

NANOINFORMÁTICA: RETOS E INICIATIVAS PARA LA GESTIÓN DE LA INFORMACIÓN GENERADA EN LA INVESTIGACIÓN NANOMÉDICA

D. DE LA IGLESIA, V. MAOJO, M. GARCÍA-REMESAL, A. GARCÍA-RUIZ, J. CRESPO

Grupo de Informática Biomédica. Departamento de Inteligencia Artificial, Facultad de Informática. Universidad Politécnica de Madrid. 28660 Boadilla del Monte, Madrid. España

Durante la última década la investigación en nanomedicina ha generado gran cantidad de datos, heterogéneos, distribuidos en múltiples fuentes de información. El uso de las Tecnologías de la Información y la Comunicación (TIC) puede facilitar la investigación médica a escala nanométrica, proporcionando mecanismos y herramientas que permitan gestionar todos esos datos de una manera inteligente. Mientras que la informática biomédica comprende el procesamiento y gestión de la información generada desde el nivel de salud pública y aplicación clínica hasta el nivel molecular, la nanoinformática extiende este ámbito para incluir el “nivel nano”, ocupándose de gestionar y analizar los resultados generados durante la investigación en nanomedicina y desarrollar nuevas líneas de trabajo en este espacio interdisciplinar. En esta nueva área científica, la nanoinformática (que podría consolidarse como una auténtica disciplina en los próximos años), el Grupo de Informática Biomédica (GIB) de la Universidad Politécnica de Madrid (UPM) participa en numerosas iniciativas, que se detallan a continuación.

1. Introducción

El conocimiento científico en biología se ha visto enriquecido con nuevas dimensiones, estrechamente ligadas a los procesos biológicos, químicos y físicos que se manifiestan a escala nanométrica. En este contexto, la nanomedicina surge como la aplicación de la nanotecnología en el ámbito médico para el diagnóstico, el tratamiento y la prevención de enfermedades [1, 2].

Con el desarrollo de la genómica, la proteómica, la nanomedicina, la biología molecular y la biología de sistemas, entre otras áreas científicas, podemos avanzar en la comprensión de los mecanismos implicados en los procesos biológicos a nivel celular, molecular y atómico. Las numerosas cuestiones de investigación que surgen son demasiado complejas para encontrar respuestas en una sola disciplina, por lo que es necesario un esfuerzo colaborativo y multidisciplinar. Desde el año 2008, el GIB lidera y participa, junto con grupos americanos y europeos, diversos proyectos e iniciativas conjuntas [3, 4], que intentan dar soporte a las necesidades existentes en el campo de la nanoinformática.

2. Nanoinformática

La investigación en nanomedicina implica la gestión y el análisis masivo de datos complejos y muy heterogéneos [5]. En este contexto, una nueva disciplina, denominada nanoinformática, surge con el objetivo de construir nuevos puentes entre la medicina, la nanotecnología y la informática mediante la aplicación de herramientas y métodos computacionales para facilitar las tareas y resolver los problemas que surgen durante la investigación nanomédica. Su principal meta es introducir nuevos enfoques en el tratamiento y la gestión de la información.

A pesar de las diferencias en características biofísicas y bioquímicas que presentan los nanomateriales respecto a otras entidades biológicas y físico-químicas que aparecen en otros niveles de organización biológica y de las correspondientes diferencias en la complejidad de la información, algunos de los métodos propios de la informática biomédica pueden ser aplicados en nanoinformática, con algunas adaptaciones.

3. Retos actuales de la nanoinformática

Durante la última década, el espectacular desarrollo de la nanotecnología ha planteado numerosos retos para la nanomedicina, que deben ser afrontados por la comunidad científica y, en particular, por los grupos de investigación que trabajan en informática biomédica y nanoinformática. En concreto, nuestro grupo ha colaborado activamente en la definición de la nanoinformática, coordinando el proyecto europeo ACTION-Grid [3, 6], pionero europeo en esta disciplina, y ha liderado la elaboración de un “White Paper” en el área [7, 8].

