Abstract

I have been involved in herd improvement for a large cooperative since the late 1980s. The concept of big data has changed considerably over the intervening years. During the late 80s to the mid-90s, herd improvement organizations were considered big-data practitioners with large computational requirements. The big-data label was primarily from maintaining a central database containing over 20 years of animal improvement data. The large computational requirements were due to the animal model and test-day model genetic evaluation systems. Labelling animal improvement organizations as big-data practitioners ceased as e-commence began to grow in the early 2000s. Since the mid-2010s, big data acquisition has been driven by the widespread adoption of artificial intelligence technologies and the biological sciences revolution. Since 2015, Livestock Improvement has been collecting data for genomic selection, milk diagnostics and the use of artificial intelligence technologies on-farm. We now have 100s of TBs of whole-genome sequence data, image data and mid-infrared spectra data. This paper will discuss the acquisition, storage and processing strategies employed to deal with this amount of data and ask the question: What will the big-data future look like for a herd improvement cooperative?