CAUTION: ANONYMIZATION IS NOT EASY

A very interesting article "Authentic Health-Care Scenarios Make Mathematics Meaningful" (May 2018, pp. 496-502) by Melissa Geist, Holly Anthony, and Twanelle Majors described how school mathematics can be helpful in processing medical data, and advises teachers to get real-life medical data by asking care providers to anonymize the data by stripping off all identifying markers such as name, patient record numbers, etc. As computer scientists well familiar with the need for privacy in processing medical information, we would like to caution that anonymizing data is not easy: as Dr. Latanya Sweeney – now from Harvard University -- has shown, often, even when we strip off all obvious identifying markers, we can still, contrary to our intensions, be able to uniquely identify a person. Some examples of such unintended identification are provided on identifiability webpage https://dataprivacylab.org/projects/identifiability/index.html maintained by Dr. Sweeney. This webpage has links to her 2000 report https://dataprivacylab.org/projects/identifiability/paper1.pdf and to other webpages describing identifiability issues – in particular, issues related to patient identifiability.

Luc Longpre and Vladik Kreinovich

longpre@utep.edu, vladik@utep.edu

University of Texas at El Paso

El Paso, Texas

May 2, 2018