Medical Informatics in a United and Healthy Europe
K.-P. Adlassnig et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-044-5-757

European Efforts in Nanoinformatics Research Applied to Nanomedicine

Stefano CHIESA [a,1], Diana DE LA IGLESIA [a], Jose CRESPO [a], Fernando MARTIN-SANCHEZ [b], Josipa KERN [c], George POTAMIAS [d], Victor MAOJO [a]
[a] *Grupo de Informatica Biomedica, Departamento de Inteligencia Artificial, Facultad de Informatica, Universidad Politecnica de Madrid, Spain*
[b] *Department of Medical Bioinformatics, Institute of Health "Carlos III", Madrid, Spain*
[c] *Department of Medical Statistics, Epidemiology and Medical Informatics, Andrija Stampar School of Public Health, Zagreb University Medical School, Croatia*
[d] *Institute of Computer Science, FORTH, Heraklion, Crete, Greece*

Abstract. Nanomedicine and nanoinformatics are emerging disciplines with substantial challenges ahead. For instance, nanomedicine involves complex and massive data analysis. Nanoinformatics could expand previous experiences in Biomedical Informatics with new features required to study different scientific biological and physical characteristics at a different level of complexity. ACTION-Grid is a project, funded by the European Commission, which aims to the creation of a collaborative environment in biomedical and nanomedical research among countries in Europe, Western Balkans, Latin America and North Africa. In this paper, we briefly review the concepts of nanomedicine and nanoinformatics and then we describe the activities of some of the ACTION-Grid consortium members considering those initiatives related to nanomedicine.

Keywords. nanoinformatics, nanomedicine, knowledge management, text mining

1. Introduction

The impact of the nano-related disciplines on the scientific research world is increasing. The possibilities offered by the application of nanotechnology to fields like medicine surely will imply a change of trend in various aspects of these areas. This growth requires a constant effort to efficiently manage this rapid progress.

The great interest of the scientific community in nanomedicine has already generated a large amount of information – i.e., documents and tools – that is normally unstructured and not organized. To improve the efficiency of research on nanomedicine-related areas, the scientific community should approach the problem considering various points of view, such as for instance:

- **Organize the current available material**. To create indexes of current research initiatives and the available resources that can be used and shared by researchers.
- **Focus the efforts on specific high-impact research directions**. Some topics can be considered more relevant for a rapid progress of the area. Research effort should be focused on these themes.

[1] Corresponding Author: Departamento de Inteligencia Artificial, Facultad de Informática, Universidad Politecnica de Madrid, 28660 Boadilla del monte, Madrid, Spain; E-mail: schiesa@infomed.dia.fi.upm.es.

1.1. Nanomedicine and Nanoinformatics

NIH defines nanomedicine as a "*highly specific medical intervention at the molecular scale for curing disease or repairing damaged tissues*". In general terms, nanomedicine is the application of nanotechnology to the medical domain. The possibilities that this area creates are impressive and range from the improvement of pharmaceutics products, making them more effective and reducing their contraindications, to the enhancement of the diagnostic precision or tissue replacement and repairing. There are also some secondary effects such as nanotoxicity, which should be analyzed before they are effectively applied in clinical routine.

1.2. Nanoparticles and Nanotoxicity

According to the ASTM standard definition a nanoparticle is any particle with lengths measured in nanometers (10^{-9} m) and range from 1 to 100 nm in two or three dimensions [1]. Recently nanoparticles started to be widely used in medical practice as diagnostic and therapeutic tools in order to understand, detect and treat diseases. To understand the toxicity of nanoparticles it is important to consider them as particles of very tiny size enabled to circulate through the blood, lymph or other paths in human, animal or any biological organism [2]. Exposure to them could have some unwanted consequences, some adverse effect, some temporary or permanent toxic result.

