When is big data too big? Making data-based models comprehensible

July 11 2016

Credit: Mary Ann Liebert, Inc., publishers

Data-driven mathematical modeling is having an enormous impact on the ability to organize and describe very large data sets, and make inferences and predictions about populations and situations based on sampling data. However, as these models become increasingly complex, the ability of users to understand and apply them represents a growing

challenge. The article "A Framework for Considering Comprehensibility in Modeling", which describes this emerging dilemma and a strategy for developing solutions, is published in *Big Data*.

Michael Gleicher, University of Wisconsin-Madison, defines comprehensibility as "the ability of the various stakeholders to understand relevant aspects of the modeling process." He suggests that comprehensibility should be a key goal in model development. However, as models become more sophisticated, tradeoffs may be inevitable—even between understandability and accuracy—in some cases, improving comprehensibility may help achieve other goals in modeling.

"Gleicher provides a holistic framework of comprehensibility that considers what the various stakeholders in a data science project do and don't understand easily and their need for comprehensibility," says *Big Data* Editor-in-Chief Vasant Dhar, Professor at the Stern School of Business and the Center for Data Science at New York University. "More broadly, the article highlights comprehensibility from a human-centric standpoint, identifying the role and needs of humans in complex data science projects."

More information: Michael Gleicher, A Framework for Considering Comprehensibility in Modeling, *Big Data* (2016). DOI: 10.1089/big.2016.0007

Provided by Mary Ann Liebert, Inc

Citation: When is big data too big? Making data-based models comprehensible (2016, July 11) retrieved 11 May 2024 from https://phys.org/news/2016-07-big-data-based-comprehensible.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.