given by the NRC of establishing **a facility that provides easy access and user support to state-of-the-art software tools for (i) storage, (ii) retrieval, (iii) automated analysis and (iv) visualization of the rapidly increasing collection of biomedical images, of various modalities, recorded in the NorMIC laboratories**. On the storage and retrieval side, images and accompanying meta-data will be stored centrally so they can easily be retrieved from any location at any time. Furthermore, the facility will turn image analysis and visualization open source software into available services for in-house users.

The proposed infrastructure will also **facilitate basic research in applied image analysis and visualization**, and by including mathematicians and computer scientists in the project the **likelihood of originality and scientific impact** will increase regarding methodological solutions and software tools for the molecular imaging user community.

We have local empirical evidence for this statement through the PhD work of Erlend Hodneland (MSc in applied mathematics). In the PhD evaluation committee report from late 2008 the following remark to his **Paper I**: Erlend Hodneland, Arvid Lundervold, Steffen Gurke, Xue-Cheng Tai, Amin Rustom and Hans-Hermann Gerdes. Automated detection of Tunnelling Nanotubes in 3D images. *Cytometry Part A* 69A:961-972 (2006) is made

*“ ...This paper describes an automated image analysis approach to detect tunneling nanotubes, TNT, a structure in cell images quite recently described in the literature ([Science] 2004, only 2 years before the publication of this paper, impressively rapid response!).*

*and .... In conclusion the PhD candidate Erlend Hodneland has presented a thesis which demonstrates the value of close collaboration between a computer science/applied mathematics fields such as image analysis and biomedical research. Image analysis tools from literature has been implemented and tested for use for solution of specific biomedical problems. In several cases the tools have also been significantly improved. In particular the tools have been adapted to handle surface stained cells and detection of tunneling nanotubes. The candidate and his coauthors have also made some general fundamental contributions to image analysis by exposing the close connections between two popular but distinct image segmentation tool boxes: level sets and active contours, suggesting and evaluating a set of approaches to how the methods can be joined within the same mathematical framework.”*

**A next step for our vision** in e-infrastructure for bioimaging informatics in Norway, not covered by this application, is to scale up and use the existing national **GRID** infrastructures for storing and processing the image data. There are currently ongoing efforts, financed through the **NOTUR** Advanced User Support programme, aiming for interfacing a set of biomedical applications with **NorGrid** and **NorStore** in collaboration with scientists from Parallab (e.g. Lundervold & Johannsen “Grid Gateway for Advanced Scientific Applications” [www.notur.no/support/advanced/2008/projects.html](http://www.notur.no/support/advanced/2008/projects.html) 2008.3; “Quantitative brain imaging in health and disease” [www.notur.no/support/advanced/2006/summary/2006lundervold.html](http://www.notur.no/support/advanced/2006/summary/2006lundervold.html) and Lundervold, Ystad, Anderlik<sup>1</sup>: [www.notur.no/publications/magazine/pdf/meta\\_2007\\_1.pdf](http://www.notur.no/publications/magazine/pdf/meta_2007_1.pdf) ) As a continuation of the present proposal, or as a parallel activity, we would like to build on the knowledge accumulated over the last ten years at Parallab in the field of Grid computing, and create the necessary software layer to enable the bio-infrastructure to use this technology to process, store and share the data produced by the imaging devices. In particular, we would like to provide access to the infrastructure through a few “gateway” nodes. These are computers that can have special hardware features (for example very capable GPUs to enable advanced visualization) as well as the software layer to create the communication between the users (and applications) and the Grid infrastructure. Such ideas are not farfetched to realize, since it is similar to the Application Server setup that has been developed at Parallab within the Advanced User Support project.

<sup>1</sup>Brain imaging and computational neuroanatomy using high performance and Grid computing. *Meta* 2007-1:6-9.

## 2. Scientific and technological environment

The **NorMIC-nodes** (consisting of imaging facilities at the Universities in Bergen, Stavanger, Oslo, Trondheim and Tromsø) and others are heavily equipped with high-end image acquisition systems like confocal, multi photon, high-throughput and spinning disc microscopes, and have highly qualified personnel running them. **MIC-Bergen**, the first

national Molecular Imaging Center, funded by FUGE 1 and now the head node in the NorMIC consortium, is localized in the Department of Biomedicine, University of Bergen ([www.uib.no/med/mic](http://www.uib.no/med/mic)). It is one of the most well-equipped nodes (cf. Fig. 3) and the mission of the Center is to maintain and develop modern technologies of biological imaging, and make them available to scientists to be applied in specific projects carried out in different research fields, such as Cell and Developmental Biology, Neurobiology, and Cancer Biology. MIC Bergen also provides a web-based booking system for end users (<http://services.cbu.uib.no/service/mic/instruments>) that keeps it simpler to operate the center, including budgeting, accounting, and optimal use of the instruments.

