**Data Science and Clouds**

We explore the principle that much of “the future” will be characterized by “Using Clouds running Data Analytics processing Big Data to solve problems in X-Informatics”. Applications (values of X) include explicitly already Astronomy, Biology, Biomedicine, Business, Chemistry, Crisis, Energy, Environment, Finance, Health, Intelligence, Lifestyle, Marketing, Medicine, Pathology, Policy, Radar, Security, Sensor, Social, Sustainability, Wealth and Wellness with more fields defined implicitly. We discuss the implications of this concept for education and research. Education requires new curricula – generically called data science – which will be hugely popular due to the many millions of jobs opening up in both “core technology” and within applications where of course there are most opportunities. We discuss possibility of using MOOC’s to jumpstart field. On research side, big data (i.e. large applications) require big (i.e. scalable) algorithms on big infrastructure running robust convenient programming environments. We discuss clustering and information visualization using dimension reduction as examples of scalable algorithms. We compare Message Passing Interface MPI and extensions of MapReduce as the core technology to execute data analytics.