There is a number of studies considering the toxicity of nanoparticles in both *"in vitro"* and *"in vivo"*. For example, Sayes and al. [3] used several types of nanoparticles for experiment with rats (*in vivo*) and selected tissue cultures (*in vitro*). However, very weak correlation between results of toxicity of particular particles in these two approaches was found. The *in vivo* approach is more complex, due to the fact that it should take care not only of dose and time dependency but also on potential multiple exposures and bioaccumulation. *In vitro* studies are more frequent, but willing the thorough understanding of kinetics and toxicology of nanoparticles, *in vivo* studies will be more appropriate. Any result identified in *in vitro* studies should be verified by *in vivo* studies. Such studies, description of their design, carrying out, and publishing results should be also standardized as is the custom in science [4, 5].

Given the complexity of interactions and secondary effects that may appear, the implication of informatics tools and systems – such as health electronic records – will be important to address information management in this area.

2. European Research in the Framework

ACTION-Grid (www.vph-action-grid.eu) is a European Commission-funded cooperative action aiming to the development of a collaborative environment among research groups in Europe, Latin America, Western Balkans and North Africa, which are working on bioinformatics, medical informatics, nanomedicine and high performance and GRID computing. In the next section we present an overview on current efforts of the ACTION-Grid Consortium members related with nanomedicine.

2.1. An Inventory of Informatics Resources

As stated above, nanomedicine implies an important information management problem that can be addressed by a novel discipline called nanoinformatics. The latter involves the development of effective tools/technologies/methods for collecting, standardizing, integrating, analyzing and visualizing information relevant to nanomedicine including those data that can be relevant such as literature, physical-chemical properties, biological, clinical and toxicological effects [6]. In such context, we developed a method for the automatic creation of an index of bioinformatics resources. This collection of tools, databases and services (in general "resources") is an index of resources, created automatically from a set of scientific documents describing resources retrieved from internet. Each of these documents is analyzed and processed to extract relevant information. The method used to process documents is based on four steps [7, 8]: i) the document is divided in sections corresponding to the title, the abstract and the article body; ii) Then the algorithm preprocesses the sentences belonging to the title and the abstract sections, eliminating stop words and reducing the remaining words to their root form; iii) the preprocessed text is given as input to a transition network that, applying morphological and syntactic patterns, extracts the name and the functionalities of the resource described in the paper given as input; iv) Finally, a second transition network uses the previous extracted information to classify the considered resource.

Tools and databases about nanotechnology applied to medicine are currently in development and in some case their classification is still impossible. Then, it is needed to define what we consider as relevant information in the nanomedicine domain.

2.2. Lab- on-Chip & Proteomics-Based Discovery of Biomarkers

The importance of lab-on-chip technology, and of the engaged proteomics based discovery of biomarkers is made evident, among other, from a recent special issue of *Nature* [9], and respective articles related to microfluidic-based diagnostic technology [10]. A biomarker is an identified protein, which is correlated with the state of a particular disease or condition [11]. Biomarkers can be used for detection, diagnosis, treatment, monitoring, prognosis and drug discovery processes. An indicative example of a relative R&D approach in the field is the work carried out in the context of the LOCCANDIA (www.loccandia.eu) project. The work concerns the validation of an application targeting plasma protein profiling for early pancreatic cancer diagnosis by means of developing an innovative nano-technology based (lab-on-a-chip) platform integrated in a full proteomics analysis chain [12]. The model combines three modules: (a) The 'bio' module of the analysis chain gets as input blood sample, and provides the 'nano' part with a selected protein mixture; (b) The 'nano' module is the part related to the dedicated Lab-on-Chip device, MS experiment and analysis of related MS-data. (c) The 'info' part is related to the supporting information technology infrastructure and is utilized through a specially designed LIMS that manages both proteomic and clinical data, i.e., an integrated *clinico-proteomics information management system* [13]. LIMS encompasses an advanced data-mining module for the analysis of MS-data and the discovery of biomarkers. Various XML (eXtensible Markup Language) standards for proteomics have been developed in order to facilitate the capture, analysis, and distribution of proteomics data, such as mzXML [14, 15]. The intent of mzXML is to encapsulate unprocessed, raw peak lists. A protein 'profile reconstruction module' for the quantification (from mass-spectrogram peaks) of targeted protein concentration

levels [16] is in the heart of the full proteomics analysis chain, realised by a dedicated module in the LOCCANDIA LIMS.

