

Almacenando datos: calidad frente a cantidad

Los datos son la representación del mundo real en un sistema de información. La calidad de un dato viene definido entonces por las características del sistema de información en el que está referenciado. Esta calidad depende de la perspectiva y las necesidades de los usuarios de ese sistema. Cuantificar la calidad de un dato permite gestionar dicha calidad, porque si algo se puede medir, se puede gestionar: para maximizar, minimizar u obtener un óptimo balance entre diferentes datos imponiendo las restricciones adecuadas a nuestros fines. Cuantificar la calidad de un dato implica determinar dimensiones para esta calidad, es decir, hallar una representación medible en el sistema de referencia. Existen muchas propuestas de conjuntos de dimensiones de calidad de un dato, y muchas dimensiones son comunes a muchas propuestas como: precisión, confiabilidad, exactitud, etc... Aunque como se ha dicho, el marco de referencia de un dato influye en su calidad y también en las dimensiones para determinarla.

En el caso un Biobanco, para cuantificar la calidad de los datos asociados a muestras es necesario comprender el sistema de referencia y determinar las dimensiones medibles de la calidad que son importantes para sus usuarios. Desde nuestro Programa hemos propuesto un primer conjunto de dimensiones adecuadas para determinar la calidad de los datos integrados en las colecciones de los biobancos pertenecientes a la RNBB. Los objetivos de esta primera aproximación son servir como guía para saber en qué punto estamos, ofrecer un sistema de referencia para encontrar soluciones que nos permitan incrementar paulatinamente la calidad de los datos en nuestras colecciones y servir de punto inicial para la mejora continua del propio sistema de referencia de calidad. Es decir, se espera que cuantificando nuestra calidad de datos, aprendamos a obtener datos cada vez más útiles y valiosos.