The Human Brain Project

It is really a data integration and analytics problem – not a data generation problem…

Thomas Lippert (Leader SP7: HPC Platform)
Boris Orth (SP 7 Project Manager)
Bernd Mohr (Task Leader T 7.2.4)

Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH
Jülich, Germany

The Human Brain Project

The goal of the Human Brain Project (HBP) is to gather all existing knowledge about the human brain and to reconstruct the brain, piece by piece, in multi-scale models and supercomputer-based simulations of these models. The resulting "virtual brain" offers the prospect of a fundamentally new understanding of the human brain and its diseases, as well as novel, brain-like computing technologies.

The HBP is developing six Information & Communication Technology (ICT) Platforms, dedicated respectively to Neuroinformatics, Brain Simulation, High-Performance Computing (HPC), Medical Informatics, Neuromorphic Computing and Neurorobotics.

In October 2013, the European Commission (EC) started supporting this vision through its Future & Emerging Technologies (FET) Flagship Initiative. The HBP's 2½ year Ramp-Up Phase, which will last until March 2016, is funded by the EU's 7th Framework Programme (FP7). This phase should be followed by a partially overlapping Operational Phase, which will be supported under the next Framework Programme, Horizon 2020.

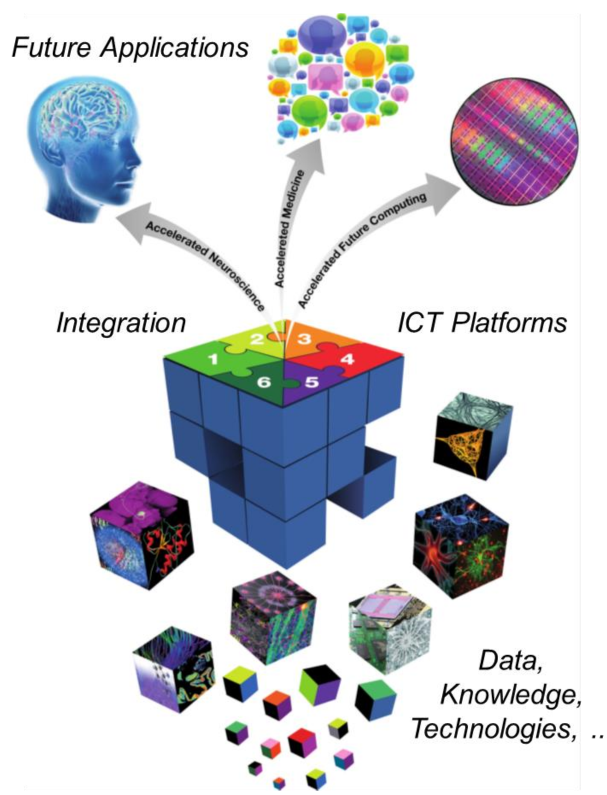

Figure 1: Goals of HBP

The HBP as a whole is planned to be implemented in three phases, spread over ten years, with an estimated total budget of more than EUR 1 billion. The project, which is coordinated by the Ecole Polytechnique Fédérale de Lausanne (EPFL) in Switzerland, already brings together 80 European and international research institutions. In 2014, more partners should join the consortium via the HBP's Competitive Call Programme.

The High Performance Computing Platform Subproject

The HPC Platform Subproject (SP 7) is one of the HBP's 12 operational subprojects. Its mission is to build and manage the hardware and software for the supercomputing and data infrastructure required to run cellular brain model simulations up to the size of a full human brain. SP 7 will make this infrastructure available to the consortium and the scientific community worldwide.

Central to the HPC Platform is the HBP Supercomputer, the project's main production system, to be located at Jülich Supercomputing Centre. The HBP Supercomputer will be built in stages, with an intermediate "pre-exascale" system of the order of 50 PFLOPs planned for the 2016-17 timeframe. Full brain simulations are expected to require exascale capabilities, which, according to most potential suppliers' roadmaps, are likely to be available in approximately 2021-22. As well as providing sufficient computing performance, the HBP Supercomputer will also need to support data-intensive interactive supercomputing and large memory footprints. Besides the HBP HPC main production system in Jülich, there will be Software Development System at CSCS, Switzerland, a Subcellular Computing System at BSC, Spain, and a Data Analytics System at Cineca, Italy.

While exascale supercomputers will become available sooner or later without any HBP intervention, it is unlikely that these future systems will meet unique HBP requirements without additional research and development (R&D). This includes topics like tightly integrated visualization, analytics, and simulation capabilities, efficient European-wide data management, dynamic resource management providing co-scheduling of heterogeneous resources, and a significant enlargement of memory capacity based on power-efficient memory technologies. In line with the EC's established HPC strategy, Forschungs-zentrum Jülich will therefore drive R&D for innovative HPC technologies that meet the specific requirements of the HBP.

