

Nanoinformatics knowledge infrastructures: bringing efficient information management to nanomedical research

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2013 Comput. Sci. Disc. 6 014011

(<http://iopscience.iop.org/1749-4699/6/1/014011>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

This content was downloaded by: ddelaiglesia

IP Address: 138.100.11.168

This content was downloaded on 28/11/2013 at 09:40

Please note that [terms and conditions apply](#).

Nanoinformatics knowledge infrastructures: bringing efficient information management to nanomedical research

D de la Iglesia¹, R E Cachau², M García-Remesal¹ and V Maojo¹

¹ Biomedical Informatics Group, Departamento de Inteligencia Artificial, Facultad de Informatica, Universidad Politecnica de Madrid, E-28660 Boadilla del Monte, Madrid, Spain

² Advanced Biomedical Computing Center, National Cancer Institute, SAIC-Frederick Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA

E-mail: diglesia@infomed.dia.fi.upm.es

Received 15 May 2013, in final form 31 October 2013

Published 27 November 2013

Computational Science & Discovery **6** (2013) 014011 (27pp)

[doi:10.1088/1749-4699/6/1/014011](https://doi.org/10.1088/1749-4699/6/1/014011)

Abstract. Nanotechnology represents an area of particular promise and significant opportunity across multiple scientific disciplines. Ongoing nanotechnology research ranges from the characterization of nanoparticles and nanomaterials to the analysis and processing of experimental data seeking correlations between nanoparticles and their functionalities and side effects. Due to their special properties, nanoparticles are suitable for cellular-level diagnostics and therapy, offering numerous applications in medicine, e.g. development of biomedical devices, tissue repair, drug delivery systems and biosensors. In nanomedicine, recent studies are producing large amounts of structural and property data, highlighting the role for computational approaches in information management. While *in vitro* and *in vivo* assays are expensive, the cost of computing is falling. Furthermore, improvements in the accuracy of computational methods (e.g. data mining, knowledge discovery, modeling and simulation) have enabled effective tools to automate the extraction, management and storage of these vast data volumes. Since this information is widely distributed, one major issue is how to locate and access data where it resides (which also poses data-sharing limitations). The novel discipline of nanoinformatics addresses the information challenges related to nanotechnology research. In this paper, we summarize the needs and challenges in the field and present an overview of extant initiatives and efforts.

Contents

1. Introduction	2
1.1. Applications of nanoparticles in medicine	3
2. Information in nanomedicine	4
2.1. From data to knowledge	5
2.2. Information standards	6
2.3. Information sources	6
3. Available resources for nanomedical researchers	9
3.1. Data sources	10
3.2. Standards, terminologies, taxonomies and ontologies	13
3.3. Modeling and simulation tools	15
3.4. Educational tools	16
3.5. Working groups on nanoinformatics	17
4. Cyberinfrastructures for nanomedical data management: current limitations and future prospects	17
4.1. Previous successes in the biomedical domain	18
4.2. Requirements for developing a nanoinformatics infrastructure	20
5. Conclusions	23
Acknowledgments	23
References	24

1. Introduction

Over the last 20 years, nanotechnology has pushed back the frontiers of science and has especially revolutionized the field of medicine (Duncan and Gaspar 2011). This novel scientific field deals with materials at the nano-dimension (macromolecules that range in size from 1 to 100 nm), offering breakthroughs in a wide range of scientific areas, such as physics, chemistry, microbiology and materials science, but also posing many challenges (Lemley 2005).

At the nanoscale, physical, chemical and functional properties differ from those at the cell, tissue or organ level (McNeil 2011). Unique characteristics emerge from nanoparticles, based on their size—they are small enough to interact with receptors with high specificity and effectiveness and large enough to carry drugs at the molecular level—and also on their quantum properties: density, stability or unexpected optical properties such as high absorption of solar radiation (Scholes and Rumbles 2006) or superparamagnetism (Sato *et al* 2011), for instance. All this represents a dramatic turnaround from a few years ago for the scientific community. In medicine, worldwide efforts to foster these novel technologies promises to bring us new insights into the development of biomarkers and personalized treatments, but also raises concerns about inefficacy in the management and analysis of all of this new information to address potential hazards and risks. In this sense, society's perception of nanotechnology requires strong consideration since it could influence industry oversight of future advancements in the area.

One current issue regarding nanoparticles is potential toxicity, a field that is still obscure to researchers. It is unclear how and what toxicity some nanoparticles have and there is a lack of extensive regulation for new products derived from them. What may be harmless at the macro or chemical level may have different effects in the body than the nanoparticle form. Possibly the largest time- and money-related costs involved in developing nanomaterials may not be in making them work, but in determining their side effects and biological impacts.

Computational methods and tools offer novel paradigms to materials discovery and design, improving properties prediction or material selection capabilities that could increase product performance and reduce time from drug discovery to marketing. The novel discipline of nanoinformatics deals with all of these information challenges related to nanotechnology research and, specifically, with nanomedicine (Maojo *et al* 2012a). Several nanoinformatics initiatives are currently identifying data gaps and research priorities in the area and

developing collaborative nanoinformatics environments where research institutions can publish descriptions of their data sources and make them publicly available over the Internet.

In this paper, we present an overview of the needs and challenges in nanoinformatics, as well as the extant initiatives and international efforts in the field. All these initiatives are fostering scientific advancement and knowledge dissemination in the nanomedical area and closing the gaps between basic research and clinical applications (de la Iglesia *et al* 2011).

1.1. Applications of nanoparticles in medicine

Groundbreaking advances in nanoscience continue to provide incremental insights into medicine. Among numerous potential applications, such as biosensors, biological carriers for drug delivery or structures for tissue repair, we illustrate in this section some examples of the more innovative research in the field of nanomedicine. We can differentiate two contexts in nanotechnology applications: diagnostics (based on molecular biosensing) and therapeutics, which follow tailored strategies.

As biomarkers and imaging contrast agents, nanoparticles can be decorated with specific diagnostic agents targeted to specific receptors (Mailänder and Landfester 2009), which allows monitoring of their distribution (Haun *et al* 2011). This property has many diagnostic biosensing applications (Yeri and Gao 2011). The large surface area of nanomaterials, chemically functionalized with targeting ligands, provides excellent levels of sensitivity for the detection of molecular targets, such as DNA, proteins, pathogens, tumor cells or enzymes (Kumbar *et al* 2008). For instance, scientists at the Northwestern University Feinberg School of Medicine have developed a bio-barcode assay (Thaxton *et al* 2009) for detecting prostate-specific antigen based on the use of gold nanoparticles and have recently defined nanocombinatorics (Giam *et al* 2012), a new lithography technique that generates nanoscale biological patterns. In the NanoSystems Biology Cancer Center at the California Institute of Technology, researchers have developed a nano-based, integrated platform with two different applications: cell sorting and multiplex detection of DNA and proteins (Bailey *et al* 2007). Photoacoustic molecular imaging is another potential application of nanoparticles as molecular markers. It consists of marking biological targets with acoustic detectors, such as carbon nanotubes and gold nanoparticles. This promising technique has been successfully applied in mice by the Molecular Imaging Program at Stanford for multimodal brain tumor imaging (Kircher *et al* 2012) and for the diagnosis of retinal diseases (de la Zerda *et al* 2012).

As delivery systems, nanoparticles are suitable for systematically releasing drugs to a specific target (cell or tissue), increasing their efficiency and improving the characterization of the treatment response. Nanocarriers can be seen as sophisticated mechanisms by which nano-sized objects can be transported through the body to target specific cells and penetrate cellular and nuclear membranes. Advanced design of nanomaterials has demonstrated that engineered nanoparticles can evade immune detection, facilitating the targeted delivery of a specific drug. Nanoparticles hide their payloads, protecting drugs from several defense mechanisms and biological barriers, such as particle disposal by phagocytic cells or renal plasma clearance (Zuckerman *et al* 2012). Nanoparticles can also be used for gene delivery (Jin *et al* 2009), leading to deactivation of genes related to a specific illness (e.g. cancer). The nanoparticle preserves its genetic payload from degrading enzymes and the immune system and carries it into the target cells, as well. In addition, advances in nanotechnology are being applied to *in vitro* regenerative medicine: tissue regeneration (Veetil and Ye 2009), cardiovascular implants (Ghanbari *et al* 2011, de Mel *et al* 2012), cell therapy (Perán *et al* 2012), wound healing and delivery of exogenous growth factors (Koria 2012) are examples.

With regards to cancer research—one of the main areas where nanotechnology discoveries are being applied—biological mechanisms are complex and drug interactions offer very different results between *in vitro* and *in vivo* assays. However, with the use of nanoparticles, diagnostics and therapeutics could be greatly improved for certain types of tumors, such as those tumors currently treated with toxic drugs—i.e. lung (Upadhyay *et al* 2012) or breast cancer (Xin *et al* 2012)—or gene therapy—ovarian (Fathabadi *et al* 2012) and brain cancer (Liu *et al* 2012). For instance, researchers at the University of Melbourne (Australia) are using certain engineered nanostructures (polymers, micelles and polymersomes) as carriers to deliver drugs that target cancer cells causing minimal biological side effects (Kamphuis *et al* 2010, Shimoni *et al* 2012). With the same objective, the continuous efforts of the National Cancer Institute (NCI) to apply nanotechnology-based

advances to early diagnosis and effective therapy against cancer (Service 2005, 2010, Ptak *et al* 2010, Shukla *et al* 2011) is also worthy of mention. There are other remarkable examples of innovative nanomedical research. At the University of North Carolina, researchers are conducting experiments that use the novel PRINT technique (Perry *et al* 2011) for nanoparticle replication and mimicry (Parrott *et al* 2012, Wang *et al* 2012), obtaining new particles that can be tuned with controlled properties, e.g. elasticity, size or shape. Such advances could be applied to the development of nanoparticles for use in the emerging area of theranostics (Kelkar and Reineke 2011, Perry *et al* 2011, Kiessling *et al* 2012), a new paradigm that combines diagnostics and therapy for individual patients in one step.

However, there is an urgent need to shorten the timeline between the acquisition of fundamental knowledge about nanomaterials, the publication of related studies and assays in scientific media and their clinical application. Social concerns about nanoparticles toxicity or nanotoxicity (the potential negative health impact of exposure to nanomaterials) could be a barrier for the agile and efficient transfer of nanotechnology from the bench to the bedside (Bellare 2011). This could be addressed by supporting responsible nanotechnology development, providing high-quality information through a better exchange system and sharing available toxicity and physico-chemical data to efficiently determine the association between exposure to nanomaterials and side effects. In both the diagnostic and treatment contexts, pharmacokinetic behavior and potential risks pose by nanoparticle-based drugs should be studied in-depth, considering both genotypic and phenotypic studies. Currently, in most cases, early clinical trials on nanoparticles are not revealing new toxic effects from the nanoparticle itself or its components. Therefore, it would be necessary to differentiate between the toxicity of the container and that of its payload (i.e. surface versus core), which are cornerstones in determining and monitoring the efficacy and adverse health effects of new nanomaterials. Indeed, in the case of nanomedicines, the development and approval of new drugs does not need to take so long and clinical trials are not necessarily vital in some cases, particularly when considering that nanomaterials are generally derived from drugs previously authorized by the US Food and Drug Administration (FDA) (Etheridge *et al* 2013) or the European Medicines Agency (Karalis and Macheras 2012). There are many drugs already approved or under commercial use, such as those for therapy against cancer—Daunoxome[®] (Petre and Dittmer 2007), Doxil[®] (Pearce *et al* 2012) and Caelyx[®] (Li *et al* 2012), among others—, bone substitutes such as CarriGen^{®3}, dental composites such as FilTek[®] (Moosavi *et al* 2012), or platforms as the Verigene[®] System (Buxton 2009), which is extensively used in hospitals for diagnostic purposes. Considering the myriad of new and increasingly rapid results generated by nanomedical research nowadays, it is obvious that the development of new tools for managing all this information has become crucial for advancing the field.

2. Information in nanomedicine

Information delivery channels are rapidly evolving, generating an increasing need to manage available data and resources. Specifically, discovery and design of new materials (Bikiaris 2007, Huh and Kwon 2011, Majzoub and Rönnebro 2012, Mentovich *et al* 2012) and recent advances in personalized medicine (Jain 2011) raise new challenges in the analysis and integration of the massive heterogeneous data generated by nanomedical research.