Instrument	Supplier/Model	Application(s)
Confocal microscopes	Leica TCS SP2 AOB5	Multi-channel 4-D acquisition, spectral detection, FRET, FRAP, <a href="#">more</a>
	Leica-TCS-SP5	Multi-channel 4-D acquisition, FRET, FRAP, <a href="#">more</a>
	Zeiss LSM 510 Meta	Multi-channel 4-D acquisition, FRET, FRAP, <a href="#">more</a>
Multiphoton	Leica SP5 MP	Deep tissue imaging and live animal imaging
Spinning disk	PerkinElmer UltraView RS	Fast confocal time-lapse imaging
Fluorescence microscope	Nikon TE 2000	Epi-fluorescence and DIC imaging through glass and plastic
Electron microscopes	Jeol JEM-1230	Transmission
	Jeol JSM-7400F	Scanning
Magnetic resonance*	7.0 Tesla MR	Small animal MR imaging
Small animal molecular imager	ART eXplore Optix	In vivo detection of near IR and GFP/luminescent probes
Flow Cytometer	Flow Cytometer	Cell sorting and in vitro detection of cellular properties and fluorescence
High throughput	BD Pathway	Fully automated image acquisition and data analysis on multi-well plates
Microdissection	P.A.L.M.	Microdissects and captures with high precision

\* Users outside Bergen, please contact partners in the [Nor-MIC](#) consortium

**Figure 3. Equipment at MIC Bergen.**

Complementary to the imaging equipment and image production in MIC Bergen, there has through the last 10 years gradually evolved a strong cross-disciplinary research network in (biomedical) image processing and analysis in Bergen. Presently, this network has two closely interrelated legs: (i) the **BBG group** (Bergen bildebehandlingsgruppe - [www.math.uib.no/BBG/BBG/BBG.html](http://www.math.uib.no/BBG/BBG/BBG.html)) consisting of professors and students in Department of Mathematics, Department of Biomedicine and Department of Informatics; and more recently, (ii) the **MedViz consortium** ([www.medviz.uib.no](http://www.medviz.uib.no)), a research cluster of groups performing interdisciplinary research in advanced image analysis and visualization consisting of nodes at the University of Bergen, Haukeland University Hospital, and Christian Michelsen Research AS. The MedViz consortium runs a monthly seminar series and organizes a yearly conference with invited national and international speakers and coordinates several imaging-related projects in Bergen involving PhDs and postdocs from many disciplines. Members of BBG have published much of their work in highly ranked international journals (e.g. IEEE

Transactions on Medical Imaging, IEEE Transaction of Image Processing, Neuroimage, Human Brain Mapping, Cytometry A), and are also editorial board members of journals in the field. From BBG there has been collaboration, student exchange, and joint publications with several international centres for image processing, e.g. at UCLA, Harvard, Saarbrücken, Clermont-Ferrand, Ljubljana, Liverpool, Lodz, Jena, Mannheim, and Singapore. Core members of BBG have also organized international imaging symposia and workshops (e.g. International Conference on PDE-Based Image Processing and Related Inverse Problems, Oslo Aug 8-12, 2005; Geometrical Partial Differential equations: Numerics and Applications, Bergen Dec 6-7, 2006; CIPR-CMA workshop on Fast level set methods with applications to segmentation and shape identification, Bergen, Jan 17-18, 2008; ClimaVis08, Bergen Aug 21, 2008, 2nd International Conference on Scale Space and Variational Methods in Computer Vision, Voss, Norway, Jun 1-5, 2009). Moreover, MedViz will organize a Special Session on Analysis and Visualization of Dynamic Images in Medical Applications at the IEEE 6<sup>th</sup> International Symposium on Image and Signal Processing and Analysis, Salzburg, Austria Sep 16-18 2009 ([www.isispa.org](http://www.isispa.org)). MedViz and BBG are also partners in the recently NRC-funded “Norwegian Interdisciplinary Research Training Programme in Medical Imaging” (Nasjonalt tverrfaglig forskerskole i medisinsk avbildning) which is coordinated from Trondheim. We have this spring also launched a new international MSc program, “Biomedical Image Sciences” at Faculty of Medicine and Dentistry, primarily intended for students with a BSc degree in the exact sciences (physics, informatics, mathematics, statistics) who want to acquire competence in biomedicine and work on biomedical imaging applications, e.g. MIC-related projects.

Based on (i) the broad range of advanced image producing equipment and several years of experience with services at MIC Bergen, both to local users and users with affiliation elsewhere, (ii) extensive previous and ongoing research in biomedical image analysis and visualization, as well as our educational activity, and (iii) ongoing and planned collaboration between BBG/MedViz and the Bergen Center for Computational Science / Parallab with their knowledge of Grid technology, their HPC resources, and NOTUR connections, we will argue that **MIC Bergen is very well suited for building up an e-infrastructure according to the present proposal – an infrastructure that will support and promote bioimaging informatics at a regional, a national, and at an international level.**

#### ***Relevance with regard to national research agendas /strategies***

In the grant letter to the leading node in the NorMIC consortium, the NRC stated that “There must be developed a strategy for how to...increase the focus on bioinformatics and data handling..” **This application is thus a direct answer to the strategic guidance from the NRC.**

**In relation to the NRCs national strategy for eInfrastructure<sup>1</sup>** this application constitutes an **advanced storage facility** with infrastructure for depositories and software for access and processing of the data. Both storage/retrieval and processing, especially the latter, requires extensive support and tailoring of software to address the specific problems related to individual data sets and the individual researchers use of these data **A central grid infrastructure storage/retrieval/analysis- facility** will also ensure future access and use of such data. This contrasts to today’s situation where most imaging data are stored by the individual researcher, in practice being lost for reprocessing by others. This application also

---

<sup>1</sup> Nasjonal Strategi for eInfrastruktur

aims to address the precarious situation with regard to storage/retrieval spelled out in the national strategy.

Through utilizing and making accessible software developed at the leading universities within the field **we will ensure access to the most relevant international resources for the Norwegian research environment.**

The establishment of a permanent facility answers the strategy recommendations on developing and retaining competence within use and operation of eInfrastructure as well as directly addressing the need for permanent establishment of such services for users.

Our application also increases the resources devoted towards providing advanced support towards specific application areas as well as developing a separate apparatus for user support towards storage and analysis. We wish to underline that appropriate software is not yet commercially available. The researchers having access to state of the art support will in our opinion have a great advantage over those still having to cope with non-coordinated storage/retrieval systems and manual analysis.