The HBP Data Integration Strategy

The HBP's ICT Platforms will make it possible to federate neuroscience data from all over the world, to integrate the data in unifying models and simulations of the brain, to validate the results against empirical data from biology and medicine, and to make them available to the world scientific community (see Figure 2). The HBP will consume existing data coming from publications, models, experiments, and patient data stored at hospitals and insurances, but will also generate new data resulting from neuroscience simulations and their analysis.

The resulting knowledge on the structure and connectivity of the brain will open up new perspectives for the development of "neuromorphic" computing systems incorporating unique characteristics of the brain such as energy-efficiency, fault-tolerance and the ability to learn. The HBP's models and simulations will also enable researchers to carry out in silico experiments on the human brain that cannot be done in vivo for practical or ethical reasons.

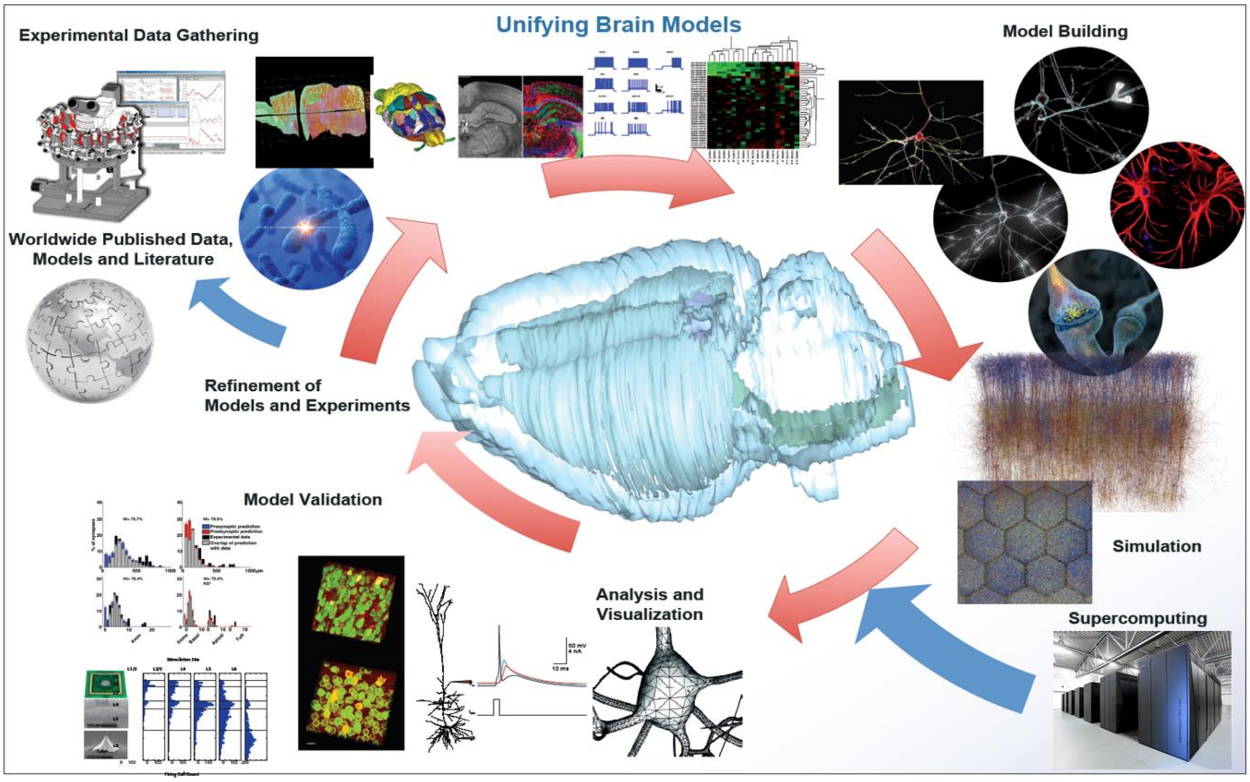

Figure 2: The HBP Data Integration Strategy [*F. Schürmann, EPFL*]

As one of many examples of the data volumes to be handled by the HBP, consider BigBrain, an Ultrahigh-Resolution 3D Human Brain Model derived from experimental data [Amunts2013]. The volume of a human cerebral cortex is ~7500 times larger than a mouse cortex, and the amount of white matter is 53,000 times larger in humans than in mice. A recently published data set of the digitized mouse brain with 1-mm resolution has a total amount of uncompressed volume data of 8 TByte [Li2010]. The creation of a volume with similar spatial resolution for the human brain would result in ~21,000 TByte. The interactive exploration (as opposed to simple storage) of such a data set is beyond the capacities of current computing. Thus, among other methodological problems, data processing becomes a major challenge for any project aiming at the reconstruction of a human brain at cellular resolution.

References

[web2013] https://www.humanbrainproject.eu/

[HBP2012] The Human Brain Project, A Report to the European Commission, https://www.humanbrainproject.eu/documents/10180/17648/TheHBPReport_LR.pdf

[Amunts2013] BigBrain: An Ultrahigh-Resolution 3D Human Brain Model, Katrin Amunts et al., Science 340, 1472 (2013)

[Li2010] A. Li et al., Science 330, 1404 (2010)