Medical Informatics
'Discipline Gebonden Onderwijs'
Bioinformatics
2010

Coordinator:
Prof. Dr. Antoine van Kampen
Bioinformatics Laboratory, Department of Clinical Epidemiology, Biostatistics and Bioinformatics, J1b-217, Tel. +31 (0)20-5667096, Email: a.h.vankampen@amc.uva.nl
www.bioinformaticslaboratory.nl

Other teachers:
Dr. Marc van Driel
Netherlands Bioinformatics Centre (NBIC), Nijmegen

Dr. Andrew Gibson
Semantic Biosystems, BioSystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam

Dr. Silvia Olabarriaga
Bioinformatics Laboratory, Academic Medical Centre

Dr. Han Rauwerda
Microarray department, Swammerdam Institute for Life Sciences, University of Amsterdam

Dr. Marco Roos
Bioinformatics Expertise Centre, LUMC, Leiden

Barbera van Schaik
Bioinformatics Laboratory, Academic Medical Centre

Dr. Frans Voorbraak
Clinical Computer science, J1B-114, Tel. +31 (0)20-5665853, Email: f.p.voorbraak@amc.uva.nl

Aim

The module Bioinformatics aims to give the student a more comprehensive definition of bioinformatics related to medical informatics. After completing the module the student will have a clear definition of the relation between bioinformatics and medical informatics. At the same time the student will have a more detailed definition of his/her role in "life sciences". Several subjects are being presented that are relevant for both bioinformatics and medical informatics. Besides, skills are gained in the preparation and leading of discussion meetings, compiling papers and presentation.

Content

Bioinformatics and medical informatics

Bioinformatics is a fast developing interdisciplinary research area which focuses on computational "management" (storage, processing, analyses, etc.) of biological data. Through new techniques and larger research efforts (such as the human genome project) biological data are made available at high speed, especially in the field of molecular biology (such as DNA sequencing, measurement of gene activity and protein structure). The (computer supported) processing and analyses of these data to create more insight in (medical-) biological processes generates a great challenge but this will eventually lead to considerable progress in medical science. It seems of crucial importance that the disciplines Bioinformatics and Medical informatics are integrated. Although the fields of study Bioinformatics and Medical informatics are related, there are also important differences. The similarities and differences are explored in this module.

E-(bio)science/GRID

E(nhanced)-bioscience refers to scientific projects within the domain of Life Sciences. Those are being performed in distributed collaborations with use of advanced technologies from computer science and ICT. The objectives of e-Bioscience include amongst others the incorporation in the Life Sciences of (large data-generating) research facilities such as measuring equipment (e.g., mass spectrometers and MRI), of large biological and medical data collections, of computer power and data storage facilities, but also facilities for e.g., scientific visualization and data analysis. The joint use of facilities is one of most essential components of e-Bioscience. Therefore, contrary to traditional collaborations, e-Bioscience uses new technologies from the computer science to extend the efficiency of data (generating) facilities within collaborations and at the same time secure the access to these facilities and the matching scientific expertise. When this will be combined with data modeling and standardization (e.g., by using the corresponding ontologies and data analysis protocols) e-Bioscience will extend the quality and reproducement of scientific research. Web and Grid technology make e-Bioscience possible.

Open source, Science Commons and WIKI

A WIKI is a (web)application, in which web documents can be edited by a community. A well-known example of a wiki is Wikipedia. Wikipedia is a project of the Wikimedia Foundation. The purpose of Wikipedia is to create a complete, freely spread and editable encyclopedia on the internet. The articles in this encyclopedia are being esteemed to have a neutral point of view. However Wikepedia does not give any guarantee in accuracy and stable quality of this information, as any visitor can edit these articles. Besides, vandalism is a problem due to the open character of the project. All together Wikipedia's are being build by many thousands of people all over the world: scientists, hobbyists, students, experts and people with good general education. Everyone can contribute to Wikipedia, also anonymous. There is no limit to contributions, some do it incidental, others daily. The full text can be edited by anybody, so if a person starts a short article, a number of other people can continue till the document is a valuable encyclopedic article. Meanwhile Wiki's are being used more and more in science.

Open-source software projects concern a specific development method for software, starting from providing of the source code of software. This way a possibility is being created to adapt a program by anyone with knowledge to understand the source code. A lot of software within bioinformatics is open-source.