In the draft *Nasjonal plan for forskningsinfrastruktur for medisinsk avbildning i Norge*<sup>2</sup> it is clearly stated that bioimaging informatics "is an integrated part of the research and that one needs to develop and tailor a broad spectrum of methods" and that this includes "automatized image analysis and different forms of Compute Aided Diagnostics (CAD) that needs to be tailored for non-expert users". The draft plan further states that "equipment and running costs for such activities should be included in the national plan" and "erection of state of the art facilities for image analysis and advanced visualization where researchers with technological and medical backgrounds can meet" are important. The draft plan has been submitted to the the Universities of Bergen and Oslo as well as NTNU.

**After the initial startup period we will integrate parts of the imaging facility with services that are provided by the High Performance Computing unit of the University of Bergen.** This collaboration will be important for meeting the needs of complex imaging problems in science that require the use of high performance scientific computing.

*The national character of the new research infrastructure.*

The proposal constitutes **new research infrastructure** that will be an integrated part of the FUGE-funded NorMIC consortium (<http://www.uib.no/med/nor-mic/>) that constitutes **the national imaging platform. All researchers from Norway within imaging will have equal access** to the facility outlined in this application and this will be communicated explicitly. Apart from our listed partners, we will seek to coordinate our activities with the FUGE-funded Bioinformatics platform when relevant. A common Norwegian bioimaging informatics facility will constitute a highly efficient manner in which to provide such services, **and support will be provided online as a grid infrastructure. We wish to underline that no such facility exists in Norway today.** The fact that this application is supported by the whole NorMIC consortium and the immense interest shown for presentations on Bioimaging informatics at the NorMIC annual symposium in November 2008 in itself illustrates the **broad national interest** for such a facility. Our effort will be aimed at the wider imaging

---

<sup>2</sup> National plan for research infrastructure for medical imaging in Norway) by the professors Olav Haraldseth (NTNU), Ole Sejersted (UiO) and Odd Helge Gilja (UIB). The draft has been elaborated as a suggested specification of national plan for research infrastructure on research areas covering imaging.

research environment in Norway in biology and basic and clinical medicine. The proposed facility will thus **lay an important part of the foundation for future leading research within imaging related research in Norway.**

### *Critical factors of the project*

The main parts of this facility will consist of personnel and software, the latter being developed/made accessible through the former. The most crucial software and hardware have already been tested in several projects at the University of Bergen. When we get an increasing number of users, we will use the expertise of the personnel at the High Performance Computing unit in Bergen help us with computer cluster related problems. We have highly qualified staff that has made considerable breakthroughs in the field and these will play a crucial role. The physical part of the facility will also be crucial, but will, apart from localities provided by the University of Bergen, mostly consist of shelf ware hardware that is easily obtained, maintained and replaced. Related to the NorMIC platform we have already highly qualified and experienced personnel within the field of storage, retrieval, analysis and visualization. The new facility will build upon these existing human resources.

## 3. Description of the Research Infrastructure

**We see a threefold solution to the, in parts 1 and 2, outlined challenges.** Firstly, images and accompanying meta-data must be stored securely and centrally so that they can easily be retrieved. Secondly, software tools for image analysis and quantification must be prepared, implemented or developed and in some cases a user friendly interface must be created to improve the accessibility to the users for practical use. Thirdly, support must be provided for both the storage/retrieval procedures, and for the development and use of advanced software tools as well as for tailoring both to specific projects. Neither of these services is to any relevant degree available in Norway as of today. We also envisage a facility that will do R&D on image analysis methods and turn findings and other available open source software into services.

**R&D challenges** in particular involves development and modification of software for

- Segmentation of cells (surface stained, cytoplasm stained and DIC images) and subcellular components
- Automated detection and localization of signals from fluorescence stained biomolecules
- Tracking of objects (cells, organelles / vesicles and molecules) in dynamic recordings
- Registration (spatial alignment) of images taken at different times or under different experimental conditions for assessment of changes in size, shape or signal intensities, or of objects recorded with different imaging modalities for multispectral analysis and pattern classification
- Analysis of functional 3D+time imaging series in small-animal MRI related to perfusion and diffusion
- Modelling of cells behaviour and biological processes – towards image-based systems biology.
- High Content Microscopy microscopy analyses (cell based screening)

With regard to service connected to storage we aim to install webservers for storing image data and database servers for storing meta-data that will enable users to search and retrieve images based on both annotated information and analysis results. Our main focus is on security but we will enable image searching across collaborating projects when this is requested from the project owners. Transfer and retrieval of data between the facility and

outside users will be performed either through a webserver, manually through Secure File Transfer Protocol (SFTP/SSH) or via custom java viewer with SFTP/SSH functionality. Conversion of image formats will be available in both Open Microscopy Environment installations as well as in our java imaging software.

### ***Technical requirements for large scale image analysis***

Most existing open source imaging software require both database accessibility and a webserver for storing and organizing images. Cell Analyst requires that the data from Cell Profiler is stored in a database. A web server for storing images is also recommended since this enables easier image retrieval when viewing the results.

Large scale image screens often have different requirements both in terms of hardware and software. It is therefore necessary to make custom scripts and software for automation of different steps in the analysis and visualization of results. We plan to implement solutions in three different ways:

1. Open Microscopy Environment (OME)
2. Cell Profiler and Cell Analyst
3. Custom analysis software

1. The Open Microscopy Environment (OME) consist of several different software components for visualising, managing, and analysing microscope image data. We plan to set up an infrastructure that will enable us to host custom installations of OME on our servers that are specifically set up in regards to the requirements of each imaging project. Different software components and algorithms from existing software (Imaris, ITK, VTK, FSL, MIPAV, Matlab) can be integrated into such systems, and enable scientists to automate all necessary steps that are required when doing large scale imaging screens. Since OME require a high degree of technical expertise and a server for installation, it is difficult for scientists who do not have an IT background to set this up by themselves.

