# INTRODUCING A METADATA CHECKLIST FOR OMICS DATA

*Edd Dumbill*[1] *and Eugene Kolker*[2,3,4]

How many times have you ignored a phone call because the number is unknown to you? Without the crucial context—the "metadata" of the data (i.e., who is on the other end)—the call is not effective. If metadata are missing from valuable life sciences data, the data can be similar to that phone call—just a string of ineffective numbers that go unused.

Some of the most valuable data gathered by the life sciences are those that describe not just one aspect of life, but multiple aspects at once. These data may include not only the now-familiar genomics (study of the DNA of the sample) but also proteomics (study of the proteins), metabolomics (study of the metabolic products made by the cells), and others. The complex and diverse "multi-omics" data are challenging the analytical abilities of data and life scientists alike while also straining the data repository systems—yet they are a critical part of our efforts to better understand biological, environmental, and medical problems.

The Data-Enabled Life Sciences Alliance (DELSA Global, delsaglobal.org) has created a multi-omics metadata checklist that is flexible yet comprehensive.[1] It provides a framework to capture the metadata needed for more complete utilization of single and multi-omics data sets.

The checklist can be used as part of a larger publication or as a stand-alone data publication[2] to inform the community of an available data resource. This will give credit to data producers, harmonize different omics data sets, and use research resources for science advancement more effectively.

We are publishing this checklist and accompanying article in the current issue of *Big Data,* along with simultaneous publication in the journal *OMICS: A Journal of Integrative Biology* (www.liebertpub.com/OMI).

## References

1. Kolker E, Özdemir V, Martens L, et al. Towards more transparent and reproducible omics studies through a common metadata checklist and data publications. Big Data 2013; 1:196–201.
2. Snyder M, Mias G, Stanberry L, Kolker E. Metadata checklist for the integrated personal omics study: proteomics and metabolomics experiments. Big Data 2013; 1:202–206.

Address correspondence to:

*Edd Dumbill*
Editor-in-Chief

*E-mail:* edumbill@acm.org

*Eugene Kolker*
Executive Editor

*E-mail:* eugene.kolker@seattlechildrens.org

[1]*Editor-in-Chief,* Big Data.
[2]*Bioinformatics and High-Throughput Analysis Laboratory, Seattle Children's Research Institute; Seattle, Washington.*
[3]*Predictive Analytics, Seattle Children's, Seattle, Washington.*
[4]*DELSA Global, Seattle, Washington.*