Human Brain Project

EBRAINS

Social, ethical & reflective work in the Human Brain Project

Scientific publications contributing to responsible neuroscience: 2013-2023

Preamble

This brochure is a compilation listing articles, books, and reports produced by the researchers working on ethical, and social issues within the Human Brain Project (HBP, the European Community Flagship Project of Information and Computing Technologies (ICT) within the framework of the Future Emerging Technologies (FET) vision). These publications start from the Ramp-up phase launched in October 2013 and cover until the end of SGA3, September 30, 2023.

The HBP is characterised by a profound and broad interdisciplinarity, and it is unique among the large international Brain Initiatives in having included social sciences and humanities in the core research from its very beginning devoting approximately 5% of the budget to this end.

The Ethics & Society research in the HBP has likewise been profoundly interdisciplinary, combining social science and foresight studies, philosophy and neuroethics, public engagement, community building and ethics support, in close collaboration with neuroscience and ICT. Our work has integrated theoretical research with practical applications and has resulted in new methodologies in neuroethics as well as in responsible research and innovation practices (RRI). In addition to scientific publications, we have jointly produced three documents analysing specific topics particularly relevant to the HBP. The goal is for each of those documents, known as Opinions, to be followed by an Action Plan to facilitate their implementation. At present we have the Opinion and Action Plan on 'Data Protection and Privacy (2017); the Opinion on 'Responsible Dual Use'- Political, Security, Intelligence and Military Research of Concern in Neuroscience and Neurotechnology (2018); and the Opinion on 'Trust and Transparency in Artificial Intelligence' (2020). To finalise, we produced a forthcoming anthology "Including society and ethics in brain science" with short reflections on our work over the course of the HBP.

The HBP has been restructured as it entered its final phase (SGA3, 2020-23). The sub-project structure has been dismantled and replaced by a work-package structure. In addition to having established a dedicated RRI work-package wherein neuroethics, foresight, public engagement and ethics support join forces, philosophy, neuroethics, stakeholder engagement and data governance are now embedded as tasks integrated within the different scientific work-packages. This is a sign of success: it could only have happened through the successful collaborative efforts of the project as a whole.

Convinced of the need to promote awareness of the ethical and societal impact of the HBP and of how the project addressed the relevant issues, and mindful of the importance of including a wide range of stakeholders, as well as EU citizens, we who have, from our different perspectives, been working with Ethics & Society issues for the entire duration of the HBP, have decided to give easy access to interested readers to what we have produced during these years, scientifically as well as practically.

We hope that this collection may be useful, notably to those interested in how social science, ethics and philosophy can become driving forces in scientific and technological development and ensure the responsible research and innovation in both theory and practice.

Kathinka Evers & Lise Bitsch

Kathinka Evers, Uppsala University, Professor of Philosophy
Leader, HBP Ethics and Society Subproject (SP12) / Leader HBP
Neuroethics & Philosophy

Lise Bitsch, Ph.D., Danish Board of Technology
Leader WP9 Responsible Research and Innovation

More Ethics & Society resources

In addition to the publications presented in the booklet, we have written about our work on the Ethics Dialogues blog (ethicsdialogues.eu) and Uppsala University's Ethics Blog (ethicsblog.crb.uu.se/tag/human-brain-project). We have developed an Ethics & Society Toolkit with tools for reflecting on ethical and socially responsible practices in brain research (ebrainsethicsandsociety.tekno.dk). In addition, the legacy of the Human Brain Project's commitment to equality, diversity and inclusion in project themes and teams is available in a toolkit that can support projects in integrating equality, diversity and inclusion in their research content and as guiding principles for team collaboration (edi-toolkit.org).

The legacy of our work is available through our online training resources covering a wide range of issues that can all be accessed on the Human Brain Project Website (humanbrainproject.eu/en/science-development/ethics-and-society/ethics-society-training-resources).