Unfortunately, researchers are facing many difficulties in their daily tasks related to information retrieval and management in nanomedicine. The number of sources is huge, and is increasing every day, but the available information is not always what is required while, in other cases, the data are not comprehensive or the quality is not sufficiently high. In addition, professionals currently working on nanotechnology have different roles and professional profiles, such as researchers, program managers, policy makers, infrastructure providers, clinicians, students and workers. Thus, the type of information to be managed is very complete and of broad scope, but also heterogeneous with regards to format and content. Information about researchers, research programs, tools, nanomaterial data and current projects is distributed among numerous sources, leading to a considerable disconnect between resources and users currently performing clinical, environmental, medical and basic research. Thus, how to manage nanoinformation is still the open question for computer scientists.

³ CarriGen[®], porous bone substitute material composed of a proprietary nanocrystalline calcium phosphate formulation that mimics the crystalline mineral structure of human bone. ETEX products. <https://secure.etexcorp.com/products/overview.html>.

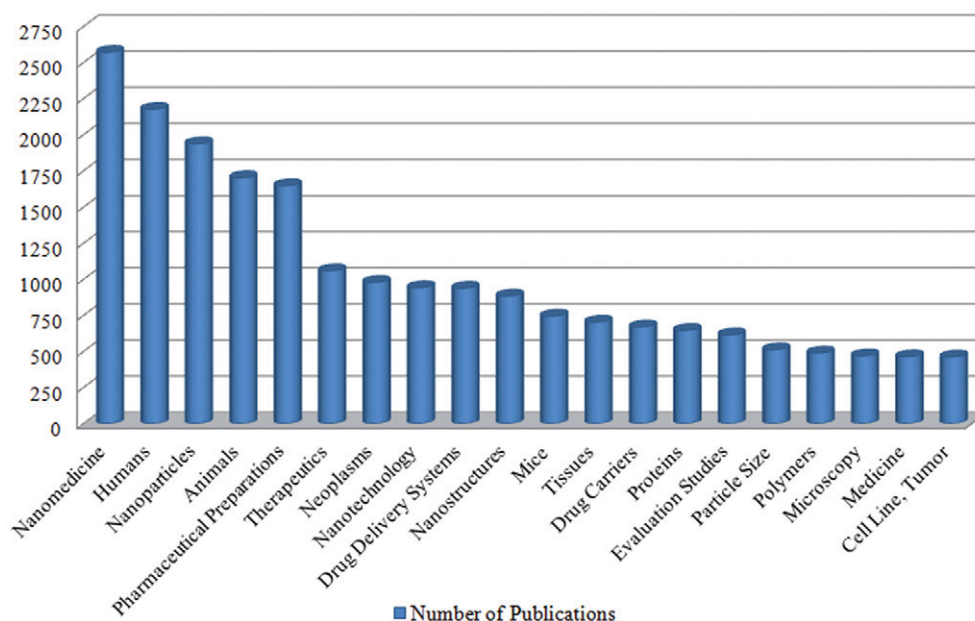


Figure 1. Top terms that appear in research publications related to nanomedicine.

Figure 1 shows the top terms appearing in nanomedical research publications, demonstrating the variety of concepts (from tissues to cells to particles) embraced by nanomedicine.

Over the last year, there has been a proliferation of computational resources related to the area of nanotechnology and, specifically, nanomedicine. As proposed elsewhere (Baker *et al* 2009, Maojo *et al* 2010, 2012b), nanoinformatics deals with how to extract useful information from all available data relevant to nanotechnology. Because this novel discipline is in its early stages, research on nanoinformatics currently faces many challenges: (i) the definition of semantics in the field; (ii) the standardization and integration of data; (iii) the interoperability of resources; (iv) the lack of consistent annotations; and (iv) the necessary integration of data analysis tools with databases. Building cooperation among these areas may facilitate the development of useful and effective nanoinformatics methods and tools.

2.1. From data to knowledge

The inference of knowledge from raw data sets involves several steps: (i) data collection; (ii) data analysis and annotation; (iii) information extraction; and, finally; and (iv) data validation by experts in the field. With regards to data collection, we currently face serious restrictions for accessing nanomaterial data, such as the unwillingness to share experiments and results or the existence of proprietary data that is not publicly available. In addition, some of the experimental results reported in the scientific literature contain numerical errors or are not properly described—in terms of protocols, experiment parameters or techniques—, so it is not possible to reproduce the experiment. Therefore, in order to extract high-quality information from raw data sets, a huge effort is needed to validate and curate existing repositories, deleting invalid and redundant data.

Knowledge inference also involves classifying and ranking information into categories and requires the definition of a specific semantic domain, based on standard terminologies and vocabularies. Specifically, knowledge discovery in nanomedicine should be driven by two means: (i) biological and clinical problems and (ii) research data and models. Biological and clinical problems raise the target questions and challenges to be addressed: what tasks need to be supported and which resources—data and services—are needed. With regards to the second point, we find data modeling to be an issue: semantics requires a solid context to be defined in order to achieve data consistency between research results and predictive models.

Informatics tools aim to define this context by providing different techniques and methodologies—models, taxonomies and ontologies—to ensure data quality and harmonization and to interpret the meaning of data, turning it into usable information and knowledge. Additionally, computational methods to store, organize

and manage the huge volumes of digital data generated during research on nanomedicine could be used to: (i) integrate structural and functional characterization research data with multiple types of ‘nano’ information (with clinical trial data, for instance); and (ii) find, through the development of data analysis methods, potential associations between the size, structure, composition and physical and chemical properties of nanoparticles, as well as the extant correlations between *in vitro* and *in vivo* assays.

2.2. Information standards

Information standards are the basis of any scientific discipline. Using a common formalism to represent the elements within a domain guarantees the creation of dynamic, interoperable information systems that share data models linked to standard terminologies. Due to the wide range and heterogeneity of fields contributing to nanotechnology, global consensus in semantics—nomenclatures, code sets, descriptions, classifications and specifications of nanomaterials and related procedures, for example—would facilitate the interoperability of data, ensuring data quality, consistency and normalization. Examples of major formal standards in nanotechnology, developed and approved by the main international standards bodies and organizations (Murashov and Howard 2008, 2011, Hatto 2009, Weidl *et al* 2010) are listed below, in table 1.

There are many other worldwide initiatives actively involved in the study and development of nanotechnology standards, such as the International Alliance for NanoEHS Harmonization⁴, the Nanosciences African Network⁵, the Asia Nano Forum⁶, the Bureau International des Poids et Mesures⁷, the International Union of Pure and Applied Chemistry⁸, the Organisation for Economic Co-operation and Development⁹ and the Versailles Project on Advanced Materials and Standards¹⁰, among many others.

2.3. Information sources

The variety of nanomedical data sources constitutes another main issue for information management in nanomedicine. Usually, these sources are disparate and heterogeneous, providing different types of data in different formats. These sources are also physically separate, forcing researchers to spend research time locating them. In addition, information sources are of different natures: structured (e.g. databases and knowledgebases) and unstructured (wikis, technical reports and scientific publications are examples). Structured sources, where data are well-defined and semantics are implicit, could be easily mined with automatic methods that are already available for the scientific community: standard queries to databases, for instance. On the other hand, unstructured sources provide data in free text format, making it difficult to automate the process of information extraction. Results from nanomedical research are usually reported in scientific publications and technical reports, resulting in large volumes of unstructured text. As can be seen in figure 2, the number of publications in the field has grown considerably in the last two decades. This led to the urgent need to develop advanced natural language processing (NLP) techniques and text mining methods to convert these unstructured sources to machine-interpretable format. This issue has been partially addressed in the biomedical domain through biomedical informatics approaches, and should be addressed in the same way in the nanomedical context, by developing and applying specific nanoinformatics methods and techniques (de la Iglesia *et al* 2009).

Although there are several nanomedical data sources, it is necessary to analyze the availability and quality of such data and properly handle the existing inconsistencies and uncertainties. For instance, protocol details are inadequately described in research publications and original datasets are rarely provided in manuscripts or deposited online. Standard methods for data validation should be developed, covering the different aspects of research conducted in nanomedicine, such as material conditions or protocol parameters. In the last 5 years,

⁴ www.nanoehsalliance.org/.

⁵ www.wcpsd.org/posters/environment/Maaza.pdf.

⁶ <http://asia-anf.org/>.

⁷ www.bipm.org/en/home/.

⁸ www.iupac.org/.

⁹ www.oecd.org.

¹⁰ www.vamas.org/.

Table 1. Summary of the main current nanotechnology standards.

	Terminology/ nomenclature	Measurement/ characterization	Materials specifications	EHS
ASTM—nanotechnology standards ^a				
E2490–09		X		
E2524–08		X		
E2525–08		X		
E2526–08		X		
E2578–07(2012)		X		
E2834–12		X		
E2859–11		X		
E2535–07				X
E2456–06	X			
British Standards Institution—nanotechnology British standards ^b				
PAS 71	X			
PAS 130	X			
PAS 131	X			
PAS 132	X			
PAS 133	X			
PAS 134	X			
PAS 135	X			
PAS 136	X			
PD 6699-1			X	
PD 6699-2				X
PD 6699-3				X
European Committee for Standardization (CEN)—nanotechnologies CEN/TC 352 ^c				
CEN ISO/TS 27687:2009	X			
EN ISO 10801:2010		X		
EN ISO 10808:2010		X		
EN ISO 29701:2010		X		
Federal Agency for Technical Regulations and Metrology (Russia) ^d				
GOST R 8.628-2007		X		
GOST R 8.629-2007		X		
GOST R 8.630-2007		X		
GOST R 8.631-2007		X		
GOST R 8.635-2007		X		
GOST R 8.636-2007		X		
GOST R 8.644-2008		X		
Institute of Electrical and Electronics Engineers—nanotechnology standards ^e				
P1650 TM -2005		X		
P1690 TM		X		

Table 1. (Continued.)

	Terminology/ nomenclature	Measurement/ characterization	Materials specifications	EHS
International Electrotechnical Commission—nanotechnology standardization for electrical and electronic products and systems ^f				
IEC/TC 113	X	X	X	
ISO—International Organization for Standardization ^g				
ISO/TC 229	X	X	X	X
ISO/TC 24/SC4		X		
ISO/TC 201		X		
ISO/TC 202		X		
TS Z 0027:2010	X			
Occupational Safety and Health Administration ^h				
1904				X
1910.132				X
1910.133				X
1910.134				X
1910.138				X
1910.141				X
1910.1200				X
1910.1450				X
United Working Group for Nanomaterials Standardization (China) ⁱ				
GB/T 19619-2004	X			
GB/T 13221-2004		X		
GB/T 19587-2004		X		
GB/T 19588-2004		X		
GB/T 19589-2004		X		
GB/T 19590-2004		X		
GB/T 19591-2004		X		

^a www.astm.org/Standards/nanotechnology-standards.html#E56.01.^b <http://shop.bsigroup.com/en/Browse-by-Subject/Nanotechnology/>.^c www.cen.eu/.^d <http://rostandart.com/>.^e <http://grouper.ieee.org/groups/nano/>.^f www.iec.ch/.^g www.iso.org/iso/iso_catalogue/catalogue_tc.htm.^h www.osha.gov/dsg/nanotechnology/nanotech_standards.html.ⁱ www.anystandards.com/.

