Data Science for the identification and control of thermonuclear plasmas

Chairs: Andrea Murari[1,2] and Riccardo Rossi[3]

1) *Consorzio RFX (CNR, ENEA, INFN, Università di Padova, Acciaierie Venete SpA), Corso Stati Uniti 4, 35127 Padova, Italy.*
2) *Istituto per la Scienza e la Tecnologia dei Plasmi, CNR, Padova, Italy*
3) *Department of Industrial Engineering, University of Rome "Tor Vergata", via del Politecnico 1, Roma, Italy*

The global energy consumption has been steadily rising at a higher rate ever since the 1950s. Given the increase in both the global population and the average standards of living, more energy is needed. Due to the impact of greenhouse gases on climate change, such an increase in energy production cannot be met simply with fossil fuels. In the medium and long term, thermonuclear fusion could potentially become a very important ingredient in a sustainable energy mix on a planetary scale. This approach to electricity generation is based on coalescing the nuclei of hydrogen isotopes; the resulting defect of mass translates in very efficient energy production.

Unfortunately, the conditions for economically viable energy production by thermonuclear fusion are not easy to achieve in the laboratory. The most promising route remains magnetic confinement, whose best configuration is the tokamak. Tokamak plasmas are very complex systems, from both a technological and physical point of view. The next generation of these devices will have to make a much more systematic use of feedback control than present day machines. The higher energy content, longer pulses and increased sophistication of the magnetic configurations are indeed all elements expected to render traditional feed forward schemes highly inadequate.

In this context, successful real time strategies will have to rely on the proper identification of the plasma and on the detailed understanding of the main instabilities, which have the potential to seriously affect the evolution of the discharges or the integrity of the devices. The relevant information will have to be extracted by the Gigabytes of data produced per second by these machines. The previous considerations have motivated an increasing interest in new data driven techniques to identify reactor relevant plasmas, to solve the required inversion problems and to plan new experiments. With regard to the instabilities, a series of causality detection and correlation tools are being developed, particularly for predicting their evolution and for controlling their severity. To achieve these goals, all the main machine-learning techniques are being deployed, from neural computation to genetic programming and decision trees.