Science Commons designs strategies and tools for faster, more efficient web-enabled scientific research. We identify unnecessary barriers to research, craft policy guidelines and legal agreements to lower those barriers, and develop technology to make research, data and materials easier to find and use. The goal is to speed the translation of data into discovery — unlocking the value of research so more people can benefit from the work scientists are doing.

Wiki, Open-source and Science Commons assume of the idea that joint effort by a large community will lead till faster and/or better results.

Academic skills

In addition to the technical aspects of this module a lot of attention will be paid to academic skills such as giving a presentation, leading discussions and writing a paper. In preparation to the master the medium of instruction during the module will be English. All presentations, discussions and papers will be done in English. The final presentation will also be done in English.

Module Planning

- Week 1. During the first week a short introduction on bioinformatics will be given as well as a concrete example of how bioinformatics is used in practice. Furthermore a few subjects will be discussed which are relevant for both Bioinformatics and Medical informatics. This first week students will have the opportunity to have discussions with different experts on the relevance of a specific subject for Medical informatics.

- Week 2-3. 3 projects will be carried out by groups of students. They will report by means of presentations. In addition two general presentations (WEB2.0 and How to write a research proposal) will be given by two students.

- Week 4. During the last week final presentations will be given on Project 1, 2, 3 and 4.

Projects: General introductions.

The information supplied and gained within the scope of Project 1 (Medical informatics and Bioinformatics), Project 2 (e-Science/GRID) and Project 3 (Open-source/Science Commons/Wiki) is essential for Project 4 (research proposal) that will be carried out by all groups. *It is imperative that all students accurately study the documents and publications mentioned below* (and not only those publications which are needed for your presentation). This enables discussion and, as said earlier, forms the basis for Project 4.

Project 1 (group A). Relation between bioinformatics and medical informatics

To gain insight on the relation between bioinformatics and medical informatics group 1 will carry out literature research and will give a presentation on this subject. The basis are the below mentioned articles. It is expected that more recent articles are being searched for and used. In particular, it is expected to the students retrieve and present (at least) one recent project from scientific literature that clearly demonstrates the added value of the integration of bioinformatics and medical informatics.

- Miller PL. (2000) Opportunities at the intersection of bioinformatics and health informatics: a case study. *J Am Med Inform Assoc*. **7(5)**, 431-8.
- Martin-Sanchez F, Iakovidis I, Norager S, Maojo V, de Groen P, Van der Lei J, Jones T, Abraham-Fuchs K, Apweiler R, Babic A, Baud R, Breton V, Cinquin P, Doupi P, Dugas M, Eils R, Engelbrecht R, Ghazal P, Jehenson P, Kulikowski C, Lampe K, De Moor G, Orphanoudakis S, Rossing N, Sarachan B, Sousa A, Spekowius G, Thireos G, Zahlmann G, Zvarova J, Hermosilla I, Vicente FJ. (2004) Synergy between medical informatics and bioinformatics: facilitating genomic medicine for future health care. *J Biomed Inform*. **37(1)**, 30-42.
- Maojo V, Kulikowski CA. (2003) Bioinformatics and medical informatics: collaborations on the road to genomic medicine? *J Am Med Inform Assoc*. **10(6)**, 515-22.
- Altman RB. (2000) The interactions between clinical informatics and bioinformatics: a case study. *J Am Med Inform Assoc*. **7(5)**, 439-43.

Project 2 (group B). E-(bio)science/GRID and applications

In this project the concept "GRID" will be studied more closely. During this presentation the following questions, amongst others, will be answered: What is GRID? What are the basis functionalities of GRID? Who has created the GRID and why? Which possibilities give GRID with respect to the Web (internet)? How to implement an operational GRID? What is the difference between a ‘data GRID’ and ‘computing GRID’?

As a start the following documents and publications can be used:

- Foster (2002) What is the grid? (pdf document)
- Foster (2001) The anatomy of the GRID (pdf document)
- Huang CH, Lanza V, Rajasekaran S, Dubitzky W. (2005) HealthGrid--bridging life science and information technology. *J Clin Monit Comput*., **19(4-5)**,259-62.