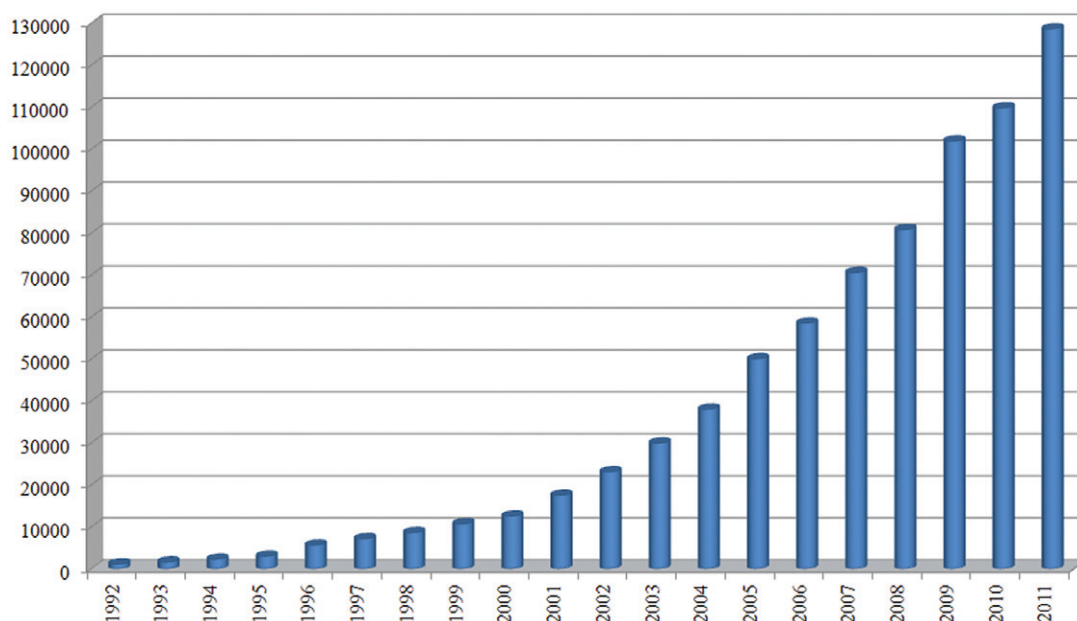


Figure 2. Number of publications with nano-prefixed terms in the title published over the last 20 years (1992–2011), following previous studies (Braun *et al* 1997, Huang *et al* 2003, Hullman and Meyer 2003, Lee *et al* 2006). Source: Web of Knowledge (Thomson Reuters)¹¹.

the number of web applications offering access to nanomaterial data sources has grown rapidly, so it has become necessary to evaluate and compare the specific features of such sources: (i) the scope of the provided data, (ii) the level of annotation of nanomaterials with relevant information, (iii) the automatic generation of catalogues from research results and (iv) the use of taxonomies and ontologies to classify the available data on nanomaterials. Another important feature of nanomaterial inventories to be evaluated, considering its use by the scientific community, is the search capabilities that it can offer. Searching for nanomaterials by name in a list sorted alphabetically is an arduous task that does not always generate the desired results. Additionally, researchers may not always know the name of the nanomaterial they are looking for (only some properties or the name of manufacturer) and, even when they do know it, it is a good idea to show them other nanomaterials that have similar characteristics. Using ontologies can assist users who are browsing the entire collection by organizing the repositories with a hierarchical category classification.

Another relevant characteristic of nanomedical data is its interdisciplinary nature. Data from chemistry, physics, medicine and other involved fields could be very different in terms of format, measurement units and vocabulary. Additionally, it generally needs to be integrated with multilevel biological data, ranging from the population level to the atomic level. The development of comprehensive data repositories linking and integrating data from diverse domains and different sources—experimental results, scientific literature, clinical trials, Electronic Health Records, genomics, proteomics, nanomaterials characterization, imaging and disease information—would lead to the design of more accurate predictive models and simulations of nanomaterials and their biological effects.

3. Available resources for nanomedical researchers

Several initiatives are currently supporting nanoinformatics research worldwide. For instance, since 2001, under the National Nanotechnology Initiative (NNI) (Roco 2003), 25 governmental agencies are sustaining nanotechnology advancements in the United States. Previous initiatives also connect to activities under the NNI, such as the Global Change Research Program¹² and the Networking and Information Technology

¹¹ <http://wokinfo.com/>.

¹² www.globalchange.gov/.

Table 2. Comparison of some data resources on nanotechnology and type of data provided.

Resource	Database	Nanomaterials characterization	Protocols	Assays	Technical reports	Scientific publications	Toxicity	Modeling tools	Interactions
Nanomaterials Registry	X	X				X	X	X	X
NIL		X				X			
ICON Nano-EHS database	X	X				X	X		
NBI Knowledgebase	X								X
CSN	X	X						X	
caNanoLab		X	X	X		X			
TOXNET						X			
NanoExPERT	X	X				X	X		
MICAD	X					X			
NCL		X	X		X	X			

Research and Development¹³, as well as other new initiatives, including the Materials Genome Initiative (National Science and Technology Council 2011), a material innovation infrastructure launched in 2011, and the Advanced Manufacturing Partnerships. In Europe, diverse projects aim at raising the levels of nanotechnology knowledge across the European Union Member States by supporting the analysis and dissemination of results from nanotechnology research, as we will see later in this paper.

The set of available nanoinformatics resources comprises databases, tools and web services useful for storing, organizing, managing and analyzing the plethora of data generated during nanotechnology research. In the next sections, we will analyze these resources, providing a comprehensive picture of the ongoing research efforts and advances in the field of nanoinformatics.

3.1. Data sources

Nowadays, there are several resources providing free access to nanomedical data. These repositories contain large amounts of heterogeneous information related to nanomedicine, including data about samples and experiments, specific studies and related protocols, nanoparticle interactions, exposure pathways, delivery targets, manufacturers, etc. Additionally, many research institutes and groups have developed elaborate nanoparticle classifications based on different parameters: nanomaterial application domain, size, function, target and toxicity, among others. The variety of ways used to categorize nanomaterials is the main barrier to standardizing common classifications. Therefore, the creation of a standard taxonomy for the classification of nanomaterials is a key element to be considered. The use of standardized data across systems and services would facilitate the updating and integration of data from many different sources. Below, we provide some examples of the interesting initiatives that are providing access to sources of information on nanomedicine and nanomaterials, as well as a comparison of their specific features (table 2).

The Nanomaterial Registry (Ostraat *et al* 2012) gives access to curated information on the interactions of nanomaterials with biological and environmental systems, including links to related publications, modeling tools, computational results and manufacturing guidance. This initiative, created and maintained by Research Triangle Institute International, is sponsored by the National Institute of Biomedical Imaging and Bioengineering, the National Institute of Environmental Health Sciences and the NCI. Users can search this database for information about a specific nanomaterial or compare the properties of multiple nanomaterials. Users can also browse by nanomaterial type (metals, fullerenes or dendrimers, e.g.), size, shape (one, two or three dimensions) and surface area. Each record contains information about the physicochemical and structural characteristics of the nanomaterial, including: size (particle size and distribution), aggregation or

¹³ www.nitrd.gov/.

agglomeration state, surface (area, charge, chemistry and reactivity), shape, composition, purity, solubility and stability. Information about the environmental and biological interactions of the nanomaterial is also provided. The Nanomaterial Registry is an ongoing project; its first beta version was released to the public at the end of June 2012¹⁴ and is updated as new information about nanomaterials is published. This resource will also offer, in future releases, tools for predictive modeling of nanomaterials. Another relevant data source is the Nanoparticle Information Library¹⁵, an ongoing project for building a web-based nanoparticle repository, developed by the National Institute for Occupational Safety and Health and the Oregon Nanoscience and Microtechnologies Institute through the Safer Nanomaterials and Nanomanufacturing Initiative¹⁶. A user can browse nanotechnology reports from a set of 88 nanoparticle records (at the time of writing). Each record contains the nanoparticle's material data, such as core element, structure or production method, the nanoparticle's image, an abstract describing the chemical and physical properties of the nanoparticle and information about pertinent scientific publications and other related data, such as its commercial availability or the sponsoring company or institution. These records also provide links to websites with extended information about the described nanoparticles. With the same goal, the International Council on Nanotechnology is developing the ICON Environmental, Health and Safety (EHS) database¹⁷. The ICON Nano-EHS database is a repository of papers related to nanoparticles. At the time of this writing, it contains over 7000 records. In this repository, the user can perform searches based on a particular characteristic of the nanoparticle—particle type, exposure pathway, method of study, exposure or hazard target, risk exposure group and production method—or based on related publications, such as paper type, target audience and content emphasis. This initiative also provides an analysis tool offering two types of analyses: (i) a simple distribution analysis (pie chart), which compares categories within a specified time range and, (ii) a time progressive distribution analysis (histogram), which compares categories over a specified time range and data grouping period. Another valuable resource is the Nanomaterial-Biological Interactions (NBI) Knowledgebase¹⁸, an initiative funded by Oregon Nanoscience and Microtechnologies Institute (ONAMI) and several US agencies: the National Science Foundation (NSF), the National Institutes of Health (NIH), the United States Environmental Protection Agency and the Air Force Research Laboratory. The NBI Knowledgebase is a web-accessible electronic library that allows users access to precise and standardized data element definitions. It serves as a repository for annotated data on nanomaterial characterization (e.g. purity, electronic and photonic properties, size, shape, charge, composition, chemical functionalization and agglomeration state), synthesis methods and nanomaterial–biological interactions (beneficial, benign or deleterious) defined at multiple levels of biological organization (e.g. molecular, cellular or organismal), thus providing the framework upon which to conduct species, route, dose and scenario extrapolations and to identify key data required to predict the biological interactions of nanomaterials. Computational and data mining tools are currently being developed and incorporated into the NBI Knowledgebase to facilitate this work.

The comparative analysis of structural complexities in nanoparticles and nanocompounds is a fundamental factor in understanding the particular properties and behaviors of nanomaterials. Regulatory approval of new engineered materials should be obtained only after a rigorous review of data from structural and functional analyses. Thus, the development of accurate models and simulations would be an essential step for the translation of results from the bench to the bedside, reducing the number of necessary phases required in clinical trials and, subsequently, the time-to-market of these manufactured nanomaterials. Following this structural approach, the Collaboratory for Structural Nanobiology (CSN)¹⁹ initiative, coordinated by one of the co-authors, provides a database of chemical and structural models of nanoparticles such as dendrimers, buckyballs, nanotubes and metallic particles. Each nanoparticle is associated with its correspondent structure file, similar to those offered by the Protein Data Bank (PDB), and related research data, such as molecular weight, simplified molecular input line entry specification or studies on van der Waals forces. The platform

¹⁴ www.nanomaterialregistry.org.

¹⁵ <http://nanoparticlelibrary.net/index.asp>.

¹⁶ www.greennano.org.

¹⁷ <http://icon.rice.edu/report.cfm>.

¹⁸ <http://nbi.oregonstate.edu/>.

¹⁹ <http://nanobiology.ncifcrf.gov>.

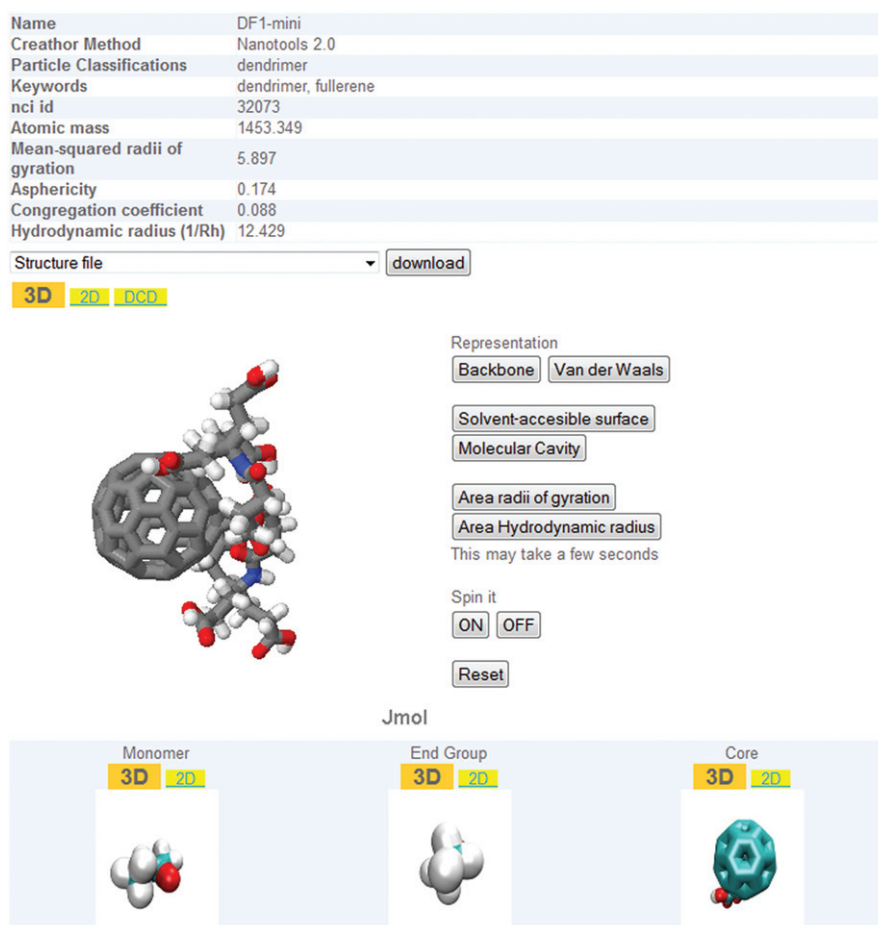


Figure 3. Example of nanoparticle record in the CSN: a 3D representation of a fullerene (DF1-mini).