2. Cell Profiler and Cell Analyst can be used for a variety of experiments and was specifically made so that scientists in life sciences could analyse their own data. It uses MySQL to store image metadata and results from analyses. MIC plans to help users with setting up Cell Profiler in two different ways. Either on their own computer connected to a MySQL database at MIC, or doing the Cell Profiler analysis for customers through OME on our servers.

3. Some project require a very specific setup, where the use of OME and Cell Profiler is too time consuming. The experiments in these projects can be automated with custom software on the servers at MIC.

Some project require a very specific setup, where the use of OME and Cell Profiler is too time consuming. The experiments in these projects can be automated with custom software on the servers at MIC. We currently have several projects at the department of biomedicine where a custom solution for large scale analyses has been developed to suit project needs. We list these as illustrative examples.

- The craniofacial development group has a custom database and software for annotating images from different developmental stages. This system is integrated with segmentation and 3D reconstruction software (Java, Java3D and VTK) and has been continuously developed since 2003.

Coordination of trigeminal axon navigation and patterning with tooth organ formation: epithelial-mesenchymal interactions, and epithelial Wnt4 and Tgfbeta1 regulate semaphorin 3a expression in the dental mesenchyme.

Kettunen P, Løes S, Furmanek T, Fjeld K, Kvinnsland IH, Behar O, Yagi T, Fujisawa H, Vainio S, Taniguchi M, Luukko K.

Development. 2005 Jan;132(2):323-34.

- A similar database that was developed for the large scale enhancer trap project at the transgenic zebrafish facility at Sars Centre at UIB.

Large-scale enhancer detection in the zebrafish genome.

Ellingsen S, Laplante MA, König M, Kikuta H, Furmanek T, Hoivik EA, Becker TS.

Development. 2005 Sep;132(17):3799-811.

- A joint project between Kai Ove Skaftnesmo from the translational cancer research group and craniofacial development group has resulted in custom software for the BD Pathway High Content Screening microscope (HCS). The software is a prototype that was applied in a screen for miRNAs that can regulate the proliferation of the U87Mg and U251Mg glioma cells.

The software from these projects is well tested and can be reused and modified to suit other projects at the Imaging facility in the future.

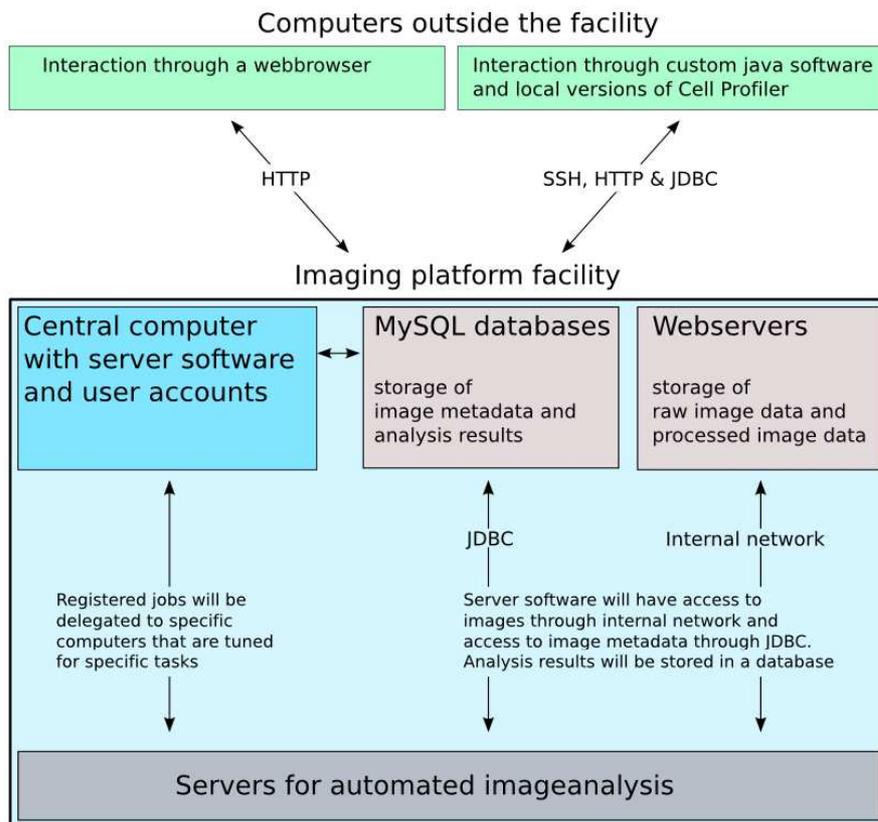
#### *Type of services and functions*

- Perform complete analysis of data where MIC will run the complete analysis of the images and present it to the user through a web server and custom made visualization software.
- Perform partial analysis of data where users will be instructed how they can run their own analyses using the system at MIC
- Host database accounts for users who analyse their own data with Cell Profiler
- Participation in open source projects by creating user-friendly interfaces for existing imaging tools that will facilitate analyses for scientist in life sciences.
- Create user documentation which is important for continuity across human resources.

With regard to support we aim to provide support on all levels ranging from

- Providing introductory courses and necessary support
- Setting up/tailoring software for specific analysis series and carrying full series of especially demanding analysis.
- Moreover, advanced courses in biomedical image processing, analysis and visualization will be connected to the new “Nasjonal tverrfaglig forskerskole i medisinsk avbildning” (Interdisciplinary Research Training Program in Medical Imaging), headed by Prof. Olav Haraldseth (NTNU), where this “Forskerskole” is already known to the NorMIC consortium and will be adapted to their educational needs.

