The neuroethical challenges of brain simulations

Christen, Markus

Originally published at:
Christen, Markus (2013). The neuroethical challenges of brain simulations. In: Meeting of the International Association for Computing and Philosophy, University of Maryland at College Park, 15 July 2013 - 17 July 2013, 1-9.

The Neuroethical Challenges of Brain Simulations

Markus Christen

Institute of Biomedical Ethics, University of Zurich, Switzerland

Introduction

In January 2013, the *Human Brain Project* was among the two scientific "Flagships" selected by the European Commission aiming to "provide a strong and broad basis for future technological innovation and economic exploitation" (FET). The project should receive up to 1 billion euros in 10 years and "should lay the technical foundations for a new model of ICT-based brain research, driving integration between data and knowledge from different disciplines, and catalysing a community effort to achieve a new understanding of the brain, new treatments for brain disease and new brain-like computing technologies" (HBP Report 2012, 3). Just a few days later, US-president Barak Obama announced a related initiative for the US, aiming to invest several billion dollars to examine the workings of the human brain (Markoff 2013). Both initiatives exemplify the huge transformation of various scientific fields that rely increasingly on computer power not only to organize data, but to generate new knowledge (Winsberg 2010). Prominent examples are climate research and cosmology. Given the historical interconnection between brain research and computer technology (see below), it is of no surprise that neuroscience uses the possibilities of today's enormous computing capacities to deal with fundamental questions of their fields.

This contribution is not about discussing the epistemological consequences of *in silico* experiments in neuroscience, which – as a general topic – is increasingly analyzed in philosophy of science and science studies (e.g., DeLanda 2011, Gramelsberger 2010). I will also not comment on the controversy whether and to what extent simulation approaches allow solving problems in neuroscience and on the specific problems large-scale funding initiatives may have on scientific practice. Rather, I discuss the ethical consequences when the brain is object of simulation approaches. This not only refers to the simulations per se, but also to the underlying restructuring of knowledge organization that accompanies such huge modeling approaches. As the initiators of the *Human Brain Project* have pointed out, "the major obstacle that hinders our understanding of the brain is the fragmentation of brain research and the data it produces" (HBP Report 2012, 3). Thus, "simulating the brain" is not only a matter of the technology and practices that are directly involved in creating simulation code, building and using the models. It also involves sophisticated processes to organize empirical and theoretical knowledge that should inform the models. The HBP promoters write: "We propose that the HBP use these techniques to generate a scaffold of *strategically selected data* [my emphasis] on the structure and functional organization of the human brain at different ages and at different levels of biological organization" (HBP Report 2012, 30). This indicates the necessity to *choose* among potentially conflicting data, which involves an important, but hidden normativity in model generation.

In the following, I discuss two claims: First, I suggest that several ethical challenges that large-scale simulations and knowledge organization within neuroscience will have, are currently not sufficiently addressed in the neuroethics community. Second, I predict that brain simulations will become equally political influential as climate models both with respect to guiding research investment allocation as well as to inform political decision making. Thus, lessons learned in climate modeling with respect to the practice of model building will be important for large-scale brain simulations.

Some historical remarks

Computers have become an indispensable tool for simulations[1] in science. In neuroscience, however, the relation between the tool for simulation and the object of simulation is bidirectional. Brain and computer were in a liaison since the conceptualization of information (Aspray 1985) and the emergence of early computers. Protagonists of the information age like Alan Turing or John von Neumann were inspired by brains when suggesting computation principles and architectures (e.g., Von Neumann 1958). For some time, during the climax of cybernetics, the computer was a model for understanding biological information processing (which is evidently no longer the case) – and the brain with its impressive energy efficiency was and still is an inspiration for building "new" computers in the context of so-called "neuromorphic engineering". Brain modelers like Chris Eliasmith (Spaun model; Eliasmith et al. 2012) have announced to make increasingly use of neuromorphic computer chips for advance their models (personal communication; February 23rd 2013). Thus, (future) brain models will not only be implemented by software programs running on (super-) computers, they may also include physical realizations of some principles of biological information processing (e.g., a combination of analog and digital technology).