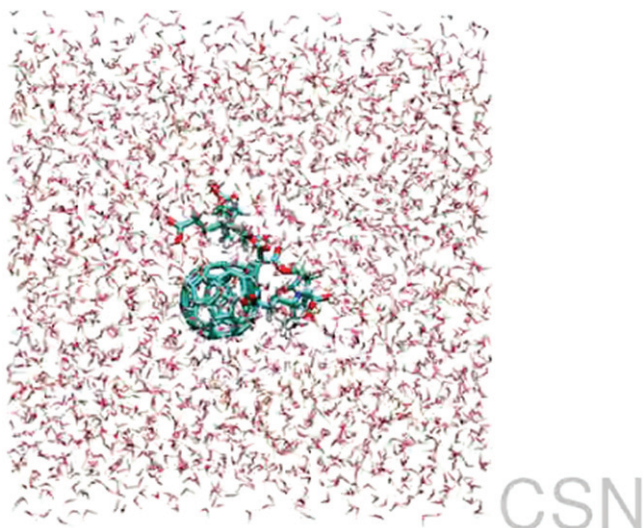


Figure 4. Sample panel of a video simulating the water molecular dynamics of DF1-mini in the CSN.

also provides several tools for the two-dimensional and three-dimensional (3D) visualization of nanoparticles (figure 3), as well as representative videos of dynamic simulations of such nanoparticles (figure 4). This kind of approach could lead to future platforms upon which to share structural and predictive models.

Another interesting data source for nanomedicine is caNanoLab²⁰, a portal for sharing information about nanomaterial characterizations resulting from cancer-related nanotechnology research, funded by the NCI. It supports the annotation of nanomaterials with data from *in vitro* assays and allows linking to remote characterization protocols. At the time of writing, the caNanoLab Portal (v1.5.3) provides access to 41 protocols and 886 samples from different sources, supplying information about the sample, such as composition, function and characterizations and providing access to more than 1000 publications related to the stored samples. In addition, the portal supplies the data submission date in the search results. It is an open-source initiative, so users are able to download the source code²¹ and try it in their own systems.

Also accessible through the web, the Toxicology Data Network (TOXNET)²² provides access to toxicology data for about 5000 chemicals, including nanomaterials. TOXNET comprises several databases from the US National Library of Medicine and some of them—the Hazardous Substances Data Bank (HSDB) and TOXLINE—allow general and specific searches about nanoparticles and nanostructures. Records can be queried by substance name or Chemical Abstract Service number. The HSDB provides chemical records about nanotoxicity by means of descriptions of physicochemical properties, toxicity studies, pharmacology, manufacturing, environmental and health safety, pharmacodynamics, etc. On the other hand, TOXLINE provides bibliographic records of scientific papers related to nanomaterials' toxicity. All of this information is presented to the user in free text, but can also be integrated with other applications via its web services Application Program Interface. NanoExPERT²³, the Nanomaterials Experiment-based Predictor of Environmental Risk and Toxicity, is another tool that provides fate and toxicity predictions for nanoparticles, as well as a database containing experimental results from peer-reviewed journals. With respect to images databases, the Molecular Imaging and Contrast Agent Database²⁴, developed by the National Center for Biotechnology Information (NCBI) at the NIH, stores imaging data about a range of nanomaterials, such as nanobubbles, quantum dots and gold nanoparticles.

Finally, it is worth mentioning the Nanotechnology Characterization Laboratory (NCL)²⁵ for conducting preclinical studies to characterize nanomaterials that are used for *in vivo* diagnostics and therapy against cancer. This outstanding nanotechnology institution, developed through a collaboration between the NCI, the FDA and the National Institute of Standards and Technology, provides access to a set of high-quality technical reports regarding liposomes, fullerenes and dendrimers. There are also other available databases dedicated to collecting transversal data about nanotechnology, such as the database of nanoscience journal publications developed by Braun (Braun *et al* 1997) and the database of nanotechnology patents, developed by Meyer (2000, 2001).

3.2. Standards, terminologies, taxonomies and ontologies

In order to address the organization and homogenization of the huge volume of information that nanotechnology research is producing nowadays, scientists need support from lexical and semantic tools, such as terminologies, vocabularies, nomenclatures, code and synonyms sets, lexicons, thesauri, ontologies, taxonomies and classifications. Specifically, there is an urgent demand for guidance in labeling and annotation of nanomaterials, as well as for the production of electronic tools for standards development and revision. Besides, since ontologies are meeting urgent needs in other areas, such as biomedicine (Maojo *et al* 2011), it is reasonable to think about their applications in the nanomedical field. Ontology is a model describing the concepts of a specific domain, as well as their categories and relations. In nanoinformatics, ontologies can be considered a semantic framework representing the nanotechnology domain. This model may contain a myriad of chemical, physical and nanotechnology concepts that can be used as controlled terminology for data management and analysis, allowing further nanoinformatics tasks, such as the annotation of experimental data

²⁰ <https://cananolab.nci.nih.gov/caNanoLab/>.

²¹ <http://ncicb.nci.nih.gov/download/downloadcalab.jsp>.

²² <http://toxnet.nlm.nih.gov/>.

²³ <https://nanoexpert.usace.army.mil>.

²⁴ www.ncbi.nlm.nih.gov/books/NBK5330/.

²⁵ <http://ncl.cancer.gov/>.

or the integration of heterogeneous data sources, to be undertaken. Considering the extensive vocabulary and the wide range of names and terms for the same concepts in this area, ontologies would provide a mechanism to standardize the representation of the nanomedical domain that could be very useful for the dissemination of information and knowledge that exists in this field in a machine-readable format and to express the metadata related to nanomaterials in a semantically interoperable way.

In 1999, Freitas proposed a primer vocabulary for nanomedicine in the Nanomedicine Book Glossary (Freitas 1999), included in his nanomedicine book series²⁶, followed by a partial taxonomy of nanomedical technologies, published in 2005 (Freitas 2005). That same year, Tanaka developed a preliminary ontology for nanoscience, a meta-level ontology for nanoscale research (Tanaka 2005), based on an analysis of the subject headings and controlled index terms of papers published in journals related to materials science, physics, chemistry, engineering, medicine and biology.

Some popular indexes, such as the Library of Congress Classification²⁷, the Dewey Decimal Classification²⁸ and the Bliss Bibliographic Classification (Thomas 1997) are already indexing nanoscience. There also are biomedical and material science indexes that include nanotechnology, such as Medical Subject Headings, a controlled vocabulary for indexing publications (Lowe and Barnett 1994), and the NCI Thesaurus (Sioutos *et al* 2007)²⁹, developed by the NCI Center for Bioinformatics as part of the Enterprise Vocabulary Services Project, among others. Additionally, standardized terminologies for nanotechnology and nanosciences can be accessed through the websites of the main international organizations for standards, as noted in previous sections of this review, e.g. the Standard Terminology Relating to Nanotechnology included in American Society for Testing and Materials (ASTM) E2456-06³⁰ (published by ASTM International, formerly known as the American Society for Testing and Materials). Other non-standardized glossaries and vocabularies, developed by research institutions and companies, can be found on the web, such as the Glossary for NanoBiotechnology³¹, developed by JPK Instruments, and the Glossary of Nanotechnology Terms³², created by the UK Institute of Nanotechnology. The Dictionary of Nanotechnology³³ expands such glossaries with single-level classification, offering a nanotechnology vocabulary that contains terms collected from different scientific sites and references and classified into different categories: general, nanomaterial and nanoelectronic, nanomedical, nanobio and bionano, carbon nanostructure, nanocharacterization, nanofabrication and molecular nanotech.

Several general taxonomies for nanotechnology can also be found on the Internet. For instance, the *Nanotechnology Taxonomy* developed by Codesta³⁴, which divides nanotechnology products based on their manufacturing process by using three criteria: manufactured state, chemistry of fabrication and means of arrangement, and the *Nanomedicine Taxonomy* developed by the Canadian NanoBusiness Alliance³⁵, which classifies terms by their area of application (e.g. biopharmaceuticals, implantable devices, surgical aids, and so on). Another example is the taxonomy developed by Wong *et al* (2007) for classifying patents based on their area of application: (i) instrumentation, (ii) chemical processes and materials, (iii) medical and biotechnology and (iv) nanoelectronics. In a similar direction, the United States Patent and Trademark Office included in its classification a hierarchical terminology named Class 977 'Nanotechnology'³⁶ to categorize existing patents in the field. Nanosafe Inc.³⁷ also provides a logical classification of nanoparticles based on their elemental composition, properties and toxicity in support of defense-related nanotechnology applications.

²⁶ www.nanomedicine.com/.

²⁷ www.loc.gov/aba/cataloging/classification/.

²⁸ www.oclc.org/dewey/resources/summaries.en.html.

²⁹ <http://ncit.nci.nih.gov/ncitbrowser/pages/home.jsf?version=12.05d>.

³⁰ www.astm.org/Standards/E2456.htm.

³¹ www.jpk.com/glossary/index.htm.

³² www.nano.org.uk/nano/glossary.

³³ www.nanodic.com/.

³⁴ <http://nanotech2k3.bravepages.com/>.

³⁵ www.pain.cz/nanomedicina/files/taxonomy.pdf.

³⁶ www.uspto.gov/web/patents/classification/uspc977/defs977.htm.

³⁷ www.nanosafeinc.com/.

The InterNano initiative, a project of the National Nanomanufacturing Network funded by the NSF, is developing a general taxonomy³⁸ that defines the nanotechnology domain. The InterNano taxonomy brings together content about nanomanufacturing from a variety of sources and is divided in eight main categories, titled: (i) Areas of application, (ii) Environment, health and safety, (iii) Informatics and standards, (iv) Nanomanufacturing characterization techniques, (v) Nanomanufacturing processes, (vi) Nanoscale objects and nanostructured materials, (vii) Social and economic impact and (viii) Tool development. Many of these categories, in turn, are divided into several subcategories down to three levels of granularity. The taxonomy can be navigated in three different ways: as a hierarchical tree, as a tag cloud or as graph with nodes, presenting the relationships between the extant terms of the taxonomy. The database contains over 1000 elements and each element of the taxonomy is tagged and linked to the portal's contents and repositories, which include descriptions, scientific publications, technical reports, techniques, methodologies and patent information related to manufacturing processes for nanostructured materials and devices. It also includes lists of organizations and contributors to these processes, as well as usage statistics. Users can submit their own processes by uploading the related information in standard forms, facilitating the dissemination of this knowledge to organizations and laboratories.

The NanoParticle Ontology (NPO)³⁹, primarily developed by Thomas *et al* (2011) and currently supported by the NCI's cancer Biomedical Informatics Grid (caBIG®) Nanotechnology Working Group (Nano WG), provides the frame of reference for representing the cancer nanotechnology domain. A user can organize and annotate knowledge underlying the preparation, chemical composition and characterization of nanomaterials involved in cancer research by browsing and navigating the ontology's hierarchies of classes. Currently, NPO researchers are focused on support for ISA-TAB-Nano⁴⁰, an extension of the standard investigation/study/assay (ISA) format⁴¹ developed by the European Bioinformatics Institute, and other annotation projects. The NPO is also being updated to include a broader range of nanotechnology concepts. The continuous advancements in the development of this ontology translate to great steps forward for nanotechnology standardization and data interoperability, paving the way toward a global ontology for nanomedicine.

Despite of the myriad varieties of glossaries, vocabularies and taxonomies focused on nanotechnology, there is an urgent need for a unique controlled vocabulary for the nano domain and, specifically, a standard specification, to facilitate the import/export of data on nanomaterials and their characterizations to/from nanotechnology resources. A current major effort to reach consensus in nomenclature and categorization is ISA-TAB-Nano, a standard, approved by ASTM International, which aims to represent and share information, including characterization data, about nanomaterials. Developed by the NCI's caBIG® Nano WG, this standard provides a framework for describing investigations, studies, materials and assays through plain text files. In addition, ISA-TAB-Nano supports nanomaterials and concepts from the NPO, facilitating annotation and integration with this ontology.