## Overview of technical equipment and software



HTTP - Hypertext Transfer Protocol is an application-level protocol for distributed, collaborative, hypermedia information systems

SSH - Secure Shell is a network protocol that allows data to be exchanged using a secure channel between two networked devices.

JDBC - Java Database Connectivity is an API for Java that provides methods for querying and updating data in a database.

Outside users will be able to interact with the imaging platform through:

- Web based userinterface using a internet browser (HTTP)
- Custom java software (SSH,HTTP, JDBC)

Users will connect to the central server to register images for analysis. The central server will delegate jobs to the imageanalysis servers. Images will be stored on imageservers. Image metadata will be stored in a MySQL database

#### 4. Plan for access and use, data and knowledge management

**The facility will be made available on a pay-per-service basis at equal terms for all university sector users in Norway.** The services will be explicitly advertised and presented in all relevant forums we can get access to and included in NorMICs marketing effort. The national and international networks of which MIC-Bergen is a partner (NorMIC, Nordic Imaging Network and Medviz) are natural first line arenas into which to disseminate our work and put our developments and service to good use towards potential users. Outside the

immediate university environment the University hospitals and the institute sector are main target for our information effort.

During the setup phase one might have to limit the number of projects. However, as projecting skills outside the facility is also a crucial goal and services become standardized, researchers and technical personnel outside the facility will over time be able to use the facility online with less assistance. **Due to the nature of the facility we can rather easily increase its capacity and acute capacity problems are not probable after the initial setting up phase of the facility.** Should however such problems arise, e.g. with regard to tailoring of software, we can address these by increasing the user fee, thus encouraging parsimony with bottleneck services.

**It is an explicit goal of the facility that software and data should become available to other than the owner for later analysis.** Our developments will be published through the relevant scientific journals and presented at international conferences and workshops. Due to using open source software and wanting as many researchers as possible to use our development to the benefit of their own research we will not seek to patent any of our developments. By providing both user friendly developments of current and future software and services suited to advanced users we will seek to disseminate our developments and service as deeply and rapidly as possible into the scientific community. It is an explicit goal to share our methods and make researchers as independent as possible in utilizing these to the benefit of their own research. An important step in this direction is to create user documentation at the level of the target users. **We will also provide support on all levels ranging from introductory courses, providing necessary introductory support, setting up/tailoring software for specific analysis series to carrying out full series of especially demanding analysis**

In cooperation with the Department of Mathematics, UiB **we will train and employ students** to do program development within image analysis and in cooperation with the Department of Informatics (UiB) we will make a similar effort within advanced visualization. Moreover, advanced courses in biomedical image processing, analysis and visualization will be connected to the new Nasjonal tverrfaglig forskerskole i medisinsk avbildning (Interdisciplinary Research Training Program in Medical Imaging), headed by Prof. Olav Haraldseth (NTNU). This Forskerskole is already known to the NorMIC consortium and will be adapted to its educational.

**Universal access to data** has in previous projects shown itself to be a difficult issue and if researchers refrain from using our facility for storage/retrieval and analysis because they fear for the security of their data we have failed our mission. Thus, we will not provide universal access to the data of any researcher without contractual consent. We will however strongly encourage universal access for the benefit of the wider scientific community once the owner has published on the data.

Despite **dissemination to the wider public** not being one of the explicit requirements of this call we nevertheless see it as important. Dissemination to the public will be carried out as an integrated part of NorMICs public dissemination effort. Worth mentioning here is that imaging data processed with bioimaging informatics software is already at display in MIC-Bergens part (60 m<sup>2</sup>) of a permanent exhibition at the Norwegian Technical Museum in Oslo.

## 5. Impact on research and innovation

**A bioimaging informatics facility will be a strong resource for the whole Norwegian community with interest in biomedical research** since it represents a common base and discussion forum for available analysis methods and necessary improvements. It should also be of note that such a project will advance the interdisciplinary work between Norwegian universities, and also across educational borders since the facility will require the daily input of biology, informatics and mathematics.

The facility will constitute a resource center for image related research on biomedical images, which will greatly assist in the quantification of biomedical processes. By offering services for the whole Norwegian research society in the field it will contribute efficiently to the biomedical research for quantification of large amount of data. **It will also be an efficient way to produce reproducible and comparable results** between research groups since the applied methods will be accessible to various research communities. This common interaction and base will also probably **enhance the collaboration between different research communities**.

**The facility offers good opportunities to incorporate the education of younger scientists** within mathematics, image processing and computer science by assigning them well-defined and specific projects related to the services of the facilities. However, to become successful, this strategy needs a strong and clear leadership defining the goals and the methods to be implemented. As mentioned above we with regard to this aim to employ students from our partner departments under tight supervision. Our partners at Departments of Mathematics and Informatics have dedicated a sizable portion of their time to this task.

**A well-functioning facility is a powerful element for creating international collaboration** since it enables easy distribution of own knowledge and software and it is a fundament for future common projects. Such collaborations are often bidirectional and the feedback from international collaborators will also contribute to the development and improvement of the facility. .

The study of mammalian cells is an example of a task with high potential results. A well-functioning high-throughput imaging system in combination with automated image analysis **e.g. represents a powerful tool for screening of cells and the effect of drugs on various cell systems**. Commonly available computational resources can also **save human time** on a larger scale since the approach for solution is not reinvented every time it occurs. By using a common facility for analysis different research communities can also **easier compare and discuss obtained results**. One can also foresee an operational system such that all users of the facility provide feedback to improve its use, accuracy and flexibility for future tasks.