The interconnection of simulation tool and simulation object and the importance of the information metaphor in neuroscience raise various questions, some of which already are discussed in the literature (e.g., Bennett & Hacker 2003; Garson 2003; Falkenburg 2012). Those epistemic issues like the meaning of the concept of 'information' in neuroscience are indeed important and should be investigated further in the context of the general discussion on the role of *in silico* experiments in research – they are, however, beyond the scope of this contribution. But it is worthwhile to mention a second aspect, the consequences of the "informatization" of neuroscience on the scientific practice. In contrast to molecular biology, where notions like the 'genetic code' became major orientations for the scientists in the 1950s and 1960s despite the vagueness of the term (Kay 2000), the first attempt to grasp neurobiological processes in terms of information theory and computation failed to become a productive orientation for research (Christen 2006). This may partly explain why there is still a profound skepticism within neuroscience with respect to simulation approaches – in particular if they try to grasp brains as a whole. However, there is bibliometric evidence that the relative importance of computational approaches within neuroscience grows (Christen 2006; see also next section), probably reflecting a change in education and training of (future) neuroscientists that, e.g., have backgrounds in physics and computational sciences.

[1] To clarify the terminology used in this contribution: Models are abstractions of real-world structures and/or processes mostly in the form of mathematical equations or algorithms (although some models are physical, e.g. in hydrology). Simulations refer to the behavior of the model in time, whereas the equations or algorithms are usually implemented on a computer, requiring in most cases numerical approximations. Simulations may specify inputs, information handling mechanisms, or outputs in order to allow for prediction, retrodiction, explanation or exploration. Due to the numerical nature of most simulation calculations, simulations can be understood as approximations of models.

There is a third lesson from this excursion into the history of the brain computer liaison. It refers to the idea that brain simulation may in the future reach a degree of complexity such that the simulation mimics brain functions that are considered to express deeply human competences like consciousness, imagination or moral concern. This is a well-known topos in science fiction literature and film, mostly in the sense that the "machine brains" endanger humanity – and it also has accompanied research communication on the intersection of neuroscience and computing ever since. The ethical importance of that point refers to the fact that, by referring to this possibility, brain simulations gain attractiveness as they allow a reference to various deep philosophical problems.[2] It is tempting for neuroethicists to use the field of brain modeling as a "playing field" for (re-)discussion these topics. The problem, however, I see here is less the fact that such advanced brain simulations are still highly speculative, but that these discussions may cover more important normative issues that refer to the methodology of simulations.

The need for brain simulations

Given the incredible complexity of the human brain that involves the interplay of various organizational levels (molecules, processes within cells, cell networks, networks of brain regions, connections of the brain to its sensory, motor and metabolic periphery, behavior of organisms with brains), it would be tempting to state that brain simulations will just not be able to deliver the promised results, thus diminishing their ethical relevance. But this opinion misses the point in two ways. First, it is based on a wrong appreciation of what (future) brain models really are. It's not primarily about simulating a specific brain process or function; it's about structuring knowledge in a specific way. Second, it's insufficient to focus the ethical analysis on the potential results of brain simulations, i.e. to make the relevance of the discussion dependent on the probability that these results actually can be achieved. The social process of generation and using simulations has normative implications, too.

It is undisputed that models have an important role in neuroscience (Gerstner et al. 2012). Mathematical and computational approaches in neuroscience have a long tradition that can be followed back to early mathematical theories of perception and of current integration by a neuronal cell membrane (e.g. the Hodgkin-Huxley model of neuronal spike generation). Their role was traditional in the sense that models and simulations were instrument to sharpen the understanding of a specific phenomenon, e.g. which mechanism captures a relevant phenomenon measured *in vivo*. There is a huge spectrum of models in neuroscience; and also on the level of large-scale brain simulations, the methodologies, aims and neurobiological fidelity differ (De Garis et al. 2010). Examples include the "Blue Brain" (Markram 2006), the "SyNAPSE project" (Systems of Neuromorphic Adaptive Plastic Scalable Electronics; Ananthanarayanan et al. 2009), a large-scale model of the mammalian thalamocortical systems (Izhikevich & Edelman 2008), and the Spaun-model (Semantic Pointer Architecture Unified Network; Eliasmith et al. 2012).