3.3. Modeling and simulation tools

Creating *in silico* models and simulation environments at the nanoscale could assist in the prediction and assessment of risks of engineered nanomaterials; for instance, modeling differences in toxicity response due to changes in particle properties. In this respect, it would be necessary to determine which characteristics are most important for building a high-quality model, and to include these characteristics in accurate modeling and simulation tools. It would also be fundamental to abandon the current theoretical approach and consider factors such as the dependences between the nanoparticle's properties and the system where they are embedded or the behavior and time-scale of nanoparticles in the human body (degradation, toxicity, side effects, etc). For this purpose, extended knowledge about the biological identity of engineered nanomaterials (function, related

³⁸ www.internano.org/component/option,com_internanotaxonomy/.

³⁹ www.nano-ontology.org/.

⁴⁰ <https://wiki.nci.nih.gov/display/ICR/ISA-TAB-Nano>.

⁴¹ <http://isa-tools.org/>.

biological effects, domain of application, relationships between particle shapes and toxicity, etc) would be a requisite (Lynch *et al* 2009).

Also to be developed are modeling and simulation tools that would allow researchers to predict new properties and potential structures to acquire knowledge about the nature and behavior of nanoparticles. These data could include: (i) nanomanufacturing models for the virtual design of engineered nanomaterials considering various requirements simultaneously: nanomaterial physicochemical characterization, functionalization of the nanomaterial surface and design and control of synthesis processes, among others; (ii) models of pharmacokinetics and pharmacodynamics; (iii) models of nanoparticles' life cycle in the body: biodistribution, metabolism disposal, immunogenicity, clearance, phagocytosis, etc; and (iv) quantitative structure–activity relationships (QSARs) for modeling nanoparticle structure, nanotoxicity and nanoparticle interactions in biological environments at the tissue, cellular, molecular and atomic level. A good example is the work carried out at the UCLA Center for the Environmental Implications of Nanomaterials (CEIN). CEIN researchers have developed nano-QSARs (Puzyn *et al* 2011) for predicting nanoparticle toxicity (Liu *et al* 2013) as well as a set of web-based high-throughput data analysis tools for rapid analysis of nanotoxicity data⁴².

3.4. Educational tools

Due to the novel and heterogeneous character of nanoinformatics, the development of interdisciplinary and cross-disciplinary education programs integrating nanomedicine, biology, materials science, physics, chemistry and informatics has become a priority (Greenberg 2009). Over the last few years, several educational resources and repositories have emerged to give training support to researchers working in the field.

For instance, nanoHUB⁴³ (Qiao *et al* 2006) is a web portal for educational purposes developed by the Network for Computational Nanotechnology and sponsored by NSF under the NNI. This modeling and simulation framework offers real-time demonstrations of a myriad of extant computational resources in nanotechnology, materials science, physics and chemistry. NanoHUB does not allow the direct upload of experimental data, but users can run specific simulations and models by selecting the values for the different parameters requested by each resource. Users can also act as contributors, uploading their own resources and making them accessible to the vast scientific community supported by nanoHUB. Additionally, the portal provides access to courses and learning material—such as extensive lecture notes, quizzes, exams, homework and exercises—video lectures, animations, presentations about the available resources, online seminars, related scientific publications and information about events and job opportunities. There are also question-and-answer forums for interacting with professors and other students. Similar educational frameworks for e-learning on nanotechnology (e.g. Nano4Me⁴⁴) are gaining force in the web.

The NCL is also conducting educational activities by publishing and publicly disseminating the recent advances in preclinical studies on nanomaterials by means of scientific bibliographies and conferences, news, open forums and press notes in scientific media. There are also plenty of web pages and other media oriented to the general public that serve to distribute this information. For instance, Nanowerk⁴⁵ and NanoTechWeb⁴⁶ provide access to nanotechnology articles and news oriented to scientists and to the general public.

In spite of the didactic initiatives mentioned above, it is still necessary to develop and sustain additional educational resources and an international network of skilled professionals, as well as the supporting infrastructure and tools to advance training in nanotechnology. This involves the organization of workshops at international scientific events that engage students, technicians and trainees in nanotechnology environmental, health and safety considerations, and the establishment of a network of websites for student recruitment and dissemination of information, e.g. webinars and courses.

⁴² <http://nanoinfo.cein.ucla.edu/public/hdat/default.aspx>.

⁴³ <http://nanohub.org/>.

⁴⁴ <http://nano4me.org/webinars.php#showW5>.

⁴⁵ www.nanowerk.com/.

⁴⁶ <http://nanotechweb.org/>.

3.5. Working groups on nanoinformatics

In order to integrate and disseminate information on global nanotechnology research efforts, it is necessary to create and sustain international groups of interest by identifying experts in the field among networks, research organizations, public agencies and the nanotechnology industry. Working groups provide an excellent opportunity to bring together individuals representing a broad range of perspectives to accomplish a common goal. While scientific communities in long-established areas, such as bioinformatics, have demonstrated the high value of collaborative approaches to advance science, such efforts by the newer nanoinformatics community are less developed. However, this is changing quickly through the organization of joint workshops and publications, as well as the creation of working groups within in various sub-areas under development. These working groups are composed of scientists from different laboratories, research institutions and universities, committed to working together and sharing complementary capabilities to enhance nanotechnology research.

For instance, the caBIG[®] Nano WG⁴⁷ is developing concerted activities and several pilot projects⁴⁸. Decisions and processes within the group are consensus-driven and representatives from the participant groups meet regularly, generally once a week. Within this set of pilots, it is worthwhile to highlight the Meta-ontology Pilot⁴⁹, related to the NPO, and the ISA-TAB-Nano pilot⁵⁰, related to the standard previously described in this paper. In Europe, there are multiple collaborative initiatives, such as QNano⁵¹, a research infrastructure for nanomaterial safety and quality testing; NanoImpact.net⁵², a network concentrating on the health and environmental impact of nanomaterials; the EU NanoSafety cluster⁵³, a group of European Commission-funded projects working on nanotoxicology and nanosafety issues and involved in the preparation of new research topics for the next European-funded Framework Programme (Horizon 2020⁵⁴); the European Technology Platform on Nanomedicine⁵⁵, a group of industrial and academic stakeholders working for the establishment of a shared vision and research lines in the area; the European Technology Platform for Micro- and NanoManufacturing⁵⁶, a collaborative platform for discussing topics of interest to industrial micro- and nanomanufacturing; MODENA⁵⁷, an action modeling nanomaterial toxicity; and NANO futures⁵⁸, a shared environment for collaboration between industry, academy and society on nanotechnology issues which is currently creating a roadmap in the area, among many others. These European groups are working cooperatively to achieve their individual scientific goals, as well as to meet the common objective to advance nanotechnology research at the European level. In the international context, it is particularly worthwhile to mention the ongoing work of the US-EU dialogue, ‘bridging nanoEHS research’⁵⁹, which is supporting nanotechnology through the establishment of worldwide Communities of Research.

4. Cyberinfrastructures for nanomedical data management: current limitations and future prospects

The existing nanotechnology platforms, data resources and tools are powerful, but some aspects are already being improved. For instance, usually each service provider’s interface is different—both in the subset of data presented and in structure—and each institution promotes its own web sites, services and data formats. This presents a challenge to nanomedical researchers, who typically need to link data from multiple online

⁴⁷ <https://wiki.nci.nih.gov/display/ICR/Nanotechnology+Working+Group>.

⁴⁸ <https://wiki.nci.nih.gov/display/ICR/Pilots>.

⁴⁹ <https://wiki.nci.nih.gov/display/ICR/Meta-Ontology+Pilot>.

⁵⁰ <https://wiki.nci.nih.gov/display/ICR/ISA-TAB-Nano#ISA-TAB-Nano-Scope%3A>.

⁵¹ www.qnano-ri.eu/.

⁵² www.nanoimpactnet.eu/.

⁵³ www.nanosafetycluster.eu/.

⁵⁴ <http://ec.europa.eu/research/horizon2020/>.

⁵⁵ <http://cordis.europa.eu/nanotechnology/nanomedicine.htm>.

⁵⁶ www.minamwebportal.eu/index.php?m1=Public-Area.

⁵⁷ www.cost.eu/domains_actions/mpns/Actions/TD1204.

⁵⁸ www.nanofutures.eu/.

⁵⁹ <http://us-eu.org/>.

sources to generate a data set for further analysis. The development of new standards and specifications, like the ISA-TAB-Nano and the NPO, is directed at harmonizing access to nanotechnology data.

Another issue is that the development of inventories of nanomaterials has been done without automating the process, thus consuming a big part of researchers' time. Inventories are manually maintained by trained personnel who conduct periodic updates based on the new published material on nanoparticles. In fact, some of these data sources are not pure databases, but web pages containing lists of nanomaterial descriptions. In some cases, users can contribute by uploading new data. Nevertheless, now that relevant resources are widely distributed over the web and data are made available in different formats, manual integration has become almost infeasible and could lead to the publication of non-comprehensive indexes that include only a portion of the available data about nanomaterials. To replace the manual development of nanomaterials repositories, there are some quite innovative approaches that could generate indexes of nanomaterials in an automatic manner. These approaches would facilitate the initial implementation of the index and, above all, allow its continuous update and curation, adding new information and deleting the obsolete data. Due to the rapid growth of information about new experiments and results in nanomedicine, the automation of the creation and update process is an important element to be considered.

Joining data sources together is quite problematic, as they frequently have insufficient or non-existent metadata describing the nanomaterials. An important consequence of the lack of metadata is that some existent nanomaterials cannot be found in all the registries. Besides, sometimes the information provided by the different sources is not directly compatible with other sources and tools. These issues are being addressed by applying information technology methods and techniques and, therefore, a comprehensive infrastructure is needed to support an interoperable network of data and analytical services. Transforming the current efforts to develop databases and tools for nanomedical research into a dedicated nanoinformatics research infrastructure requires an understanding of the needs of the scientific community, the need for efficient and high-quality management of information at the nano level and an understanding of the strengths and weakness of experimental results and clinical data. This multidisciplinary, open and free infrastructure should provide a new way for scientists to immediately access and use data from nanomedical research, saving them from investing time in fruitless web searches. With such an infrastructure, researchers can focus on moving forward with their research, not on acquiring the necessary resources for data analysis and management.

For instance, in the United States, the NNI⁶⁰ has led to public funding for nanotechnology research (see figure 5). According to a recent NNI statement⁶¹, published on 14 May 2012, the establishment of a Nanotechnology Knowledge Infrastructure is a priority research line for the forthcoming years. The Subcommittee on Nanoscale Science, Engineering and Technology, within the National Science and Technology Council's Committee on Technology, announced, through the publication of a White Paper⁶², the fourth Nanotechnology Signature Initiative to stimulate the development of and improvements in the tools and methods required to access, organize and gather breakthroughs from huge amounts of data generated by nanotechnology, beginning to weigh in on just what data really need to be intelligently managed and stored.

4.1. Previous successes in the biomedical domain

Computational resources and infrastructures for biomedicine have successfully given support to address a major part of the challenges existing in the field of biomedicine. For instance, the development of databases, tools and services has accelerated advancements and findings in the omics fields. Researchers have benefited from these resources, adapting their professional skills to the new technological requirements for managing biomedical information. This has been the case for numerous existing omics databases that have proliferated in recent years, providing a common framework for research on genomics (e.g. Online Mendelian Inheritance in Man (Hamosh *et al* 2005), Ensembl GenBank (Benson *et al* 2012) and Artemis (Rutherford *et al* 2000)); proteomics (e.g. PRoteomics IDentifications Database (Martens *et al* 2005), Database of Interacting Proteins

⁶⁰ www.nano.gov/.

⁶¹ www.nano.gov/node/819.

⁶² www.nano.gov/sites/default/files/nki_nsi_white_paper_-_may_14_2012_secured.pdf.