A continuación se detallan los retos principales existentes en este campo, así como las tareas e iniciativas desarrolladas por el GIB en relación con dichos retos.

b. Información relevante en nanomedicina. Fuentes de datos

A menudo, los investigadores se encuentran con dificultades para distinguir qué fuentes de información son las más fiables, así como para identificar qué datos específicos son relevantes para su investigación. En este contexto, la nanoinformática pretende, entre otros objetivos, automatizar la búsqueda y gestión de información sobre los estudios, proyectos y resultados de investigaciones en el área. Para ello, alguno de los retos y prioridades que se plantean son:

- El análisis de la información proporcionada por las diversas fuentes de datos disponibles, comparando dichas fuentes e identificando posibles carencias.
- La definición de los criterios que definen la calidad de un estudio en nanomedicina, permitiendo el desarrollo de herramientas que faciliten la identificación automática de los estudios que cumplen dichos criterios, así como la definición de la información mínima necesaria para que los conjuntos de datos sean completos.
- La identificación de los posibles usuarios de estas herramientas, así como de sus criterios para evaluar la calidad e utilidad de los estudios.
- La identificación del uso que se dará a la información extraída.

c. Estándares e interoperabilidad

Dado al alto número de ensayos, estudios y experimentos que se generan en la actualidad, el intercambio de datos y resultados de las investigaciones en nanomedicina es un tema crucial para el desarrollo de esta disciplina.

En este campo existen diversos problemas, a los que la nanoinformática puede dar solución. Por ejemplo, las mediciones de los experimentos realizados dependen de protocolos no estandarizados, así como del tipo de tecnología empleado. Además, las descripciones de los diversos nanomateriales no siguen ningún formato ni estándar. Es decir, no existe un intercambio de datos significativo entre la comunidad científica debido a la falta de normalización.

Uno de los retos principales de la nanoinformática es el establecimiento de redes de colaboración para la conexión de aplicaciones y herramientas, así como para el intercambio de datos, teniendo como objetivo final la interoperabilidad entre los sistemas y servicios ya existentes.

Como ejemplo, podemos citar la iniciativa Nano-TAB (<http://sites.google.com/site/cabignanowg/data-sharing-and-nanotechnology-standards/nanotab>), que consiste en una especificación para facilitar el intercambio de datos entre recursos de nanotecnología y, a su vez, proporciona a los investigadores guías y directrices para la representación y caracterización de nanomateriales, lo cual permitirá comparar dichos materiales posteriormente.

d. Clasificación de nanopartículas

Otro reto importante es la clasificación de las nanopartículas [9]. Existen diferentes clasificaciones pero el alcance de las mismas y los criterios en que se basan son muy diversos. El principal reto es aquí es definir el tipo de datos e información necesarios y utilizados durante la investigación en nanomedicina, con un especial énfasis en las características físicas, químicas y estructurales de las nanopartículas.

Un ejemplo actual de herramienta nanoinformática para la clasificación de nanopartículas es la Nanoparticle Ontology, una ontología desarrollada para representar el conocimiento relativo a la preparación, composición química y caracterización de los nanomateriales empleados en la investigación del cáncer [10].

Por otra parte, un estudio reciente [11] demuestra que el potencial toxicológico y patogénico de un material cambia según la forma de dicho material—en concreto, cuando toma una forma que dificulta el procesado por las células fagocíticas. El Grupo de Informática Biomédica está trabajando actualmente en el desarrollo de ontologías de formas geométricas, que permitan clasificar las nanopartículas en base a su estructura tridimensional [12].

e. Repositorios de datos y herramientas

Los objetivos principales de la nanoinformática en este campo son:

- Crear repositorios de datos de carácter global y cooperativo, mediante la colaboración de los grupos que trabajan en el área. Como ejemplo de estos repositorios podemos citar el Nanomaterial Biological Interaction repository (<http://nbi.oregonstate.edu/>) — una base de conocimiento de datos anotados sobre caracterización de nanomateriales (pureza, tamaño, forma, carga, composición, etc.), métodos de síntesis e interacciones biológicas de los nanomateriales definidas a diferentes niveles (molecular, celular, a nivel de órgano)— y la Nanoparticle Library (www.nanoparticlelibrary.net) —un repositorio que contiene información sobre nanopartículas, incluyendo su composición, sus características físicas y químicas, su tamaño y morfología, sus aplicaciones potenciales, etc.
- Desarrollar herramientas para la visualización y el análisis de datos originados durante la investigación en nanomedicina, así como formar a la comunidad científica en el uso de dichas herramientas.