There are certainly controversies with respect to the relevance of these models for particular neuroscientific research questions. Those controversies, however, may hide that models and simulations more and more obtain new functions within neuroscience, namely as being the most

[2] For example, the authors of the HBP report write: "Whatever the results of the project, they would profoundly influence current beliefs about the human mind, identity, personhood, and our capacity for control" (HBP Report 2012, 93).

promising instrument to integrate knowledge gained on all levels of neuronal organization. It is indeed hard to imagine how neuroscience would be able to address the "big" questions of its field without this integrative perspective – an insight that in social neuroscience lead to the notion of the "multi-level analysis" (Cacioppo & Decety 2011). Thus, it seems inevitable that models and simulations – both with respect to specific brain processes as well as a tool to organize knowledge – will play a central role in neuroscience. This is indeed the major aim of the *Human Brain Project*, as Markram points out: "We are not building a model; we are building a data integration strategy to render biologically realistic models. Blue Brain is a strategy to generically build brain models" (personal communication, February 21st 2013; for Christen 2013).

There are three major points that have to be emphasized here: First, models understood in that way will become a *predictive tool*, i.e. they are used to guide what kind of experimental measurements should be obtained in order to test theories. Second, models will become a *communication tool*, i.e. they generate a new type of evidence (visualizations, movies) that is both relevant in communication between scientists working on the various levels of brain organization as well as for informing the public. Third, modeling of this type requires a *knowledge model*, i.e. a structured access to data and data interpretations across all levels that will, due to the enormous number of publications in neuroscience,[3] indispensably rely on automatized procedures of text mining and the like. As I will show now, underestimated ethical challenges of brain simulations will refer to those three novel roles of simulations within neuroscience.

The ethical challenges of brain simulations

The ethical assessment of large-scale research projects is traditionally output-oriented, i.e. one analyses benefits and risks of potential results weighted with their likelihood of occurrence using some semi-structured normative reference scheme (e.g. principlism; Beauchamp & Childress 2012). A paradigmatic example is the Human Genome Project's Ethical, Legal and Social Issues (ELSI) program. Almost all of the more than 190 project financed in the ELSI frame had this orientation, referring e.g. to handling of genetic information, bench-to-bedside issues, or informed consent (Meslin et al. 1997, ELSI 2000).

Within ethics, it is certainly important to deal with potential consequences that the results of scientific endeavors may have. For example, referring to the Human Brain Project, advances in neuromorphic computing and neurorobotics could allow for a higher degree of machine autonomy that may challenge our notion of responsibility (Christen 2004). But a focus on negative (and positive) consequences of the result of research is incomplete due to two reasons: First, evaluating the ethical relevance of the results requires an understanding of their genesis and the (often hidden) normative decisions that have been made in that process. This may also include a proactive approach along the notion of value-sensitive design (Friedman 1997). Second, the research process itself may have side-effects that are not directly related to the intended goal of the project, e.g. with respect to funding allocation, training, "working philosophies" and the like. Those side-effects can be of normative importance, too.

---

[3] The promoters of the Human Brain Project estimate that the "publication body" relevant for the project consists of at least 30 million papers (HBP 2012, p. 37)

Referring to the novel role of simulations within neuroscience mentioned above, there are three ethical challenges for brain simulation. First, predictive models are an important extension of the classical hypothetic-deductive approach in science, because the deduction relies on a very complex process that may not anymore be understandable for the experimenter conducting the empirical tests. When simulations are guiding experiments, they also can misguide them, i.e. it will be essential to enable for close and intense collaboration between modelers and those who use them.

Second, simulations as communication tools rely on conventions how to visualize the output generated by the simulation. It is crucial to understand that the step from the stream of numbers to the fancy pictures or movies that visualize the simulation result is accompanied by various decisions, allowing e.g. whether or not one can distinguish simulation results from empirical measurements.

Third, the built-up of knowledge models, i.e. structured access to data and publications referring to the phenomena one wants to model, is connected to normative decisions – namely with respect to what should be included in these knowledge libraries and what not. This selection procedure is different compared to the traditional one relying on peer review, as latter allows for contradicting knowledge. But when creating models, at some time one has to choose which mechanism one wants to build in and which one not. If this concerns only one or very few mechanisms, this may be unproblematic. But the more complex the models are and the more one has to rely on techniques like parametrization for making, e.g., the model feasible with respect to computation time, the more such decisions are necessary. Thus, there will be an incentive to "clean" the knowledge base from conflicting data, requiring a careful governance of building up the knowledge base for such large-scale simulations. This problem is aggravated when "confirmed" knowledge is questioned again – but when this knowledge is already deeply implemented in simulation code, the effort for change is large, resulting in a temptation to neglect this discrepancy.