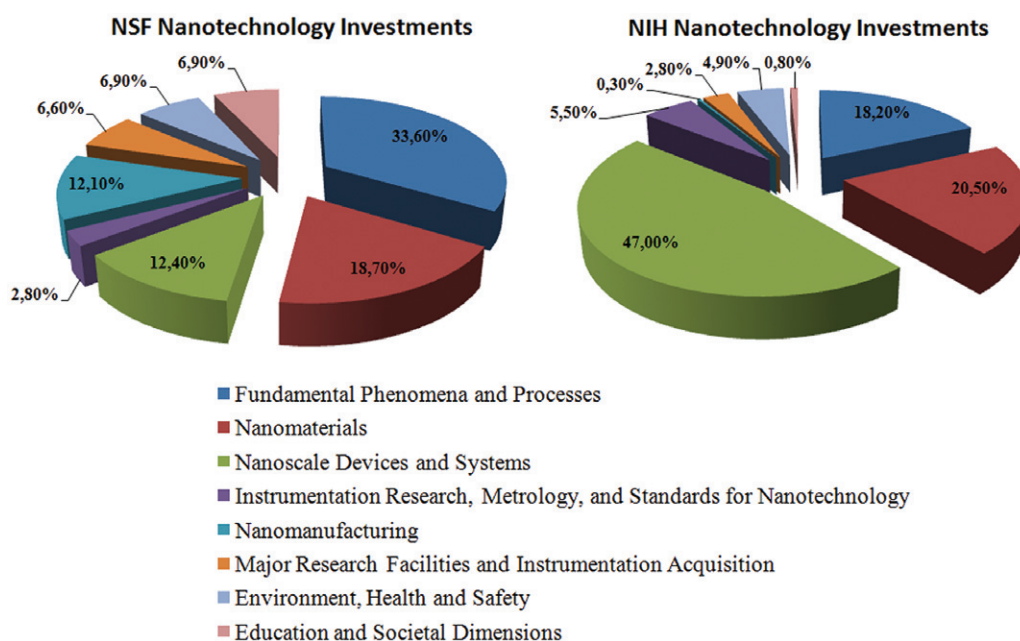


Figure 5. Comparison of the NNI investment distribution (2011–2013) into the different nano areas between two US governmental agencies: the NSF (covering general research on nanotechnology) and the NIH (focused on nanomedicine). Source: NNI Dashboard⁶³.

(Xenarios *et al* 2002), Swiss-Prot (Bairoch and Apweiler 2000) and Universal Protein Resource (Bairoch *et al* 2005)); epigenomics (e.g. NCBI Epigenomics); metabolomics (e.g. Human Metabolome Database (Wishart *et al* 2007)); and others.

An exemplary model to emulate would be the development strategy for the PDB (Berman *et al* 2000), the outstanding, worldwide protein structure database. The PDB has grown over the decades, since its creation in 1971, as a result of international collaboration between researchers who have deposited their protein structures in the database. This open strategy—where researchers can submit their own work, following a set of standards and guidelines, and research institutions act as storage, data processing and distribution centers for PDB data (Berman *et al* 2003)—led to the cooperative development, population and curation of this database.

In biomedicine, several application development projects have emerged for constructing and managing computational pipelines from existing bioinformatics databases and tools. A good example is iTools⁶⁴ (Dinov *et al* 2008), a system for classification, categorization and integration of different computational biology resources. This system has several functionalities, such as storage, searching, traversal and retrieval of resources and their corresponding metadata description. iTools infrastructure includes an ontology of computational biology resources, a database of descriptions of computational biology resources and an infrastructure for integration and interoperability of meta-algorithms for biomedical computing. Each resource is annotated with metadata that includes a short description, its authors, version, date and stage, the ontology classification of the resource, a set of keywords and the implementation language. In the ontologies' area, it is worthwhile to mention the BioPortal (Noy *et al* 2009), a compendium of ontologies for life sciences that is constantly updated with new ontologies and new versions of the existent ones. Regarding the indexing of bioinformatics databases and tools, there are some remarkable ontologies attempting to generate a standardized classification of bioinformatics resources, like the Biomedical Resource Ontology⁶⁵ (Musen *et al* 2008), a controlled terminology defining bioinformatics resources types under development by NIH researchers who have an interest in the classification of such resources. Medical imaging is another field that has grown considerably due to successful advances in the techniques for image acquisition, analysis, storage and

⁶³ <http://nanodashboard.nano.gov/>.

⁶⁴ <http://itools.loni.ucla.edu/>.

⁶⁵ <http://biportal.bioontology.org/>.

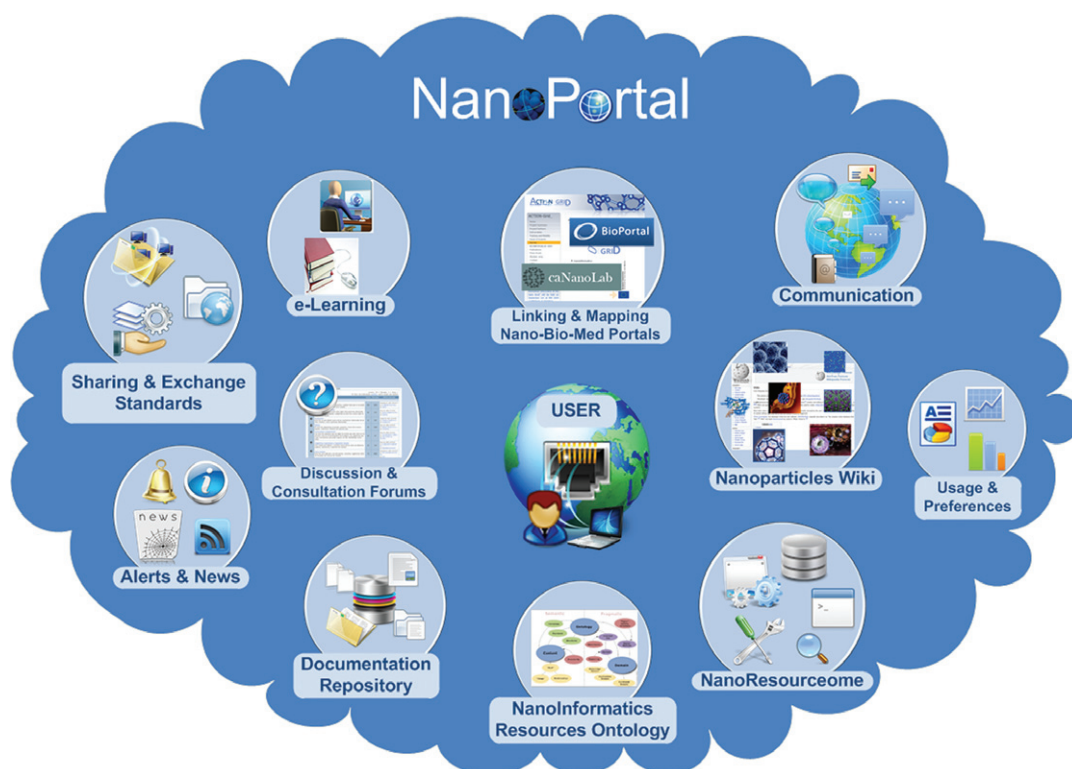


Figure 6. The NanoPortal: A hypothetical, multidisciplinary, open infrastructure enabling uniform access to a myriad of resources on nanotechnology.

annotation. A renowned resource in this field that should be mentioned is Digital Imaging and Communications In Medicine (DICOM)⁶⁶, the standard for exchanging medical images.

All of these resources are highly accessed by biomedical researchers on a daily basis, which indicates the relevant role for computational tools and infrastructures in conducting and reporting experimental research. While it is certainly correct that investigators typically do not control the systems in the sense of design, administration, etc, they do maintain the data. Thus, these sorts of approaches should be followed in the nanomedical area to develop and publish nanoinformatics resources.

4.2. Requirements for developing a nanoinformatics infrastructure

This section describes the desired elements of a hypothetical, multidisciplinary, open infrastructure, a research framework including a set of applications enabling easy access to a myriad of nanomaterials, enhancing cooperation among different institutions and researchers (see figure 6). This infrastructure should provide users with uniform access to a wide set of sources of nanomaterial data, not only collecting and integrating information from different databases but also offering information (metadata) about the source of the data. The nanoportal should offer services for sharing data and resources between researchers—e.g. scientific articles, demos, access to nanomedicine databases, ontologies, terminologies. It would also serve as dissemination channel for characterization data and aim at looking across characterization profiles and techniques. In order to accelerate the discussion and validation of nanomedicine research data, the infrastructure should support services for data annotation, data curation and data quality rating, which could improve the quality of published experimental results. Curation will rely on the literature descriptions as well as in registered experiments and results. The infrastructure should also help researchers to keep track of their experiments (techniques, parameters, protocols details) and then apply computational techniques and methods to analyze and inference knowledge from such data. We believe that such an open-environment approach could foster the development

⁶⁶ <http://dabsoft.ch/dicom/index.html>.

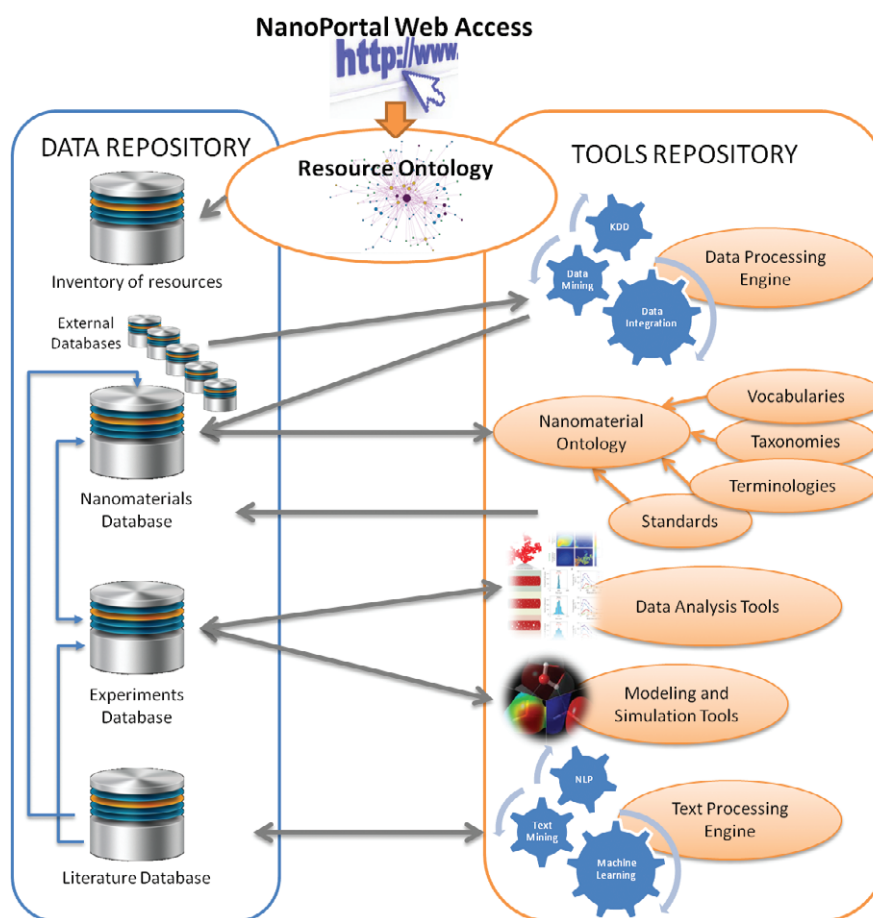


Figure 7. Schematic representation of the main components of the infrastructure and their interconnections.

of novel integrative tools, accessible over the Internet, to share nanotechnology knowledge and resources through a geographically distributed virtual community.

This infrastructure could be divided into two different parts (figure 7). The first one would consist of a *Data Repository* to support the exchange of nano research data, publish data generated by the research community and make these data freely available. All these data would be stored in a database to be queried over the Internet. Each nanomaterial would appear annotated with its corresponding information and the whole data set would be wrapped by an upper-level ontology. The second part would be the *Tool Repository* containing tools and services which would be classified into several categories: acquisition, calibration and quality assurance; data upload, storage and management; analysis and processing; data sharing and collaboration; visualization and ‘atlas’ tools; and standards, terminologies and ontologies. For this purpose, the automatic assignment of categories to the nanomaterials could be carried out with the use of text mining and NLP techniques to extract relevant information from papers describing them. Clustering engines could also help in this categorization by the application of standard mechanisms and algorithms, such as normalization, feature selection, consensus clustering and multidimensional scaling, among others (García-Remesal *et al* 2010). Nanomaterials would be clustered into sets so as to discover useful patterns (García-Remesal *et al* 2013). Data and tools gathered in the repositories should be committed to an open-source policy and the software should be freely available, allowing users to share their tools.