Furthermore the concept 'e-(Bio)science' will be studied more closely. During the presentation the following questions, amongst others, will be answered: What is e-

(Bio)science? What are the (inter)national developments? What are existing applications and areas? Does e-(Bio)science has any future? Is e-(bio)science the solution for distributed and multi-disciplinary collaborations? The students discuss 2 concrete applications of e-Bioscience in detail, for example, BIRN (www.nbirn.net) and Integrative Biology (www.integrativebiology.ac.uk)).

As a start the following papers can be used:

- Rauwerda H, Roos M, Hertzberger BO, Breit TM. (2006) The promise of a virtual lab in drug discovery. *Drug Discov Today*. **11(5-6)**, 228-36.
- Olabarriaga SD, Nederveen AJ, Snel JG, Belleman RG. (2006) Towards a virtual laboratory for FMRI data management and analysis. *Stud Health Technol Inform.*, **120**, 43-54.
- Hey T, Trefethen A. (2003) e-Science and its implications. *Philos Transact A Math Phys Eng Sci.*, **361(1809)**, 1809-25.

Project 3 (group C). Open-source, Science Commons and Wiki

In this project the advantages and disadvantages of Open-source, Science Commons and Wiki for science (especially Bioinformatics and Medical informatics) will be studied more closely. What does open-source mean exactly? Which open-source licenses exist? Does open-source mean that exploitation of results is being excluded? How is open-source and Wiki used in science? Which advantages and disadvantages apply? Does Science Commons apply to medical informatics?

As a start the following publications be used:

- Wikipedia (http://www.wikipedia.org)
- Mons B, Ashburner M, Chichester C, van Mulligen E, Weeber M, den Dunnen J, van Ommen GJ, Musen M, Cockerill M, Hermjakob H, Mons A, Packer A, Pacheco R, Lewis S, Berkeley A, Melton W, Barris N, Wales J, Meijssen G, Moeller E, Roes PJ, Borner K, Bairoch A. (2008) Calling on a million minds for community annotation in WikiProteins. *Genome Biology*, **9(5)**, R89.
- Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C. (2008) WikiPathways: pathway editing for the people. *PLoS Biology*, **6(7)**, e184.
- Streeter JL, Lu MT, Rybicki FJ. (2007) Informatics in radiology: RadiologyWiki.org: the free radiology resource that anyone can edit. *Radiographics*, **27(4),** 1193-2000.
- Koru G, El Emam K, Neisa A, Umarji M. (2007) A survey of quality assurance practices in biomedical open source software projects. *J Med Internet Res.*, **9(2)**, e8.
- Open source software: a primer for health care leaders. http://www.chcf.org/documents/healthit/OpenSourcePrimer.pdf
- Science Commons: http://sciencecommons.org/

Project 4 (all groups). Research at the interface of Medical informatics and Bioinformatics

Define a research proposal in which the expertise of both Bioinformatics and Medical informatics is needed. Show how this research uses e-Science/GRID and/or open-source/Science Commons/WIKI and any possible other technologies. Explain the application (i.e. clarify the need for this application or what advantages for this application are with

respect the current situation). Describe in a detailed and realistic way the contribution of the BI and MI and why collaboration is necessary. Explain the modernized aspects of the research and how much time, money, men power, etc. are needed to realize this.

Schedule

Week 1

During this week the objective of the bioinformatics module will be explained. Besides, several presentations will be given on relevant subjects for both Bioinformatics and Medical informatics.

Mo 4 January (11.00-13.00); L-021

Introduction (Prof. Dr. Antoine van Kampen)

Mo 4 January (13.00-15.00); L-021

An example bioinformatics project: high throughput sequencing (Barbera van Schaik)

Tue 5 January (10.00-12.00); L-021

The BioExpert project (Dr. Andrew Gibson)

Tue 5 January (14.00-16.00); **Location to be announced**

Scientific visualization and e-BioLab (Dr. Han Rauwerda)

Wed 6 January (13.00-15.00); L-021

e-Science for Medical Imaging (Dr. Silvia Olabarriaga)

Thu 7 January (13.00-15.00); L-021

Taverna/MyExperiment (Dr. Marco Roos)

Fri 8 January (13.00-15.00); L-023

Open Source/WIKI (Dr. Marc van Driel**)**

Week 2

Mo 11 January (10.00-11.00); L-021

Discussion progress Project 1, 2 and 3. Every group prepares a powerpoint presentation of 5 minutes to give a first explanation about MI/BI, e-Science and Open-source/Science Commons/Wiki respectively. At least indicate the importance of these topics for Medical informatics. The gained knowledge can be brought along in Project 4. Indicate what literature and other sources have been referred to.

