Published in final edited form as:

Nat Nanotechnol. 2013 February ; 8(2): 73–74. doi:10.1038/nnano.2013.12.

Standardizing data

Nathan A. Baker^{1,*}, Juli D. Klemm², Stacey L. Harper³, Sharon Gaheen⁴, Mervi Heiskanen², Philippe Rocca-Serra⁵, and Susanna-Assunta Sansone^{5,†}

¹Computational and Statistical Analytics Division, Pacific Northwest National Laboratory, MSID K7-28, PO Box 999, Richland, Washington 99352, USA

²Center for Biomedical Informatics and Information Technology, National Cancer Institute, Maryland 20852, USA

³Department of Environmental and Molecular Toxicology, School of Chemical, Biological and Environmental Engineering, Oregon State University, Oregon 97331, USA

⁴SAIC-Frederick, Frederick National Laboratory for Cancer Research, Information Systems Program, 6110 Executive Boulevard, Suite 250, Rockville, Maryland 20852, USA

⁵University of Oxford e-Research Centre, Oxford OX1 3QG, UK

To the Editor

To rationally design nanomaterials with improved efficacy and safety it is critical to understand and exploit the physicochemical properties that cause a biological response¹. Data mining and computer simulation are important for deriving information about the behaviour of nanomaterials, but the datasets needed to support such studies are sparse and stored across a variety of repositories and resources. The need for more coherence and structure in the conduct of nanotechnology research has been suggested before² and was highlighted in your recent Editorial³. However, the lack of common reporting standards and non-uniformity of information reported are also significant barriers to data sharing and re-use. The Nanotechnology Working Group (Nano WG) of the US National Institutes of Health National Cancer Informatics Program has been focused on addressing these issues.

The Nano WG — which includes representatives from over 20 organizations including government agencies, academia, industry, standards organizations and alliances — has developed ‘ISA-TAB-Nano’^{4,5}, a general framework for representing and integrating diverse types of data related to the description and characterization of nanomaterials using spreadsheet or TAB-delimited files. Nanoparticle characterization studies have many of the same challenges as omics-based (metabolomics, genomics and functional genomics, for example) assays such as high data volume and variety, multiple experimental end points, and complex protocols and study samples. Therefore, ISA-TAB-Nano is based on the ‘ISA-TAB’ format developed and used by the ISA Commons⁶ to share datasets in a diverse set of life sciences and in particular omics data. The ISA-TAB-Nano extension uses the three

primary files of ISA-TAB — investigation, study and assay (ISA) files — as well as an additional file called the material file.

Delivering a community-driven specification for nanotechnology data is only the first phase of the process. To be useful, ISA-TAB-Nano must be implemented in tools and by databases to assist researchers in reporting their data while shielding them from unnecessary complexity. Our next step is to extend the open source ISA software suite to provide user-oriented tools for the collection, curation and storage of data that is compliant with the ISA-TAB-Nano specification. Future work will also focus on using the ISA-TAB-Nano format to support emerging standards in the characterization of nanomaterials in biological research^{7,8}. ISA-TAB-Nano development is a community-driven effort and we welcome new contributions, collaborations and expertise.

References

1. Morris J, et al. Nature Nanotech. 2011; 6:73–77.
2. Schrurs F, Lison D. Nature Nanotech. 2012; 7:546–548.
3. Nature Nanotech. 2012; 7:545.
4. Thomas DG, et al. BMC Biotechnol. in the press.
5. <https://wiki.nci.nih.gov/display/ICR/ISA-TAB-Nano>
6. Sansone SA, et al. Nature Genet. 2012; 44:121–126. [PubMed: 22281772]
7. MINChar Initiative. <http://characterizationmatters.org>
8. Nanomaterial Registry. <https://www.nanomaterialregistry.org>