- Utilizar metadatos para la descripción de las herramientas y de los datos generados por las mismas. La descripción de las entradas y salidas producidas por las herramientas de análisis y visualización sería muy útil para la construcción de flujos de trabajo. Como ejemplo de iniciativas de este tipo, se puede citar el portal web Nanohub (<http://nanohub.org/>), el cual proporciona documentos XML describiendo las entradas y salidas de cada una de sus herramientas y servicios.
- Elaborar un sistema de ranking de datos y fuentes para la investigación en nanomedicina.
- Nuevos métodos automatizados de búsqueda de información y análisis de la literatura científica.
- Desarrollar repositorios de métodos computacionales, herramientas informáticas y servicios para facilitar las tareas de búsqueda de información de los investigadores. En el Grupo de Informática Biomédica de la UPM estamos desarrollando actualmente un inventario de recursos nanoinformáticos con información anotada sobre cada uno de los recursos [13,14].
- Desarrollar terminologías biomédicas y metatesauros con información integrada de diversas fuentes que cubran los términos utilizados por las instituciones y grupos de investigación más relevantes en las áreas de aplicación clínica, investigación básica, investigación translacional y salud pública. Como ejemplo de iniciativa, se puede citar el metatesauro del National Cancer Institute (NCI) (<http://ncimeta.nci.nih.gov/>) que contiene más de 20 millones de conceptos.

f. Nanotoxicología

Otra de las líneas de investigación relevantes en nanomedicina es el análisis y tratamiento de la toxicidad de los nanomateriales. El problema principal es la falta de un criterio estandarizado de caracterización de nanopartículas orientado a la identificación de la toxicidad y riesgos de los nanomateriales [15]. En esta dirección, nuestro grupo ha desarrollado una base de datos de nanopartículas asociadas a sus aplicaciones médicas y a sus efectos tóxicos [16, 17] y trabaja con un grupo del Instituto de Salud Carlos III —dirigido por el Dr. Fernando Martín Sánchez— en esta línea. Esta base de datos ha sido creada automáticamente a partir de la literatura científica existente en el área y apoyándose en el uso de dos ontologías: la Nanoparticle Ontology, descrita previamente, y el Foundational Model of Anatomy (FMA) [18].

g. Modelización y Simulación

Dada la complejidad y la gran diversidad biológica entre individuos aparentemente similares, los modelos matemáticos son, claramente, de alta importancia para la experimentación y el desarrollo efectivo de la investigación en nanomedicina y, en especial, de la medicina personalizada.

La modelización matemática y la simulación de los fenómenos biológicos, físicos y químicos es crucial para [19]:

- La identificación de los parámetros principales que rigen la evolución espacial y temporal de los sistemas biológicos.
- La interpretación del rol de esos parámetros y la cuantificación y evaluación de sus efectos.

- La predicción de la evolución del sistema, sin necesidad de recurrir a experimentos largos y costosos.

Por otra parte, los modelos matemáticos y simulaciones también pueden ser útiles para diseñar e inspirar nuevos experimentos, así como para predecir los resultados de los mismos. En este sentido, hay abundantes sinergias con el llamado “Virtual Physiological Human” (VPH), un macroprograma de investigación lanzado por la Comisión Europea para crear modelos y simulaciones de todo el cuerpo humano.

5. Conclusiones

Los desafíos que plantea la nanomedicina implican el desarrollo de nuevas estrategias científicas, así como el impulso del diálogo que permita establecer las prioridades en esta nueva área, discutiendo futuras iniciativas y actividades de cooperación. En este sentido, el uso de las Tecnologías de la Información puede facilitar y fomentar la investigación médica en la escala “nano”.

Desde una perspectiva tecnológica, es realista prever que la nanotecnología continuará logrando grandes avances para su aplicación en el ámbito médico: detección temprana de enfermedades, regeneración de tejidos, administración de fármacos, biomarcadores, etc. Esto permitirá a los pacientes recibir tratamientos personalizados, a través de la integración y el análisis detallado de sus datos en todos los niveles de organización biológica, incluyendo el nivel “nano”.

Estos desafíos también implican el establecimiento de las bases de un futuro programa de investigación en nanoinformática, así como la promoción de nuevas comunidades científicas de investigación interdisciplinar. La nanoinformática brinda aquí la oportunidad de utilizar la creciente cantidad de herramientas y métodos informáticos que existen en la actualidad, para poder usar más eficazmente la gran riqueza de datos e información que la investigación nanomédica está produciendo.

6. Agradecimientos

Este trabajo ha sido financiado, en parte, por la Comisión Europea a través de los proyectos ACTION-Grid (FP7-ICT-2007-224176) y ACGT (FP6-2005-IST-026996); por el Ministerio de Ciencia e Innovación mediante los proyectos FIS/AES PS09/00069 y COMBIOMED-RETICS; por la Red Ibero-NBIC (CYTED 209RT0366) y por la Comunidad de Madrid.