Presentation "How to write a research proposal". One of the students will provide a presentation of 30 minutes (including 5 minutes of discussion) about this topic as a preparation to project 4. Use

http://www.meaning.ca/archives/archive/art_how_to_write_P_Wong.htm as a starting point, but also search for other (literature) resources.

Thu 14 January (10.00-12.00); L-021

Discussion progress Project 4. The 3 groups explain their idea for a research proposal on the interface of Bioinformatics and Medical informatics. What is the scientific phrasing of the question? What are the contributions of the Bioinformatics and Medical informatics? How to continue? Every group prepares a powerpoint presentation of 30 minutes.

Week 3

Thu 21 January (10.00-12.00); A1.09

Discussion progress Project 1, 2, 3 and 4. No powerpoint presentation.

Presentation "WEB 2.0". One of the students prepares a 30 minute powerpoint presentation about WEB 2.0. What is WEB 2.0? How does it relate to medical informatics and bioinformatics? Lots of material can be found on the web and in the literature.

Week 4

Mo 25 January (10.00-12.00); L-021

Final presentation Project 1 by group 1: Presentation on project 1 (Medical Informatics and Bioinformatics). Presentation duration 30 minutes, 30 minutes discussion.

Thu 28 January (10.00-12.00); L-030

Final presentation Project 2 by group 2 and Project 3 by group 3: Presentation on project 2 (e-Science/Grid) and project 3 (Open Source/Science Commons/WIKI) (presentation duration 30 minutes, 30 minutes discussion).

Fr 29 January (14.00-17.00); K01-122

Final presentation Project 4 by group 1, 2 en 3: Final presentations on project 4. Presentation duration 30 minutes, 30 minutes discussion. The presentation needs to be given in an enthusiastic way. Show clearly what the scientific question is and how collaboration between MI/BI has added value in this field. Clearly explain the role of the bioinformaticians and the medical computer scientist.

Presentations and discussion

During this module each group will give a number of presentations. In principle every student will have his/her turn as speaker or discussion leader.

As every student could have downloaded a number of basis articles from Blackboard about Bioinformatics/Medical informatics, e-(bio)science/GRID and Open-source/Science Commons/WIKI during week 1 in preparation of week 2-4, substantive discussions can take place during or after the presentations. In that perspective it is expected that every student studies each subject (bioinformatics/medical informatics, e-(bio)science/GRID and Open-source/Science Commons/Wiki) and prepares questions.

The presentations will mainly be assessed on content and less on form. Preceding every presentation one of the teachers will direct a "chairman" who will lead the discussion and questions in a good way. Therefore, the person who is giving the discussion is not the discussion leader. Just as during a real conference it is the job of the chairman to ask the speaker (that person who gives the presentation) several questions in case there are few questions coming from the audience, to stimulate questions in that way.

Report

Every student should write a report about the module in which (like in the final presentation) all aspects from this module will be brought up. The paper will give a perfect summary on the gained knowledge during this module. Both Medical informatics and Bioinformatics are being defined and specified in the paper. Also the relation between Bioinformatics and Medical informatics will be discussed.
Further all discussed aspects (e-Bioscience, Grid, Open-source, Science Commons, Wiki, etc) will be pursued and explained in which way they could be relevant for research projects on the interface of Bioinformatics an Medical informatics. It is stated what the role of the medical informaticians could be in such a project. The discussion in the report must give a personal vision on the role of the medical informatician in the future in relation to the bioinformatics.

Examination

Every group will be graded for the complete project, mainly based on the report (25%) and the presentation/defense (25%). This grade defines for 50% the grades of the students who are part of that particular group. The other 50% of the student's grade will be determined based on his/her participation during the classes.

Please send both the final presentation and the written paper before Friday **5 February 12h00 to** f.p.voorbraak@amc.uva.nl an to a.h.vankampen@amc.uva.nl.

