水 研 機 構 研 報, 第 45 号, 33 − 33, 平 成 29 年
Bull. Jap. Fish. Res. Edu. Agen. No. 45, 33 − 33, 2017

Big Data in Agriculture and the USDA/ARS Initiative

Jeffrey SILVERSTEIN*

Abstract: In recent years, technological achievements have permitted the relatively inexpensive and rapid production of vast amounts of data. The large and often complex datasets produced in the scientific sphere demand new approaches to gain value and to turn data into information. Management of the growth in the volume, variety, and velocity of data is often referred to as the 'Big Data problem'. The Agricultural Research Service of the U.S. Department of Agriculture (USDA/ARS) has long been a strong, science-based, problem solving agency. In the past, our computational infrastructure has primarily been based on meeting administrative needs and security requirements, whereas computational, analytical, and sharing of scientific data were regarded as secondary priorities. The Big Data problem has required a reassessment of scientific computing needs. In February 2013, we held a workshop led by ARS scientists to assess scientific data needs, and this resulted in a $25 million initiative to develop the USDA/ARS capacity to collect, share, and analyze Big Data. Our Big Data initiative contains three elements. First, we have developed a dedicated scientific research network (SCInet) which will leverage Internet2 to facilitate large scale transfer of research data at high-speeds with low latency. Second, we have constructed a high-performance computing (HPC) system with high memory, high processing capacity and the potential to burst computational workloads to commercial resources when necessary. SCInet and the HPC have been largely constructed and connected in the first 18 months of the initiative. Finally, we are developing a virtual research support core of individuals who will provide scientific computational and informatics support. These individuals will work with agency scientists to facilitate specific projects and will also develop standardized solutions and training for common challenges. Key steps of this initiative, 1) determining the needs of the agency, 2) developing a plan to connect more than 90 locations in the agency, 3) building the technical capacity in our labs, and 4) controlling costs due to restrictive budgets are challenges that will be shared and discussed.

Key words: Big Data, Internet2, informatics

2017年 2 月28日受理（Received on February 28, 2017）
*1 Southeast Area, US Department of Agriculture, Agricultural Research Service, P.O. Box 225, Stoneville MS 38776, U.S.A.
E-mail: jeff.silverstein@ars.usda.gov