## 6. Partners and scientific institutions

Beside **MIC-Bergen, the leading node of the NorMIC consortium**, the partners in the project are the following persons and their institutions:

**- Tomasz Furmanek, Department of Biomedicine, UiB**

Furmanek is a PhD student working with large scale image analysis and 3D visualization of gene expression in developmental biology. He has a Master degree in molecular biology where he worked with 3D visualization and quantification of viability and gene expression in thick tissue samples. He has worked as a programmer on several scientific projects during the past years:

- Design and implementation of image database, integration with fish room database and custom online screening software at Sars Centre, UIB. (Ellingsen S, Large-scale enhancer detection in the zebrafish genome. Development. 2005)
- Internet based education at the Department of Chemistry and Department of Molecular Biology.
- Evolutionary Image Array (EvoArray) software, for investigating the complete set of expressed microRNAs and corresponding mRNA targets in various developmental stages and tissues. (Johansen SD, 2009, Large-scale sequence analyses of Atlantic cod, New Biotechnology, Review In Press April 2009)
- Proficient in Java, Java Advanced Imaging, Java 3D, databases and web based user interfaces.

**Role:** Partner on implementing the database system for metadata and image storage and retrieval, and incorporation of software solutions for high-throughput image analysis.

**Furmanek represents a nationally unique combination of information technology skills and an extensive biomedical education and experience, and is one of the best to implement the initial phase of the proposed infrastructure.**

#### **- Dr. Erlend Hodneland, Department of Biomedicine & The MedViz consortium, UiB**

Hodneland has his background from applied mathematics, fluid mechanics and image processing. In his master thesis, he worked on segmentation of MRI images, and in his PhD work he has designed a comprehensive and advanced software system and evaluation framework for segmentation and analysis of cells marked with fluorescent markers. During this work, also novel image analysis methods (implemented in MATLAB) were contributed to the cell biology community. This work has given him a broad experience in everyday challenges in the analysis of biological systems. The PhD work has also been an important example of true interdisciplinary work between mathematics, image processing and cell biology, shown to be very useful and stimulating for both research communities.

**Role:** Partner on interfacing the advanced cell segmentation routines and image analysis algorithms he and others have developed to add important functionality to the infrastructure.

#### **- Prof. Arvid Lundervold, Department of Biomedicine / MIC-Bergen & MedViz, UiB**

Arvid Lundervold has educational background in both mathematics and medicine. Currently he is a professor in medical information technology at the University of Bergen, Department of Biomedicine, and head of the Neuroinformatics and Image Analysis Laboratory in the Neuroscience Research group. Lundervold is also member of MIC-Bergen, BBG and MedViz, and affiliated as a researcher with the Radiology Department, Haukeland University Hospital. Since 1987, he has participated in a broad range of biomedical imaging projects related to quantitative image analysis and pattern recognition in both Norway and in the

European context of COST, and supervised several MSc and PhD students with background from applied mathematics, physics, computer science, physiology and medicine.

**Role:** Project leader and interdisciplinary coordinator.

**- Prof. Antonella Zanna Munthe-Kaas, Department of Mathematics & MedViz, UiB**

Prof. Zanna Munthe-Kaas is a core member of the BBG image processing group and member of the MedViz consortium leader team. She is involved in several interdisciplinary biomedical imaging projects and a main driving force for image processing research and education at the Department of Mathematics.

**Role:** Partner on advanced image processing, analysis and mathematical modelling with participation in algorithmic developments and supervision of student projects that contributes to the build up of the infrastructure.

**- Prof. Klaus Johannsen, BCCS / Parallab & Department of Mathematics, UiB**

Prof. Johannsen is head of the High Performance Computing Center (Parallab) in Bergen and adjunct professor in applied mathematics. His field of research is partial differential equations, numerical simulation and parallel computing. In Parallab, there is also valuable competence in Grid computing (e.g. Dr. Csaba Anderlik), and Grid technologies are already involved in medical applications related to brain imaging and functional kidney imaging through the NOTUR Advanced User Support program..

**Role:** Partner on high performance computing and Grid technologies. Prof. Johannsen will contribute to design and give advice regarding infrastructure build-up in order to be scaleable and suitable for Grid solutions that can serve a high volume of users and imaging project at a national level.

**- Assoc. Prof. Ivan Viola, Department of Informatics & MedViz, UiB**

Ivan Viola has long-lasting research experience and many publications and contributions in the field of illustrative and medical visualization, both as project manager at the Institute of Computer Graphics and Algorithms, Vienna University of Technology, and more recently as an associate professor in the Visualization group, headed by Prof. Helwig Hauser, Department of Informatics, UoB.

**Role:** Partner on advanced visualization technologies. He will contribute to the design of software solutions and supervise student projects that will add much-wanted visualization functionality to the infrastructure.

Along with the project and according to needs, all partners will bring in their national and international research network which can be relevant for establishing or improving specific components of the infrastructure (e.g. Grid middleware, image processing algorithms, visualization tools). The infrastructure project will also profit from the ERASMUS student exchange agreements that UoB has with image processing labs in France and Poland, as many programming tasks can be very suited for short-term student projects.

## 7. User groups and international cooperation

We think that **the molecular imaging community needs, the level of national interest and participation and the potential user groups** are rather well covered in the above parts of the application and we thus recapitulate briefly. The bioimaging informatics facility will be of great interest to and available for all researchers within the relevant fields. The relevance and level of interest was well illustrated at NorMIC annual symposium where the acclamation towards presentations of **software for automated image analysis** was immense and a clear need explicitly stated. The FUGE-funded NorMIC consortium also voices its explicit support for improving the eInfrastructure within imaging in Norway in the attached letter of support.

