

THE MEETING OF INFORMATICS WITH BRAIN MEDICINE: A VISION FOR THE FUTURE FROM THE PERSPECTIVE OF THE HUMAN BRAIN PROJECT

Richard Frackowiak – HBP & EPFL, Lausanne

We now know that a single human gene mutation may present with any of multiple phenotypes, and *vice versa*, that a range of genetic abnormalities may cause a single disease phenotype. These observations lead to the conclusion that a deeper understanding is needed of the way changes at one spatial or temporal level of brain organisation integrate and translate into others, eventually resulting in behaviour and cognition or their abnormalities.

The **basic idea** is that it is now possible to look for rules underlying the functional and structural organisation of the human brain, exhaustively, at all spatial scales, and eventually perhaps at all spatio-temporal scales. The **methodological approach** is to federate and integrate existing knowledge from bottom up using recent advances in information technologies, notably supercomputing and distributed and interactive data basing. The **theory** is that rules and constraints determining a particular structural and functional organisation at one level will limit what organisational principles are possible at the next. It has, for example, been shown that one can construct *in silico* models that look and behave remarkably like their *ex vivo* counterparts, up to the level of the cortical mini-column. The **ambition** therefore is to link genetic and proteomic levels by determining the rules that govern the segregation of protein expression. From protein expression we can start to extract rules that determine cellular morphology, which in turn predicts connectivity, and so on, until the mechanisms of emergent properties are discovered by a constructive process of reconstruction and predictive simulation, not as isolated modules but as interacting biological entities.

Materials scientists discovered that the apparent simplicity of matter emerges from complex statistical combinations of subatomic particles and waves and have generated a rich theory to situate and interpret all new experimental results. They have done so by means of systematic deconstruction based on experimental hypothesis refutation – the classical reductionist approach. So it is perhaps unsurprising that in the case of the human brain, a similar approach has generated a mass of knowledge but no overarching theory explaining emergent behaviour. The response of brain scientists has been to generate more knowledge, hoping that a theory will eventually reveal itself. We argue that the degree of complexity represented by organic matter warrants an additional, complementary set of tools – tools that federate, integrate mine and simulate those data, thus discovering new patterns and rules that govern its organisation. The application of such tools is justified, we think, by one simple fact: every organic object, however complex, is generated from just four building blocks, the four base pairs of DNA.

The HBP's goal is to generate a draft blueprint that describes how the brain is constructed across all levels. The blueprint will provide a framework within which

new and old theories can be tested and new hypotheses generated. An immediate aim is to create a federation of hospital databases, whilst respecting the imperatives of security and privacy of individuals. To this end the Medical Informatics Platform of the Human Brain Project brings together scientists and engineers from multiple domains including computer science, informatics, statistics, mathematics, clinical science, neurologists, psychiatrists, neuro-imagers and the like. We have specified and built together software that provides an Open Source distributed platform for use by European and other clinician scientists, population health scientists, epidemiologists, health economists and the like. All these and other health care related disciplines should in the long run be able to benefit from the very large, constantly accruing masses of clinical data that reside, currently grossly underused, in hospital databases. The federation will be distributed, eschewing the need for data warehousing and guaranteeing local control of access to data by every hospital, as presently

Another critical medicine-related research aim of the federated platform will be to use big data techniques and algorithms to fill the multi-dimensional brain disease space, which ranges from psychiatry through behavioural disorders to neurological diseases, with groups of like patients characterised by combinations of homogeneous features. We call such a group of specific features a “disease signature”. A disease signature brings together, in a systematic way, clinical features (phenomenology) and the results of genetic, biological, physiological and anatomical test results (biology). This reclassification should in the long-term supplant the symptom and syndrome based DSM and ICD disease catalogues by ones based on the same clinical features supplemented by patterns of abnormal investigations. This strategy has as its primary aim the definition of more precise diagnoses that reduce error variance in diagnostic categories for construction of clinical trial cohorts. A biological underpinning of phenomenological features may in combination with improved understanding of the functional and structural organisation of the human brain emanating from other parts of the Human Brain Project facilitate other aspects of brain medicine, such as identification of biological treatment targets, planning of potential benefits, risks and side effects of proposed treatments, as well as considerable reduction of clinical cohort sizes associated with more precise disease signature based definition of diagnosis and prognosis.