2.3. Micro and Nano-Arrays

Microarrays represent a bridging development between micro and nanotechnology. At the microarray laboratory of the Medical Bioinformatics Department of ISCIII (MBD), we have been working for several years [17] in this scientific and technological interface between micro and nano technologies. The main efforts carried out up to now have been focused on two different lines. The first one is oriented towards improving the quality of the chemistry that it is used on the surfaces of the microarrays. The group has been testing different methodologies for treating glass slides to increase their affinity for the biological molecules that are more frequently immobilised (nucleic acids and proteins). The second approach is represented by the use of biological machinery and robotics for the in-situ synthesis of biological molecules on the microarray surface. The basis for this approach relies on the immobilization of biological molecules that can act as a substrate for the synthesis. We have applied this technology for the development of diagnostic assays [18] and the detection of microbial pathogens.

2.4. Knowledge Management in Nanomedicine

Knowledge management is a key component that will facilitate further advances in the field, classifying and making easily available the different resources to the research community. Our group has been working in structuring a framework that can accommodate and organize the increasing number of available resources (systems, methods, data, software tools and standards) in nanomedicine [19]. The in-house knowledge management system (BIKMAS) that MBD has been using for several years [20] has been recently adapted to be able to incorporate and process all the new information coming from nanomedicine research and development. Using this new version of the software, it is possible to provide partners with up-to-date information related to advances of the application of nanotechnology in medicine.

2.5. Converging Technologies: Nano Bio Info Cogno (NBIC)

At the MBD, researchers are working in a model called Info-POC. Info-POC stands for Information-enhanced Point-Of-Care diagnostic system. This research aims to demonstrate the concept of convergence among NBIC technologies, with the goal of creating a new research framework for the development of diagnostic methods at the point of care, in the context of genomic medicine. The systems includes: (1) the study of new "*in info*" diagnosis and treatment methods, based on biomolecular automata. (2) The analysis of differential gene expression in colon cancer and the detection of microorganisms "*in vitro*" by mean of microarrays. They will be complemented with the biomolecular automata systems for diagnosis and with processes of information retrieval from Internet with the aim of putting into context the results of the system. (3) the design of new "*in silico*" methods, based on the Semantic Web, for enhancing private and public data sources retrieval and mining. These results, previously began in a Spanish national network called COMBIOMED, will be integrated with computerized protocols for diagnosis, treatment and management of patients with particular genetic (cancer) and infectious diseases.

3. Results and Conclusions

The ACTION-Grid project aims to achieve important objectives to foster research in all the mentioned areas and facilitate knowledge exchanges among heterogeneous research groups. Currently, preliminary results are already available, contributing to the construction of the resources index. Each ACTION-Grid consortium members is also obtaining results on their specified research lines as described in the previous sections of this article. Nanoinformatics can become a key topic for future medical research activities. The scientific community should facilitate and enhance synergies between this discipline and classical disciplines like medical informatics and bioinformatics. Lessons learned for various decades in these two disciplines can be valuable to develop a sound agenda and strong foundations for nanoinformatics.