Experiences from climate modeling show that these concerns are not merely of a theoretical kind and that they do have ethical consequences. A study of Lahsen (2005) identified several pitfalls of climate modeling that are relevant for us. First, with respect to collaboration among modelers and empirical scientists, it was found that model developers typically are also model users. Because of the complexity of the models and of the phenomena they seek to represent, model developers build only parts of a model, integrating sub-models and representational schemes ('parameters') developed by other modeling groups. Even scientists ('model users') who are not primarily model developers typically modify the models they have obtained from elsewhere. This difficulty with distinguishing developers from users also complicates clear identification of the exact site of production. This increased specialization has reduced the amount of time model developers have to study the atmosphere by means of empirical data. In the meantime, it has also been observed that the empiricists whose role is checking models against empirical knowledge have been alienated from the models. Empiricists live in a culture that also involves humility about the accuracy of forecasts of atmospheric conditions, which they trace to experiences of regularly seeing synoptic and numerical weather forecasts proven wrong. They complain that model developers often freeze others out and tend to be resistant to critical input, living in a 'fortress mentality'.

Second, visualization indeed matters, as it has often been observed that (e.g. at conferences) it was unclear whether overhead charts and figures were based on observations or simulations.

This confusion of simulations with real data within the atmospheric sciences may be part of a more general phenomenon: similar conflation of simulations with 'observations', 'samples', and 'data' has been identified in studies of scientists in other fields of research (Dowling 1999). Simulation techniques may especially encourage such conflation, however. For example, Stefan Helmreich's ethnographic study of artificial life simulators (1998) revealed the powerful effect of simulations on the imagination of their creators and users.

Third, the psychological and social investment in models and the social worlds of which the modelers are a part can reduce their critical distance from their own creations. Although such personal and professional investments are not unique to the field of modeling, the crucial role of experiments to validate models aggravates failures in critical distance, in particular when models obtain roles in predictions and as "guiders" for experiments. It is thus probable that this can have effects on building up the knowledge base that accompanies the modeling process. This problem is aggravated by the finding that, already in a time where codes were simpler, model codes are seldom subjected to peer review (Bankes 1993) and large-scale model studies are never replicated in their entirety by other scientists, because this would require them to re-implement the identical conceptual models. Replication in science is generally difficult (Collins & Pinch, 1993), and in the field of climate modeling, the exact reproduction of a climate model outcome will never happen due to the 'internal model variability' that results in chaotic dynamic perturbations. The nearest climate models come to close scrutiny of their subcomponents is in the comparison of international peer reviewed studies and a variability of models that converge in their findings.

Conclusion

In summary, the comparison with the experiences in climate modeling indeed shows that collaboration between modelers and empirical scientists are tricky, that visualizations tend to blur important differences and that various psychological mechanisms are at work that may undermine the critical function of the knowledge base that underlies the modeling process. It is thus not surprising that some of the critique raised against the political implications of climate modeling relied on these issues, e.g. with respect to the "climate-gate" controversy. There is admittedly much more to say on this debate – but the point here is merely that the practice of modeling involves various critical issues that have the potential to undermine the function of models and simulations, in particular when they obtain political relevance. Given the enormous burden brain-related diseases have, it is likely that brain simulations will obtain such a political role, for example with respect to guide resource allocation for research in neurodegenerative diseases. It may even be possible that – in combination with approaches in personalized medicine – future brain simulations guide therapy decisions in individual patients, making the ethical impact immediate.

It's important to mention here that these problems can be addressed, but that in particular the community of neuroethics, whose task is to critically accompany these new developments, may lack the competences in doing so. This relates to the fact that most promoters of neuroethics have a background in medicine and – as a member of the community since several years – thus focus on issues that relate to primarily medical problems like enhancement, incidental findings, or side effects of neurological interventions. But addressing the ethical challenges of brain simulation will require experts that grew up in a quite different culture shaped by information technology and physics. Thus, the challenges of brain simulations are also a challenge for those, whose job it will be helping to avoid the ethical pitfalls in a methodology that will transform neuroscience significantly.

References

Ananthanarayanan R, Esser SK, Simon HD, Modha DS (2009): The Cat is Out of the Bag: Cortical Simulations with 109 Neurons, 1013 Synapses. in Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis 2009, ACM: Portland, Oregon. p. 1-12.