Discussion meetings

Instruction discussion leaders

- You are responsible for the discussion: introduce structure, first list the issues the listeners wish to discuss, switch to a next subject in time (f.e. if it's clear that further discussion on the current subject make no sense anymore), etc.
- Think of your own structure during the discussion
- Carefully read the articles on the different subjects and try to find relevant articles yourself. Try to understand everything, if necessary by searching somewhere else.

(Start in time with reading!) If you do not understand everything, despite of doing your best, it is not a disaster. The knowledge you hereby gain gives an excellent base to participate in the discussion and to formulate questions.

- Make a list of discussion topics. These topics do not necessarily need to contain controversial issues; it is e.g., also possible to clarify unclear aspects in the text by discussing it together.

Instruction listeners

- Do not wish to be a passive listener.
- Carefully read the articles on the different subjects and try to find relevant articles yourself. Try to understand everything, if necessary by searching somewhere else. (Start in time with reading!) If you do not understand everything, despite of doing your best, it is not a disaster. The knowledge you hereby gain gives an excellent base to participate in the discussion and to formulate questions.
- Make a list of discussion topics yourself (if you are able to have a discussion on the most of your topics you read and thought about, and your fellow students necessarily not, there might be a chance that you look relatively smarter).

Presentation and -paper

Instructions for the setup and presentation of research proposal (Project 4)

- Inform the audience shortly after starting the presentation, how the presentation is structured, such that the audience knows what to expect roughly
- Indicate if you wish the audience to interrupt for questions/remarks, or that you prefer them to wait till the discussion.
- In any case clarify the content of your intended research proposal and why you have chosen it.
- Give sufficient background information. That way the audience understands your intended research, however make sure not to give a lecture on the subject.
- Remember that your intended research needs to be chosen in such a way that it can be handled in the 3-weeks module, however a team of advanced MIK students could handle a lot in 3 weeks.
- Describe in detail the method how aim to implement in your intended research proposal, including a workplan.
- Make sure this method is realistic (an extensive international survey is hard to realize within 3 weeks).
- Be prepared that it might be wise to slightly adjust of even drastically change your research proposal or presentation because of certain remarks of teachers or fellow students.

Instructions for other presentations

- Inform the audience shortly after starting the presentation, how the presentation is structured, in which case the audience knows what to expect roughly
- Indicate if you wish the audience to interrupt for questions/remarks, or that you prefer them to wait till the discussion.
- Roughly describe the context of your subject, in which you, for example, briefly mention interesting web-sites.
- Remember that your audience did not go deeply into the subject in a manner you have done and therefore you will have to explain it to them in simple terms.
- Not everything you tell needs to be pointed out on a slide!
- There is no need to tell everything you have learned during your research. Sketch your outlines.
- You may refer to your paper for more detailed information. However, your presentation needs to be understandable and convincing.
- During the presentation you may try to influence the discussion following the presentation by explicitly stating that some subjects could be discussion topics. However, it might be better to reflect what possibly will be discussed and already anticipate on this during the presentation (a most obvious question during the presentation could be, for example, "What is the relevancy of the subject for medical informatics and v.v.?"
- **If the content of your presentation is fine, it is alright to creatively use the audiovisual facilities of Powerpoint, however make sure this does not distract the audience (please note that a beamer will be made available, but you yourself will have to bring necessary dolby surround system if needed)**.

Instructions for your report

- The report does not need to be extensive; quality is of greater importance that quantity (a MIK-Master thesis, excluding possible enclosures, is supposed to count not more than 20.000 words, i.e., 40 pages each 500 words, and needs to describe the result of 32 working weeks. **Approximately 5000 words should be considered fine for the report about this module)**
- The report does not need to look superb; content is much more important than form. Make sure that the paper is readable through a clear and detailed presentation and reasoning.
- State your name, student number and email address on your title page.
- Insert a table of contents.
- Clearly structure your paper by dividing it in chapters, starting with an introduction and ending with a conclusion.
- Roughly describe the context of your subject in the introduction.
- Proceed from the point of view that the reader did not deeply go into the subject. Try to explain specialistic concepts shortly and detailed.
- Do not forget to mention your sources (references).
- There is no need to tell everything you have learned during your research. Sketch your outlines.