**In order to highlight some specific examples from the immediate research environment** we like to mention the following fields and research issues as highly relevant for a facility as described in this application: Quantitative analysis, automated segmentation, and co-registration of 3-D and 3-D+time structural and function small-animal high field magnetic resonance image acquisitions; High-throughput microscopic scanning and quantitative image analysis of breast cancer cells; Proper analysis methods for colocalization of vesicles in mammalian cells expressing specific proteins; Image-based analysis (and modeling) of the division and growth of epithelial cells in the formation of blood vessels, and the influence of drugs on this process; Development of matrices for reconstruction of human tissue; Quantitative analysis (PCA) of molecular movements from molecular simulations in Amber; Large scale image analysis and 3D reconstruction in craniofacial development biology; Custom software for large scale analysis of miRNA screens using BD Pathway High Content Screening microscope.

Moreover, from a previous FUGE 2 channel 3 application (that was not funded): **“Advanced Quantitative Analysis and Modelling in Molecular Imaging using High Performance Computing”** we addressed the following three biology-driven applications: **(i) *Quantitative Imaging of Synaptic Plasticity using Multi-Photon Laser Scanning Microscopy (MPLSM) combined with Electrophysiology***, **(ii) *Quantitative Imaging of Membrane Dynamics in the Living Cell***, and **(iii) *Quantitative Imaging and Modelling of Tunnelling Nanotubes (TNT) -based Cellular Networks***. In the planning of all these sub-projects both MIC Bergen and several professors at the Department of Biomedicine and at the Department of Mathematics, as well as Bergen Center for Computational Science / Parallab and international partners in Germany and UK were involved. This situation demonstrates that **with funding of the proposed e-infrastructure, a revitalization of such interdisciplinary project plans can be put into action on a relatively short time-frame, and thus speed up cutting-edge research within molecular cell biology and bioimaging informatics.**

Most software for large scale imaging is developed as open source projects. Designing our system around software like Open Microscopy Environment will **enable us to participate in several scientific open source projects**. Instead of using resources on designing local solutions from scratch, **we can base our systems on well tested source code and the experience of leading international research groups**. The open source community is a classic example of well functioning **international collaboration** where groups with similar goals are able to produce effective solutions. Many of the imaging tools that are published as software articles are registered as open source projects since this is a requirement in many scientific journals. This enables sharing of expertise and resources on an international level.

## 8. Management plan and localisation

### *Project management and responsibilities*

The bioimaging Informatics facility will be managed as an integrated part of MIC-Bergen and its scientific leader will be Professor Arvid Lundervold and as leader of both MIC-Bergen and NorMIC, Associate Professor Frits Alan Thorsen will also be strongly involved. The facility will benefit from MIC-Bergens website, booking and invoicing system as well as ample administrative support and highly skilled technicians within imaging techniques.

The Bioimaging informatics facility will have a 50% researchers position. It is planned for this position to be covered by Erlend Hodneland or a person of similar qualifications in order to enable further development of his software designs into user friendly and fully fledged image analysis software (see publications 3, 6 and 8 in Hodnelands attached CV). Thomasz Furmanek will have a central role as the facilitys main programmer. He will be responsible for coordinating the initiation of systems for storage and retrieval as well as for modifying the current suite of image analysis software mentioned in part 3 of this project description. He will also work towards modifying the designs of Erlend Hodneland..Another important responsibility will be to adapt/tailor software for specific projects. Over time he will retreat into more of a supervisory position delegating tasks to other affiliated programmers and a pool of students from the Departments of Mathematics and Informatics, seeking to ensure recruitment to the field.

Of our partners at the Department of Mathematics, professor Antonella Zanna Munthe-Kaas will train and supervise students in image analysis programme development for our facility and Senior Researcher Klaus Johannsen Parallab/Department of Mathematics will function as an advisor regarding high performance computing and GRID opportunities. Our partner at the Department of Informatics, Associate professor Ivan Viola, will train and supervise students in programme development within advanced visualization. Senior

The facility will be continually upgraded and most of costs associated with upgrades are in terms of soft-, not hardware. Thus our most important resource will be that of alert, enthusiastic and competent staff that will develop the field further and integrate into the facility progress made in-house as well as elsewhere.

### *Justification of the suggested localisation*

MIC-Bergen is set up with a broad range of imaging techniques including Confocal-, Multiphoton-, Spinning Disc- (live cell), Fluorescence- and Electron microscopes, two Small Animal Molecular Imagers, a High Content/High-Throughput microscope and a small animal MRI. A small animal PET/CT will soon be added. MIC-Bergens users come mainly from the western region but also from the rest of Norway. Further, through its status as the leading node within the NorMIC consortium, its association with the the whole animal molecular imaging network (NTNU professor Olav Haraldseth and UiO/Rikshospitalet professor Ole Sejersted), its active membership in the Nordic Imaging Network<sup>3</sup> and its close ties with Medviz, MIC-Bergen has a ready network outside of its immediate users to which to

---

<sup>3</sup> MIC-Bergen will e.g. host its annual symposium in 2009

disseminate the methods and services of the suggested facility. Though the Bioimaging Informatics facility is meant to cater for the whole country, we see it as important, especially initially, to have close direct ties to a wide range of potential users

Another moment of great importance is the links to young and forth coming as well as well established researchers within the field. Such highly competent and motivated personnel is an absolute necessity for the establishment and maintenance of such a facility.