Aspray W (1985): The scientific conceptualization of information: a survey. Annals of the History of Computing 7(2): 117-140

Bankes S (1993): Exploratory Modeling for Policy Analysis. Operations Research 41(3): 435–49.

Beauchamp TL, Childress JF (2012): The Principles of Biomedical Ethics. Oxford, Oxford University Press.

Bennett MR, Hacker PMS (2003): Philosophical foundations in neuroscience. Blackwell Publishing, Malden MA.

Cacioppo JT, Decety J (2011): Social neuroscience: challenges and opportunities in the study of complex behavior. Ann. N.Y. Acad. Sci. 1224: 162–173

Christen M (2013): Gehirn-Simulationen – ein hindernisreicher Erkenntnisweg. NZZ, 27.02.2013.

Christen M (2006): The role of spike patterns in neuronal information processing. A historically embedded conceptual clarification. ETH-Diss No. 16464. Reprinted 2012, Südwestdeutschen Verlag für Hochschulschriften.

Christen M (2004): Schuldige Maschinen? Autonome Systeme als Herausforderung für das Konzept der Verantwortung. Jahrbuch für Wissenschaft und Ethik 9: 163-191.

Collins H, Pinch T (1993): The Golem: What Everyone Should Know About Science. Cambridge: Cambridge University Press

De Garis H, Shuo C, Goertzel B, Ruiting L (2010):A world survey of artificial brain projects, Part I: Large-scale brain simulations. Neurocomputing 74, 3–29

DeLanda M (2010): Philosophy and Simulation. The emergence of Synthetic Reason. London: Continuum.

Dowling D (1999): Experimenting on Theories. Science in Context 12(2): 261–73

Eliasmith C, Stewart TC, Choo X, Bekolay T, DeWolf T, Tang C, Rasmussen D (2012): A Large-Scale Model of the Functioning Brain. Science 338, 1202-1205.

ELSI (2000): A Review and Analysis of the Ethical, Legal, and Social Implications (ELSI) Research Programs at the National Institutes of Health and the Department of Energy. Available via: http://www.genome.gov/Pages/Research/DER/ELSI/erpeg_report.pdf (accessed on March 7th 2013)

Falkenburg B (2012): Mythos Determinismus. Wieviel erklärt uns die Hirnforschung? Springer, Heidelberg.

FET Flagship Initiatives: http://cordis.europa.eu/fp7/ict/programme/fet/flagship/home_en.html (accessed on March 8th 2013).

Friedman B (Ed.) (1997): Human values and the design of computer technology. Cambridge University Press, New York.

Garson J (2003): The introduction of information into neurobiology. Philosophy of Science 70: 926-936.

Gerstner W, Sprekeler H, Deco G (2012): Theory and Simulation in Neuroscience. Science 338, 60-65.

Gramelsberger G (2010): Computerexperimente. Zum Wandel der Wissenschaft im Zeitalter des Computers. Bielefeld: Transcript Verlag.

HBP Report (2012): The Human Brain Project. A Report to the European Commission. Available at: http://www.humanbrainproject.eu/files/HBP_flagship.pdf (accessed on February 14th 2013)

Helmreich S (1998): Silicon Second Nature: Culturing Artificial Life in a Digital World. Berkeley, CA: University of California Press

Izhikevich EM, Edelman GM (2008): Large-scale model of mammalian thalamocortical systems. PNAS 105(9), 3593–3598

Kay LE (2000): Who wrote the book if life? A history of the genetic code. Stanford University Press, Stanford.

Lahsen M (2005): Seductive Simulations? Uncertainty Distribution Around Climate Models. Social Studies of Science 35: 895–922

Markoff J (2013): Obama Seeking to Boost Study of Human Brain. New York Times, February 17th 2013 (http://www.nytimes.com/2013/02/18/science/project-seeks-to-build-map-of-human-brain.html?pagewanted=all&_r=0, accessed on March 8th 2013)

Markram H (2006): The Blue Brain Project. Nature Reviews: Neuroscience 7, 153-160

Meslin EM, Thomson EJ, Boyer JT (1997): Bioethics Inside the Beltway. The Ethical, Legal, and Social Implications Research Program at the National Human Genome Research Institute. Kennedy Institute Ethics Journal 7(3): 291-298.

Von Neumann J (1958): The computer and the brain. Yale University Press, New Haven, London.

Winsberg E (2010): Science in the age of computer simulation. Chicago: University of Chicago Press.