Contratto di Filiera Stalla Modello

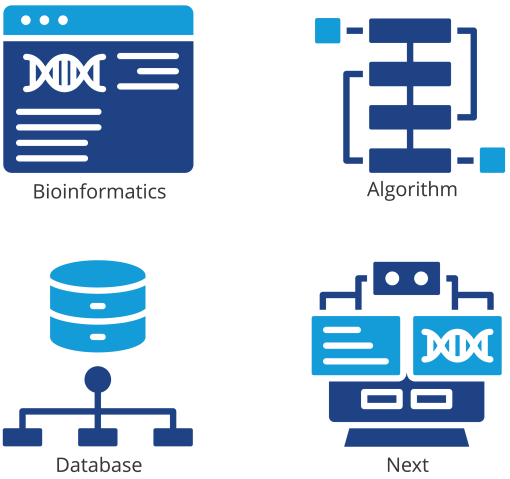
WP2 – OF 2.1 Data storage and data management system implemented and operational to meet the project's needs

Recipient: UNIBO

Authors: Valentina Indio

OF2.1 Description

An integrated data storage and data management system dedicated to handling the data generated within the Stalla Modello project has been implemented and made fully operational. The infrastructure was designed to support the requirements arising from genomic, metagenomic, and metataxonomic sequencing activities involving dairy cattle, milk, dairy products, and the environmental matrices analysed during the project. Given the large volumes of data generated by next-generation sequencing (NGS) technologies, the system was developed to ensure adequate storage capacity, reliability, security, and operational continuity.



The platform is based on servers running the Rocky Linux operating system and is integrated with a centralised NetApp storage infrastructure with an overall capacity of approximately 120 TB. The system enables the structured storage of raw data generated by sequencing runs, intermediate files produced by bioinformatics pipelines, and the final results of genomic and microbiological analyses. It is also directly connected to the Illumina NextSeq 1000 sequencing platform, from which it automatically acquires and stores the data generated during sequencing runs, ensuring uninterrupted data transfer and immediate file availability for subsequent bioinformatics processing.

The infrastructure also includes dedicated systems designed to enhance storage protection and reliability, based on Broadcom/LSI MegaRAID RAID controllers, thereby minimising the risk of data loss. Dataset synchronisation and replication tools based on rsync have also been implemented to support the secure transfer and alignment of data across the different components of the infrastructure.

The implemented system therefore serves as the central data management platform for the VeLaSeq genomics facility and represents a strategic infrastructure for the collection, organisation, preservation, and exploitation of the genomic data generated. The availability of a dedicated platform also promotes the standardisation of data management procedures and supports the research and development activities planned within the project.

Contacts:
Alessandra De Cesare
alessandra.decesare@unibo.it
Tel. +39 051 2097583
Cell. +39 331 1369358

Valentina Indio
valentina.indio2@unibo.it
Tel. 051 2097516


VeLaSeq