Referencias

- [1] K. K. Jain. *The Handbook of Nanomedicine*. 1st ed. Humana (2008).
- [2] S. Singh. *Nanomedicine-nanoscale drugs and delivery systems*. *J Nanosci Nanotechnol*. **10**(12):7906-18 (2010).
- [3] ACTION-Grid: International cooperative action on grid computing and biomedical informatics between the European Union, Latin America, the Western Balkans and North Africa (FP7-224176). Disponible en: <http://www.vph-action-grid.eu>.

- [4] V. Maojo, D. de la Iglesia, F. Martín-Sánchez, M. García Remesal, J. Crespo et al. *Nanoinformatics in Europe: The ACTION Grid White Paper. Nanoinformatics 2010 Conference*. 3-5 Noviembre 2010. Arlington, Virginia (USA).
- [5] K. Riehemann, S. W. Schneider, T. A. Luger, B. Godin, M. Ferrari y H. Fuchs. *Nanomedicine-Challenge and Perspectives. Angew Chem Int Ed Engl.* **48**(5):872-897 (2009).
- [6] D. de la Iglesia, V. Maojo, S. Chiesa, F. Martín-Sánchez, J. Kern, G. Potamias, J. Crespo, M. García-Remesal, S. Keuchkerian, C. Kulikowski y J. A. Mitchell. *International efforts in nanoinformatics research applied to nanomedicine. Methods Inf Med.* **50**(1):84-95 (2011).
- [7] V. Maojo, F. Martín-Sánchez et al. The ACTION-Grid White Paper: Linking Biomedical Informatics, Grid Computing and Nanomedicine (2010). Disponible en: <http://www.action-grid.eu/index.php?url=whitepaper>.
- [8] V. Maojo, F. Martín-Sánchez, C. Kulikowski, A. Rodríguez-Paton y M. Fritts. *Nanoinformatics and DNA-based computing: catalyzing nanomedicine. Pediatr Res.* **67**(5):481-9 (2010).
- [9] D. Berube, C. Cummings, M. Cacciatore, D. Scheufele y J. Kalin. *Characteristics and classification of nanoparticles: Expert Delphi survey. Nanotoxicology*. 2010 Sep 30. En prensa.
- [10] D. G. Thomas, R. V. Pappu y N. A. Baker. *NanoParticle Ontology for cancer nanotechnology research. J Biomed Inform.* 2010 Mar 6. En prensa.
- [11] R. F. Hamilton, N. Wu, D. Porter, M. Buford, M. Wolfarth y A. Holian. *Particle length-dependent titanium dioxide nanomaterials toxicity and bioactivity. Part Fibre Toxicol.* 2009;**6**:35.
- [12] V. Maojo, J. Crespo, M. García-Remesal, D. de la Iglesia, D. Pérez-Rey y C. Kulikowski. *Biomedical Ontologies: Toward scientific debate. Methods Inf Med.* (2011). En prensa.
- [13] S. Chiesa, M. García-Remesal, G. de la Calle, D. de la Iglesia, V. Bankauskaite y V. Maojo. *Building an Index of Nanomedical Resources: An Automatic Approach Based on Text Mining. Proceedings of the KES 2008;* **2**: 50-57 (2008).
- [14] G. Calle, M. García-Remesal, S. Chiesa, D. de la Iglesia y V. Maojo. *BIRI: a new approach for automatically discovering and indexing available public bioinformatics resources from the literature. BMC Bioinformatics* **10**(1):320 (2009).
- [15] G. Oberdörster, A. Maynard, K. Donaldson, V. Castranova et al.. *Principles for characterizing the potential human health effects from exposure to nanomaterials: elements of a screening strategy. Part Fibre Toxicol.* **2**:8 (2005).
- [16] Linares Quintana, A. *Herramienta para el descubrimiento de efectos tóxicos de nanomateriales en aplicaciones médicas a partir de la literatura*. Facultad de Informática, Universidad Politécnica de Madrid; Mayo 2010.
- [17] V. Maojo, M. García-Remesal, D. de la Iglesia, J. Crespo et al. *Intracellular Drug Delivery: fundamentals and applications Capítulo: Nanoinformatics: developing advanced informatics applications for Nanomedicine*. Ed: A. Prokop (2011). En prensa.
- [18] J. H. Sakamoto, A. L. van de Ven, B. Godin, E. Blanco et al. *Enabling individualized therapy through nanotechnology. Pharmacol Res.* **62**(2):57-89 (2010).
- [19] C. Rosse y J. V. L. Mejino. *A reference ontology for biomedical informatics: the Foundational Model of Anatomy. J Biomed Inform.* **36**:478-500 (2003).