The creation of such a nanoinformatics platform would also involve the mandatory development of the following elements:

- *Terminologies, thesauri and vocabularies defining the nanotechnology domain and, specifically, the nanomedical semantic context:* the proposed data model would use nanomedical domain information

captured in terminologies, thesauri and taxonomies, which provide information related to nanomaterials, such as chemical descriptors, toxicity, size and structure. The organized description of nanomaterials through these properties will enable users to make advanced searches, such as all nanomaterials that perform a specific function or have a concrete composition.

- *Standards that will allow data and systems interoperability*: having a controlled terminology is a key factor for indexing and integrating nano sources; therefore, the platform should be supported by a set of well-established international standards.
- *Taxonomies and ontologies modeling the scientific knowledge existent in the nanomedical domain, including the design and implementation of services for searching and managing such taxonomies and ontologies*: for organizational purposes, the ontology could be separated into two different components: the domain ontology, an annotation vocabulary that would include descriptions of nanomaterial data types and their relationships, and the service ontology, which would describe the operational features of services within the domain ontology. The final goal of this ontology would be to capture the semantics of the relationship between the different nanomaterials, providing a contextual framework to share data and develop interoperable resources using this information.
- *Databases of related resources, bibliography, data sources, etc., including services for generating use statistics and monitoring user preferences*: as the nanotechnology field grows, it has become harder to find and access all the knowledge that is relevant to the discipline. In addition, non-specific search engines—such as Google—produce too general and massive results, so, in most cases, researchers are unable to find the best resource to obtain the target information necessary for their research. Therefore, it is necessary to solve the gap between the rising information needs of nanomedical researchers and the growing number of information sources available through the web, and create inventories indexing, describing and linking the existing data. As we have seen before, over the last years, several research groups and laboratories have developed directories of resources which work as ‘yellow pages’ for the biomedical scientific community. The same approach should be followed by the nanomedical community, promoting the dissemination and sharing of the currently available resources in the field. Then, it would be needed to describe potential applications of these indexes, including the automatic orchestration of nanoinformatics workflows, or novel methods to integrate public-online nanomedical databases and tools.
- Tools for the analysis and visualization of experimental results, accompanied by reference data sets for testing new algorithms.
- Tools for nanotechnology-oriented simulation and modeling.
- Communication services between users, including the development of tools for knowledge and resources exchange among researchers working in the area: This brokerage system would allow users to register their own nanomaterials and tools and provide an inventory for use by others in acquiring the desired data by means of a syntactically and semantically matching process, by entering words or phrases. This is a flexible novel paradigm already applied in biomedicine, which successfully integrates knowledge obtained from databases using Semantic Web technologies, sharing a common ontology and that can be accessed through normalized URLs.
- Documentation repositories providing researchers with intelligent search tools, as well as services for version control.
- Resources and tools for management of recent contents, alerts and news published in the area.
- Learning tools and consultation forums.

Finally, this hypothetical cyber-infrastructure should also support different views for different type of users, for instance: (i) small data users versus massive data users; (ii) novice users versus expert users; (iii) researcher’s view versus patient’s view versus administration’s view versus clinician’s view; and (iv) basic research versus translational research.

5. Conclusions

As discussed above, nanomedical research is producing huge volumes of information due to the numerous advances in the nanotechnology field. Due to the fact that the nature of this information is heterogeneous and has different levels of granularity, researches are facing many challenges in analyzing all these data. Consequently, the use of computational methods for the design and development of innovative analysis tools has progressively gained higher relevance, requiring new publicly accessible resources that facilitate the cooperative work between research groups working in the field.

In this paper, we have analyzed the urgent need to use automatic approaches to efficiently manage the new data generated during nanomedical research. This knowledge must be organized and structured, by means of the creation of standard domain models, and annotated by labeling data using existing terminologies, vocabularies and ontologies. The constant publication of new results in the area also demands a continuous update of data resources, which cannot be done manually, thus resources providing new methods for automatically updating these resources are crucial.

One of the main challenges in nanoinformatics is the variety and heterogeneity of the nanomedical domain, where users have very different needs, depending on their training and professional role. In order to understand the current problems and challenges and to develop the required applications to address them, nanoinformatics researchers should have a fluid communication with the scientific community conducting research in the area. This research involves professionals with varied profiles working in very different areas: chemistry, physics, medicine, the pharmaceutical industry, among others (including the materials science and engineering community), working closely with technology developers and manufacturers. It is necessary to consider social factors, as well, since they affect the acceptance of nanomedicine by consumers and its impact upon them in a global way. With this in mind, a great effort should go into undertaking and analyzing nanotoxicity studies, taking into account the results of previous studies, and the results, both positive and negative, managed as a complete information set and made easily available. Computational methods for modeling and simulation in nanomedicine are vital for understanding the behavior and interactions of engineered nanoparticles, as well as their biological effects within the human body and the environment.

Finally, due to the large number of new nanomaterials that are being produced by manufacturing processes, an open approach is necessary with regards to the publication of new research data. A major portion of the information about nanoparticles studies, assays and clinical trials can be found in the scientific literature, but there is a lack of publicly available databases and repositories of structured information. In order to accelerate advancements and improvements in nanomedicine and nanoinformatics, easy availability of experimental and clinical data is essential, allowing for inference of information from data and the subsequent integration of research results. Thereby, current informatics methods and techniques for text mining and NLP can help to characterize and classify nanomaterials, as well as identify data correlations from current research results.

By adopting such a big-picture approach, we can foster a community of integrated and coordinated research groups, facilitating the creation and support of interdisciplinary teams. Scientific groups collaborating through a nano-platform or cyber-infrastructure should include basic and translational scientists, clinical researchers and computer scientists. This kind of platform would also promote the establishment of public-private partnerships across the different areas involved, leading to long-term sustainability of funded research programs. The definitive establishment of nanotechnology in medicine and industry depends on all of these factors where, as we have seen, nanoinformatics is critical to improving the attainment of research results.

Acknowledgments

The work of the authors in this area of research has been partially funded by the European Commission (the ACTION-Grid Support Action, FP7-224176), the Spanish Ministry of Economy and Competitiveness (FIS/AES PS09/00069, RETICS COMBIOMED RD07/0067/0006, Ibero-NBIC CYTED 209RT0366) and the Consejo Social of the Universidad Politécnica de Madrid. This work has also been funded in part with

funds from the NCI-NIH (contract no. HHSN261200800001E). The contents of this publication do not necessarily reflect the views or policies of the DHHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

- Bailey R C, Kwong G A, Radu C G, Witte O N and Heath J R 2007 DNA-encoded antibody libraries: a unified platform for multiplexed cell sorting and detection of genes and proteins *J. Am. Chem. Soc.* **129** 1959–67
- Bairoch A and Apweiler R 2000 The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000 *Nucleic Acids Res.* **28** 45–8
- Bairoch A *et al* 2005 The Universal Protein Resource (UniProt) *Nucleic Acids Res.* **33** D154–9
- Baker N A, Fritts M, Guccione S, Paik D S, Pappu R V, Patri A, Rubin D, Shaw S Y and Thomas D G 2009 *Nanotechnology Informatics White Paper* caBIG National Cancer Institute February 2009 (non-peer reviewed manuscript)
- Bellare J R 2011 Nanotechnology and nanomedicine for healthcare: challenges in translating innovations from bench to bedside *J. Biomed. Nanotechnol.* **7** 36–7
- Benson D A, Karsch-Mizrachi I, Clark K, Lipman D J, Ostell J and Sayers E W 2012 GenBank *Nucleic Acids Res.* **40** D48–53
- Berman H M, Henrick K and Nakamura H 2003 Announcing the worldwide protein data bank *Nature Struct. Biol.* **10** 980
- Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N and Bourne P E 2000 The protein data bank *Nucleic Acids Res.* **28** 235–42
- Bikiaris D, Koutris E and Karavas E 2007 New aspects in sustained drug release formulations *Recent Pat. Drug Deliv. Formul.* **1** 201–13
- Braun T, Schubert A and Zsindely S 1997 Nanoscience and nanotechnology on the balance *Scientometrics* **38** 321–5
- Buxton D B 2009 Current status of nanotechnology approaches for cardiovascular disease: a personal perspective *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **1** 149–55
- de la Zerda A, Bodapati S, Teed R, May S Y, Tabakman S M, Liu Z, Khuri-Yakub B T, Chen X, Dai H and Gambhir S S 2012 Family of enhanced photoacoustic imaging agents for high-sensitivity and multiplexing studies in living mice *ACS Nano* **6** 4694–701
- de la Iglesia D, Chiesa S, Kern J, Maojo V, Martin-Sanchez F, Potamias G, Moustakis V and Mitchell J A 2009 Nanoinformatics: new challenges for biomedical informatics at the nano level *Stud. Health Technol. Inform.* **150** 987–91
- de la Iglesia D *et al* 2011 International efforts in nanoinformatics research applied to nanomedicine *Methods Inform. Med.* **50** 84–95
- de Mel A, Chaloupka K, Malam Y, Darbyshire A, Cousins B and Seifalian A M 2012 A silver nanocomposite biomaterial for blood-contacting implants *J. Biomed. Mater. Res. A* **100** 2348–57
- Dinov I D *et al* 2008 iTools: a framework for classification, categorization and integration of computational biology resources *PLoS ONE* **3** e2265
- Duncan R and Gaspar R 2011 Nanomedicine(s) under the microscope *Mol. Pharm.* **8** 2101–41
- Etheridge M L, Campbell S A, Erdman A G, Haynes C L, Wolf S M and McCullough J 2013 The big picture on nanomedicine: the state of investigational and approved nanomedicine products *Nanomedicine* **9** 1–14
- Fathabadi E G, Shelling A N and Al-Kassas R 2012 Nanocarrier systems for delivery of siRNA to ovarian cancer tissues *Expert Opin. Drug Deliv.* **9** 743–54
- Freitas R A Jr 1999 *Nanomedicine, Volume I: Basic Capabilities* (Georgetown, TX: Landes Bioscience) Available at: www.nanomedicine.com/NMI/Glossary.htm
- Freitas R A Jr 2005 What is nanomedicine? *Nanomedicine* **1** 2–9
- García-Remesal M, Cuevas A, Pérez-Rey D, Martín L, Anguita A, de la Iglesia D, de la Calle G, Crespo J and Maojo V 2010 PubDNA finder: a web database linking full-text articles to sequences of nucleic acids *Bioinformatics* **26** 2801–2
- García-Remesal M, García-Ruiz A, Pérez-Rey D, de la Iglesia D and Maojo V 2013 Using nanoinformatics methods for automatically identifying relevant nanotoxicology entities from the literature *Biomed. Res. Int.* **2013** 410294
- Ghanbari H, Cousins B G and Seifalian A M 2011 A nanocage for nanomedicine: polyhedral oligomeric silsesquioxane (POSS) *Macromol. Rapid Commun.* **32** 1032–46