**MIC-Bergen as a location thus has both the whole range of instrumentation as well as established contacts with a wide range of potential users both locally and nationally. It has personnel with high motivation and key competences both in its immediate vicinity and as partners and a good recruitment ground. It is in our view an ideal site for such a start up facility.**

#### *Host institutions long term planning and research strategy*

Even though the Bioimaging Informatics facility will be primarily the responsibility of MIC-Bergen, the imaging- and bioinformatics clusters at UIB and HUS are strong and in rapid development. The UiBs interest in the maintenance of MIC-Bergen is e.g. illustrated through earmarked positions and funding and through the priorities of the Faculty of Medecine and Odontology (MOF) and Department of Biomedicine.

MOF in its Research Report (*Forskningsmelding*), which is an attachment to its strategic plan, states that MIC-Bergen should be “continued in accordance with the intention in the FUGE programme”. As previously mentioned FUGE has voiced the importance of “a strategy for how to...increase the focus on bioinformatics and data handling..”. For the Department of Biomedicine, the host department, MIC-Bergen is “an important part of its strategic effort” and is “advantageous in terms of giving highly competent staff and high quality services” and is “rational as one does not want several environments with suboptimal competence”. The emphasis upon MIC-Bergens imaging services is also shown through the departments localization of approx 2,5 mill NOK annually (5 % of its basic budget) towards MIC-Bergen. It is also worth mentioning that the Department used 700 000 NOK of its strategic funds on equipment upgrades for MIB-Bergen in 2009.

## 9. Time-schedule and deliverables

The first task for the facility will be to modify the exsisting systems we have for storage and image analysis to suit it for processing data from external users. The facility already has access to a robust and flexible system for data handling that only requires small changes to work with our planned setup. This is a stand-alone goal since the data handling can be used by anyone, regardless if they use any further services provided by the facility.

The second task will be to install a custom installation of Open Microscopy Environment where Cell Profiler, Matlab and other software can be integrated. A proper documentation and must be provided at all stages. All setup will be done with automated scripts so that new servers can be added in a flexible and effective way.

The third task will be to choose a common field of research like cell segmentation and integrate available systems for cell segmentation into the services. Successive fields will be elaborated upon

When the basic parts of our system is operational, we will focus on automating tasks and customizing the user-interfaces on our main server which will register new jobs for analysis and delegate tasks to our grid of computers.

As soon as certain services for analysis are available, these need to be published on a website to potential users and a proper training must be offered. Over time, extended biomedical tasks related to image processing can also be integrated into the facility. However, one should also keep in mind the need for fast aided help in connection to very specific problems of little common interest. In these cases, one should also be able to offer analysis help which not necessarily is included into the public part of the facility.

Tasks	From	To
Modify and establish system for data handling (stor./retr.)	2009 3	2009 3
Custom installation of Open Microscopy Environment	2009 3	2009 3
User documentation on data handling and OMERO	2009 4	2009 4
Integrating/Customizing applications for fields of research	2009 4	2012 2
Automating tasks and delegating task to our computer grid	2009 4	2012 2
Customizing user interfaces	2009 4	2012 2
Offering training to users	2009 4	2012 2
Offering courses to users	2010 1	2012 2
Offering help with complete analysis	2010 1	2012 2
Integrating students into operation	2010 2	2012 2

## 10. Budget and funding plan

### Total Costs in categories (in NOK 1000)

	2009	2010	2011	2012	Sum
<b>Payroll and indirect expenses</b>	1063	2211	2299	1196	6769
<b>Procurement of R&amp;D services</b>					0
<b>Equipment</b>					0
<b>Other operating expenses</b>	233	119	124	65	541
<b>Totals</b>	1296	2330	2423	1260	7310

Payroll and indirect expenses: 10% of a professor at Department of Biomedicine (IBM), 10% of a professor at Department of Mathematics, 5% of an associated professor at Department of Informatics, 15% of a Senior Researcher at Parallab Bergen Center for Computational Science (BCCR), 5 % of a professor at Department of Mathematics, 10% of a research coordinator at IBM, 10% of a senior engineer at IBM, 50% researcher at IBM (whose costs and running costs are based on the NRC standard for postdocs), the equivalent of a 100% position as programmer/program developer at IBM and the equivalent of a 50% research assistant position to be used to employ several students part time at Department of Mathematics and Informatics. The latter will initially be a single highly skilled computer engineer whose position will be scaled back as services come up and running. After the initial phase it is our intention to include computer science students on a project basis.

Other operating expenses covers a one time investment in hardware of 200k NOK in 2009 and updates of hardware of 50k NOK pr year thereafter as well as software licences are set at 70k NOK pr year.

Note that for all expences a increase in price and salaries of 4% is included.

### Finacial plan (in NOK 1000)

	2009	2010	2011	2012	Sum
<b>Own financing</b>	277	576	599	312	1764
<b>International funding</b>					0
<b>Other public funding</b>	0	125	187	125	436
<b>Other private funding</b>					0
<b>From Research Council</b>	1019	1630	1637	824	5111
<b>Totals</b>	1296	2330	2423	1260	7310

Own financing covers a 10% professor position at IBM, 10% of a professor at Department of Mathematics, 5% of an associated professor at Department of Informatics 10% research coordinator and a 10% senior engineer. Costs for general infrastructure/localities is not listed but will be covered by the host institution. Other public sector funding constitutes gradually increasing user fees peaking at 20% of running expences (researcher+programmer+hardware and software).

Funding after the project period from the host institution will depend on in which form MIC-Bergen is continued and the degree to which the Bioimaging Informatics facility will be deemed successful. To promise funding after 2012 in this application would thus entail little substance. We do however wish to point out that the main cost of such a facility is in the first years of its operation and that there are examples of e.g. storage/retrieval facilities, once well established, having run for several years thereafter with minimal resources.