Table of Contents

Giacomo Domeniconi's Projects ................................................................. 3
Giacomo Domeniconi's Projects

- **GenData 2020**: Prevention, treatment, management and cure of diseases are all underpinned by the fundamental understanding of their causes, processes and impacts. The technology for fast DNA sequencing appears as the main innovation factor of the next decade: high throughput devices will soon enable reading the whole genome much faster, at higher resolution, and at lower cost, thereby giving us the data to answer fundamental biological questions and open the ground to personalized genetic medicine. While genetic sequencing is “mature” - future advances will concern the number and length of sequences produced per unit of time or the precision of nucleotide identification - a quantum leap is now needed for building the computing infrastructure at the receiving end of DNA sequencing machines. In particular, current genomic data management is struggling on the “initial” problem of storing the data which are fast produced by biologists in their laboratories. A powerful data infrastructure is required for going beyond pure storage, and enabling viewing, querying, analyzing, mining, and searching over a world-wide available collection of genetic data. The vision of the GenData 2020 project is that it is now possible to build the abstractions, models, and protocols for supporting a network of genomic data, made available by genome servers located in the major biologist laboratories in the world. The huge amount of data and the diversity of the platforms and formats yields to a major data management challenge: how to model and store genetic data so as to foster their accessibility.