References

[1] ASTM International (2006) *Terminology for Nanotechnology*, ASTM E 2456-06.
[2] Buzea, C., Pacheco, I.I., Robbie, K. (2007) Nanomaterials and nanoparticles: Sources and toxicity. *Biointerphases* 2(4):MR17–MR71.
[3] Sayes, C.M., Reed, K.L., Warheit, D.B. (2007) Assessing toxicity of fine and nanoparticles: Comparing in vitro measurements to in vivo pulmonary toxicity profiles. *Toxicological Sciences* 97(1):163–180.
[4] Talmon, J., Ammenwerth, E., Brender, J. et al. (2009) STARE-HI – Statement on reporting of evaluation studies in Health Informatics. *International Journal of Medical Informatics* 78(1):1–9.
[5] STROBE Statement. *STrengthening the Reporting of OBservational studies in Epidemiology*, http://www.strobe-statement.org/Checklist.html.
[6] López-Alonso, V., Hermosilla-Gimeno, I., Lopez-Campos, G. et al. (2008) Action GRID: Assessing the impact of nanotechnology on biomedical informatics. *AMIA Annual Symposium Proceedings 2008*, 1046.
[7] Chiesa, S., García-Remesal, G. et al. (2008) Building an index of nanomedical resources: An automatic approach based on text mining. In *Proceedings of the 12th International Conference on Knowledge-Based Intelligent Information and Engineering Systems*, Pt II, Springer, Berlin, 50–57.
[8] de la Calle, G., Garcia-Remesal, M., Maojo, V. (2008) A method for indexing biomedical resources over the internet. *Studies in Health Technology and Informatics* 136:163–168.
[9] Nature (2006) Lab on a chip. *Nature 442* (special issue), 367–418.
[10] Yager, P., Edwards, T., Fu, E., Helton, K. et al. (2006) Microfluidic diagnostic technologies for global public health. *Nature* 442:412–418.
[11] Morris, S.K., Coombes, K.R., Koomen, J. et al. (2005) Feature extraction and quantification for mass spectrometry in biomedical applications using the mean spectrum. *Bioinformatics* 21:1764–1775.
[12] Jordan, B. et al. (2008) LOCCANDIA: Lab-on-Chip Based Protein profiling for Cancer Diagnosis. In *Proceedings of pHealth2008*, Valencia, Spain, http://www.phealth2008.com/Events/papers/p18.pdf.
[13] Kalaitzakis, M., Kritsotakis, V., Kondylakis, H. et al. (2008) An Integrated Clinico-Proteomics Information Management and Analysis Platform. In *Proceedings of CBMS 2008*, IEEE, 218–220.
[14] Pedrioliet, N.G. et al. (2004) A common open representation of mass spectrometry data and its application to proteomics research. *Nature Biotechnology* 22(11):1459–1466.
[15] Lin, S.M. et al. (2005) What is mzXML good for? *Expert Review of Proteomics* 2(6):839–845.
[16] Paulus, C. et al. (2007) Chromatographic alignment combined with chemometrics profile reconstruction approaches applied to LC-MS data. In *Proceedings of the IEEE EMBS 2007*, 5983–5986.
[17] Lopez-Campos, G., Coiras, M., Sanchez Merino, J.P. et al. (2007) Oligonucleotide microarray design for detection and serotyping of human respiratory adenoviruses by using a virtual amplicon retrieval software. *Journal of Virological Methods* 145(2):127–136.
[18] Lopez-Campos, G.H., Garcia-Albert, L. et al. (2006) Analysis and management of HIV peptide microarray experiments. *Methods of Information in Medicine* 45(2):158–162.
[19] Martín-Sanchez, F., López-Alonso, V., Hermosilla-Gimeno, I. et al. (2008) A primer in knowledge management for nanoinformatics in medicine. In *Proceedings of the 12th International Conference on Knowledge-Based Intelligent Information and Engineering Systems*, Pt II, Springer, Berlin, 66–72.
[20] López Alonso, V., Moreno López, L. et al. (2002) Description of an Ontology for supporting BIKMAS, a Biomedical Informatics Knowledge Management System. In *Proceedings of the Annual Congress of Knowledge-Based Intelligent Information and Engineering Systems KES 2002*, IOS Press, 955–959.