Content

BOOKS AND BOOK CHAPTERS 75

Ethics & Society Anthology

Bitsch, L., Salles, A., Evers, K., Changeux, J-P., Stahl, B, Aicardi, C., Burton Datta, S., Mahfoud, T., Reinsborough, M., Rose, N., Klüver, L., Ladegaard, S.F., Alves, E., Nordfalk, F., Bådum, N., Eke, D., Knight, W., Grasenick, K., Romero, P.F.,, Ulnicane, I., Rosemann, A., Ogoh, G., Matar, A., Fernow, J., Bringedal, B., Christen, M., Domingo-Ferrer, J., von Schomberg, R., Illes, J., Rommelfanger, K. (2023). Ethics and Society in Brain Research: Implementing Responsible Research and Innovation (RRI) in the Human Brain Project (HBP). Zenodo. doi: 10.5281/zenodo.7736402
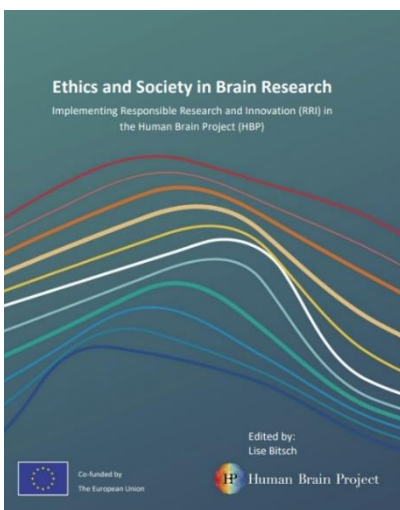
About the anthology: Attention to ethical and social issues were part of the Human Brain Project's work from the very beginning in 2013. Accordingly, a group of HBP researchers from the social sciences and the humanities created several structures and mechanisms and used various conceptual and empirical methods to develop activities and to identify, reflect upon, and manage the ethical and social issues raised by brain research, its outputs, and applications.

With this collection of essays, we aim to present our work in an accessible format, with the ambition of sharing the research and its outputs with diverse stakeholder communities, including policymakers, civil society -and interest organisations, research, and expert communities outside our peer communities. The collection includes short essays by our HBP colleagues who describe and reflect on their work at different stages of our developmental history. In the process, they offer key findings, reflection points, and lessons learned

Opinions and Action Papers

Abstract: The Ethics and Society Subproject has developed this Opinion in order to clarify lessons the Human Brain Project (HBP) can draw from the current discussion of artificial intelligence, in particular the social and ethical aspects of AI, and outline areas where it could usefully contribute. The EU and numerous other bodies are promoting and implementing a wide range of policies aimed to ensure that AI is beneficial - that it serves society. The HBP as a leading project bringing together neuroscience and ICT is in an excellent position to contribute to and to benefit from these discussions. This Opinion therefore highlights some key aspects of the discussion, shows its relevance to the HBP and develops a list of six recommendations.

Abstract: Current and newly emerging insights and technologies arising from research in brain sciences increase capabilities to access, assess and affect thought, emotion and behaviour. While much of this research and development is directed towards clinical use, it also has applications in other settings, notably in the political, security, intelligence and military (PSIM) domains. This is often referred to in terms of 'Dual Use'. Many of these potential uses raise important social and ethical questions which demand the attention of all those involved in the research, administration, management and regulation of neuroscience research and related technological developments, including those in information and communication technologies (ICT) and robotics. In this Opinion, we suggest that we can increase our ability to identify which programmes and projects of research, development and innovation are 'of concern' by applying the principles of Responsible Research and Innovation (RRI) to the concept of 'dual use' and distinguishing between 'responsible' and 'irresponsible' systems of research and technological development. We therefore use the term 'dual use research of concern' (DURC) to refer to neuroscience research and technological innovations, and brain inspired developments in information and communication technologies, for use in the political, security, intelligence and military domains, which are either directly of concern because of their potential for use in ways that threaten the peace, health, safety, security and well-being of citizens, or are undertaken without responsible regard to such potential uses. To ensure ongoing attention to these

issues, the Opinion proposes recommendations for the Human Brain Project, the European Union and the wider neuroscience and ICT community.

