



17. Data Governance How do Organizations Deal with Stewardship of Farmer Data?

Title presentation

Data Ownership, Privacy, Use, Sharing and Stewardship - The U.S. Experience

Author(s)

João Dürr

Institution for which the first author of this abstract is working

Council on Dairy Cattle Breeding

Abstract

The Council on Dairy Cattle Breeding (CDCB) is a non-profit collaboration between four sectors of the U.S. dairy industry: dairy records providers, dairy records processing centers, Purebred Dairy Cattle Association (PDCA) and National Association of Animal Breeders (NAAB). The purpose of CDCB is to host the national cooperators database (NCD) on behalf of the dairy community and use data analytics to provide value back to dairy producers through genetic evaluations and management information. The NCD is composed of three basic data types: animal relationships (pedigrees), management and performance records (phenotypes) and single nucleotide polymorphisms (SNP) markers (genotypes). Phenotypes and pedigrees have been collected over a century by dairy herd improvement (DHI) services contracted by dairy producers to enhance decision-making process at the herd. Dairy producers own and control the use of data generated from their herds, and access to data is regulated by agreements signed with the DHI service providers. Genomic nominators provide tissue sampling and data collection services and serve as conduits between farmers, genotyping laboratories and CDCB for genomic predictions. Control over the use of genomic data is also exerted by animal owners and regulated by commercial agreements between service providers and users. The NCD receives data from a multitude of sources, and CDCB has material license agreements with each one establishing data access and use limits. A material transfer research data exchange agreement between the Agricultural Research Service (ARS) of U.S. Department of Agriculture (USDA) and CDCB allows ARS researchers access to the NCD for research purposes only. Other research organizations can only access data from the NCD if formally authorized by the data controllers. International data sharing initiatives mainly involve genotypes and are negotiated by data controllers and executed by CDCB accordingly. CDCB is also investing in novel phenotypes data generation projects such as feed efficiency, which require a new business model to secure sustainable data flow. Dairy herds data from sensors is growing rapidly in the industry but in a disordered process that lacks standards, quality assurance, and means to be properly integrated to the existing data systems. Dairy herd data belong to dairy producers who invest in technology and pay for all services. Therefore, dairy data must serve primarily to improve the decision-making process at the herd level and the role of farmer based organizations such as CDCB is to facilitate an effective integration of the existing and emerging data streams.