- Giam L R, Massich M D, Hao L, Shin Wong L, Mader C C and Mirkin C A 2012 Scanning probe-enabled nanocombinatorics define the relationship between fibronectin feature size and stem cell fate *Proc. Natl Acad. Sci. USA* **109** 4377–82
- Greenberg A 2009 Integrating nanoscience into the classroom: perspectives on nanoscience education projects *ACS Nano* **3** 762–9
- Hamosh A, Scott A F, Amberger J S, Bocchini C A and McKusick V A 2005 Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders *Nucleic Acids Res.* **33** D514–7
- Hatto P 2009 International standards for risk management in nanotechnology *Nature Nanotechnol.* **4** 205
- Haun J B, Yoon T J, Lee H and Weissleder R 2011 Molecular detection of biomarkers and cells using magnetic nanoparticles and diagnostic magnetic resonance *Methods Mol. Biol.* **726** 33–49
- Huang Z, Chen H, Yip A, Ng G, Guo F, Chen Z K and Roco M C 2003 Longitudinal patent analysis for nanoscale science and engineering: country, institution and technology field *J. Nanopart. Res.* **5** 333–63
- Huh A J and Kwon Y J 2011 ‘Nanoantibiotics’: a new paradigm for treating infectious diseases using nanomaterials in the antibiotics resistant era *J. Control Release* **156** 128–45
- Hullman A and Meyer M 2003 Publications and patents in nanotechnology: an overview of previous studies and the state of the art *Scientometrics* **58** 507–27
- Jain K K 2011 Nanobiotechnology and personalized medicine *Prog. Mol. Biol. Transl. Sci.* **104** 325–54
- Jin S, Leach J C and Ye K 2009 Nanoparticle-mediated gene delivery *Methods Mol. Biol.* **544** 547–57
- Kamphuis M M, Johnston A P, Such G K, Dam H H, Evans R A, Scott A M, Nice E C, Heath J K and Caruso F 2010 Targeting of cancer cells using click-functionalized polymer capsules *J. Am. Chem. Soc.* **132** 15881–3
- Karalis V and Macheras P 2012 Current regulatory approaches of bioequivalence testing *Expert Opin. Drug Metable Toxicol.* **8** 929–42
- Kelkar S S and Reineke T M 2011 Theranostics: combining imaging and therapy *Bioconjug. Chem.* **22** 1879–903
- Kiessling F, Fokong S, Koczera P, Lederle W and Lammers T 2012 Ultrasound microbubbles for molecular diagnosis, therapy, and theranostics *J. Nucl. Med.* **53** 345–8
- Kircher M F et al 2012 A brain tumor molecular imaging strategy using a new triple-modality MRI-photoacoustic-Raman nanoparticle *Nature Med.* **18** 829–34
- Koria P 2012 Delivery of growth factors for tissue regeneration and wound healing *BioDrugs* **26** 163–75
- Kumbar S G, Kofron M D, Nair L S and Laurencin C T 2008 Cell Behavior Toward Nanostructured Surfaces *Biomedical Nanostructures* ed K Gonsalves, C Halberstadt, C T Laurencin and L Nair (New York: Wiley-Interscience) 261–95
- Lee L L, Chan C K, Ngaim M and Ramakrishna S 2006 Nanotechnology patent landscape 2006 *Nano* **1** 101–13
- Lemley M A 2005 Patenting nanotechnology *Stanford Law Rev.* **58** 601–30
- Li Y, Zheng X, Cao Z, Xu W, Zhang J and Gong M 2012 Self-assembled peptide (CADY-1) improved the clinical application of doxorubicin *Int. J. Pharm.* **434** 209–14
- Liu R, Zhang H Y, Ji Z X, Rallo R, Xia T, Chang C H, Nel A and Cohen Y 2013 Development of structure–activity relationship for metal oxide nanoparticles *Nanoscale* **5** 5644–53
- Liu S, Guo Y, Huang R, Li J, Huang S, Kuang Y, Han L and Jiang C 2012 Gene and doxorubicin co-delivery system for targeting therapy of glioma *Biomaterials* **33** 4907–16
- Lowe H J and Barnett G O 1994 Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches *JAMA* **271** 1103–8
- Lynch I, Salvati A and Dawson K A 2009 Protein–nanoparticle interactions: what does the cell see? *Nature Nanotechnol.* **4** 546–7
- Mailänder V and Landfester K 2009 Interaction of nanoparticles with cells *Biomacromolecules* **10** 2379–400
- Majzoub E H and Rönnebro E C E 2012 Methodology of materials discovery in complex metal hydrides using experimental and computational tools *Mater. Sci. Eng. Rep.* **73** 15–26
- Maojo V et al 2010 *The ACTION-Grid White Paper: Linking Biomedical Informatics, Grid Computing and Nanomedicine* Available at: www.action-grid.eu/index.php?url=whitepaper (non-peer reviewed manuscript)
- Maojo V, Crespo J, García-Remesal M, de la Iglesia D, Perez-Rey D and Kulikowski C 2011 Biomedical ontologies: toward scientific debate *Methods Inform. Med.* **50** 203–16
- Maojo V, Fritts M, de la Iglesia D, Cachau R E, Garcia-Remesal M, Mitchell J A and Kulikowski C 2012a Nanoinformatics: a new area of research in nanomedicine *Int. J. Nanomed.* **7** 3867–90
- Maojo V et al 2012b Nanoinformatics: developing new computing applications for nanomedicine *Comput. Sci. Eng.* **94** 521–39
- Martens L, Hermjakob H, Jones P, Adamski M, Taylor C, States D, Gevaert K, Vandekerckhove J and Apweiler R 2005 PRIDE: the proteomics identifications database. *Proteomics* **5** 3537–45

- McNeil S E 2011 Characterization of nanoparticles intended for drug delivery (*Methods in Molecular Biology* vol 697) ed S E McNeil (New York: Humana Press/Springer Science+Business Media) pp 3–8
- Mentovich E D, Livanov K, Prusty D, Sowwan M and Richter S 2012 DNA-nanoparticle assemblies go organic: macroscopic polymeric materials with nanosized features *J. Nanobiotechnol.* **10** 21
- Meyer M 2000 Patent citations in a novel field of technology: what can they tell about Interactions of emerging communities of science and technology? *Scientometrics* **48** 151–78
- Meyer M 2001 Patent citation analysis in a novel field of technology: an exploration of nano-science and nano-technology *Scientometrics* **51** 163–83
- Moosavi H, Zeynali M and Pour Z H 2012 Fracture resistance of premolars restored by various types and placement techniques of resin composites *Int. J. Dent.* **2012** 973641
- Murashov V and Howard J 2008 The US must help set international standards for nanotechnology *Nature Nanotechnol.* **3** 635–6
- Murashov V and Howard J 2011 Nanotechnology standards *Nanostructure Science and Technology* ed V Murashov and J Howard (New York: Springer Science+Business Media) pp 1–20
- Musen M A, Shah N H, Noy N F, Dai B Y, Dorf M, Griffith N, Buntrok J, Jonquet C, Montegut M J and Rubin D L 2008 BioPortal: ontologies and data resources with the click of a mouse *AMIA Annu. Symp. Proc.* 1223–4
- National Science and Technology Council 2011 Materials genome initiative for global competitiveness *White Paper* June 2011 Available at: www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf
- Noy N F et al 2009 BioPortal: ontologies and integrated data resources at the click of a mouse *Nucleic Acids Res.* **37** W170–3
- Ostraat M L, Mills K, Murry D and Guzan K 2012 Nanomaterial Registry: an authoritative resource for assessing environmental and biological interactions of nanomaterials *TechConnect WORLD: Summit and Innovation Showcase 2012 (Santa Clara, CA 18–21 June 2012)*
- Parrott M C, Finniss M, Luft J C, Pandya A, Gullapalli A, Napier M E and Desimone J M 2012 Incorporation and Controlled release of silyl ether prodrugs from PRINT nanoparticles *J. Am. Chem. Soc.* **134** 7978–82
- Pearce T R, Shroff K and Kokkoli E 2012 Peptide targeted lipid nanoparticles for anticancer drug delivery *Adv. Mater.* **24** 3803–22
- Perán M, García M A, López-Ruiz E, Bustamante M, Jiménez G, Madeddu R and Marchal J A 2012 Functionalized nanostructures with application in regenerative medicine *Int. J. Mol. Sci.* **13** 3847–86
- Perry J L, Herlihy K P, Napier M E and Desimone J M 2011 PRINT: a novel platform toward shape and size specific nanoparticle theranostics *Acc. Chem. Res.* **44** 990–8
- Petre C E and Dittmer D P 2007 Liposomal daunorubicin as treatment for Kaposi's sarcoma *Int. J. Nanomed.* **2** 277–88
- Ptak K, Farrell D, Panaro N J, Grodzinski P and Barker A D 2010 The NCI alliance for nanotechnology in cancer: achievement and path forward *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2** 450–60
- Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari T P, Michalkova A, Hwang H M, Toropov A, Leszczynska D and Leszczynski J 2011 Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles *Nature Nanotechnol.* **6** 175–8
- Qiao W, McLennan M, Kennell R, Ebert D S and Klimeck G 2006 Hub-based simulation and graphics hardware accelerated visualization for nanotechnology applications *IEEE Trans. Vis. Comput. Graph.* **12** 1061–8
- Roco M C 2003 Nanotechnology: convergence with modern biology and medicine *Curr. Opin. Biotechnol.* **14** 337–46
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M A and Barrell B 2000 Artemis: sequence visualization and annotation *Bioinformatics* **16** 944–5
- Sato K, Yokosuka S, Takigami Y, Hirakuri K, Fujioka K, Manome Y, Sukegawa H, Iwai H and Fukata N 2011 Size-tunable silicon/iron oxide hybrid nanoparticles with fluorescence, superparamagnetism, and biocompatibility *J. Am. Chem. Soc.* **133** 18626–33
- Scholes G D and Rumbles G 2006 Excitons in nanoscale systems *Nature Mater.* **5** 683–96
- Service R F 2005 Materials and biology. Nanotechnology takes aim at cancer *Science* **310** 1132–4
- Service R F 2010 Nanotechnology nanoparticle Trojan horses gallop from the lab into the clinic *Science* **330** 314–5
- Shimoni O, Postma A, Yan Y, Scott A M, Heath J K, Nice E C, Zelikin A N and Caruso F 2012 Macromolecule functionalization of disulfide-bonded polymer hydrogel capsules and cancer cell targeting *ACS Nano* **6** 1463–72
- Shukla G C, Haque F, Tor Y, Wilhelmsson L M, Toulmé J J, Isambert H, Guo P, Rossi J J, Tenenbaum S A and Shapiro B A 2011 A boost for the emerging field of RNA nanotechnology *ACS Nano* **5** 3405–18
- Sioutos N, de Coronado S, Haber M W, Hartel F W, Shaiu W L and Wright L W 2007 NCI thesaurus: a semantic model integrating cancer-related clinical and molecular information *J. Biomed. Inform.* **40** 30–43

- Tanaka M 2005 *Toward a Proposed Ontology for Nanoscience CAIS/ACSI 2005: Data, Information, and Knowledge in a Networked World* (London: The University of Western Ontario)
- Thaxton C S *et al* 2009 Nanoparticle-based bio-barcode assay redefines 'undetectable' PSA and biochemical recurrence after radical prostatectomy *Proc. Natl Acad. Sci. USA* **106** 18437–42
- Thomas A R 1997 Bibliographical classification: the ideas and achievements of Henry E Bliss *Cataloging Classification Q.* **25** 51–104
- Thomas D G, Pappu R V and Baker N A 2011 NanoParticle ontology for cancer nanotechnology research *J. Biomed. Inform.* **44** 59–74
- Upadhyay D, Scalia S, Vogel R, Wheate N, Salama R O, Young P M, Traini D and Chrzanowski W 2012 Magnetised thermo responsive lipid vehicles for targeted and controlled lung drug delivery *Pharm. Res.* **29** 2456–67
- Veetil J V and Ye K 2009 Tailored carbon nanotubes for tissue engineering applications. *Biotechnol. Prog.* **25** 709–21
- Wang Y, Byrne J D, Napier M E and DeSimone J M 2012 Engineering nanomedicines using stimuli-responsive biomaterials *Adv. Drug. Deliv. Rev.* **64** 1021–30
- Weidl T, Klein G and Zöllner R 2010 *International Handbook on Regulating Nanotechnologies* ed G A Hodge, D M Bowman and A D Maynar (Cheltenham: Edward Elgar Publishing) pp 462–83
- Wishart D S *et al* 2007 HMDB: the human metabolome database *Nucleic Acids Res.* **35** D521–6
- Wong P K, Ho Y P and Chan C K 2007 Internationalization and evolution of application areas of an emerging technology: the case of nanotechnology *Scientometrics* **70** 715–37
- Xenarios I, Salwinski L, Duan X J, Higney P, Kim S M and Eisenberg D 2002 DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions *Nucleic Acids Res.* **30** 303–5
- Xin G, Zhao X, Duan X, Ning Q, Meng M, Meng D and Liu L 2012 Antitumor effect of a generation 4 polyamidoamine dendrimer/cyclooxygenase-2 antisense oligodeoxynucleotide complex on breast cancer in vitro and in vivo *Cancer Biother. Radiopharm.* **27** 77–87
- Yeri A and Gao D 2011 Biosensing using nanoelectromechanical systems (*Biomedical Nanotechnology Methods and Protocols, Methods in Molecular Biology* vol 726) ed S J Hurst (Hatfield, Hertfordshire: Humana Press/Springer Science+Business Media) pp 119–40
- Zuckerman J E, Choi C H, Han H and Davis M E 2012 Polycation-siRNA nanoparticles can disassemble at the kidney glomerular basement membrane *Proc. Natl Acad. Sci. USA* **109** 3137–42