Salles, A., Stahl, B., Bjaalie, J., Domingo-Ferrer, J., Rose, N., Rainey S., Spranger, T. Opinion and Action Plan on 'Data Protection and Privacy (2017) Ethics and Society, The Human Brain Project. doi: 10.5281/zenodo.4588466

Abstract: A fuller understanding of the human brain, better diagnoses and treatment of brain disorders, as well as the development of new brain-like technologies are all goals of the Human Brain Project (HBP). Realizing these goals requires the collection, storage, curation, and analysis of data of various sorts over extended periods of time. Securing privacy interests and advancing data protection measures are key concerns of the HBP. The HBP needs to comply with national and European data protection legislation and must go beyond existing legal protections and show ethical sensitivity to privacy concerns, even when such concerns fall outside regulatory frameworks. Recommendations made, in this opinion and action plan, includes measures to ensure data protection in data governance structures, adopting a privacy model when anonymizing data, privacy by design in systems development, exploring ICO tools for privacy management and data protection and the promotion of trust and transparency.

Editorials

Stahl BC, Aicardi C, Brooks L, Craigon PJ, Cunden M, Burton SD, de Heaver M, de Saille S, Dolby S, Dowthwaite L, Eke, D, Hughes S, Keene P, Kuh V, Portillo V, Shanley D, Smallman M, Smith M, Stilgoe J, Ulnicane I, Wagner C & Webb H (2023) Assessing Responsible Innovation Training. *Journal of Responsible Technology*. doi: 1016/j.jrt.2023.100063

Abstract: There is broad agreement that one important aspect of responsible innovation (RI) is to provide training on its principles and practices to current and future researchers and innovators, notably including doctoral students. Much less agreement can be observed concerning the question of what this training should consist of, how it should be delivered and how it could be assessed. The increasing institutional embedding of RI leads to calls for the alignment of RI training with training in other subjects. One can therefore observe a push towards the official assessment of RI training, for example in the recent call for proposals for centres for doctoral training by UK Research and Innovation. This editorial article takes its point of departure from the recognition that the RI community will need to react to the call for assessment of RI training. It provides an overview of the background and open questions around RI training and assessment as a background of examples of RI training assessment at doctoral level. There is unlikely to be one right way of assessing RI training across institutions and disciplines, but we expect that the examples provided in

this article can help RI scholars and practitioners orient their training and its assessment in ways that are academically viable as well as supportive of the overall aims of RI.

Ulnicane, I., & Erkkila, T. (2023) Politics and policy of Artificial Intelligence. Review of Policy Research, 40(5). doi: 10.1111/ropr.12574

Abstract: While recent discussions about Artificial Intelligence (AI) as one of the most powerful technologies of our times tend to portray it as a predominantly technical issue, it also has major social, political and cultural implications. So far these have been mostly studied from ethical, legal and economic perspectives, while politics and policy have received less attention. To address this gap, this special issue brings together nine research articles to advance the studies of politics and policy of AI by identifying emerging themes and setting out future research agenda. Diverse but complementary contributions in this special issue speak to five overarching themes: understanding the AI as co-shaped by technology and politics; highlighting the role of ideas in AI politics and policy; examining the distribution of power; interrogating the relationship between novel technology and continuity in politics and policy; and exploring interactions among developments at local, national, regional and global levels. This special issue demonstrates that AI policy is not an apolitical field that can be dealt with just by relying on knowledge and expertise but requires an open debate among alternative views, ideas, values and interests.

Preprints & live papers

Farisco, M,. Formisano, R., Gosseries, O., Kato, Y., Koboyashi, S., Laureys, S,. Lejeune, N., Martial, C., Matar, Amal., Morrisey, A-M., Schnakers, C., Yakufujiang, M., Yamaki, T., Veeramuthu, V., Zandalasini, M., Zasler, N., Magliacano, A., Estraneo, A. International survey on the implementation of the European and American guidelines on disorders of consciousness. Journal of Neurology. (2023) (in press)

Abstract: Diagnostic, prognostic, and therapeutic procedures for patients with prolonged disorders of consciousness (pDoCs) vary significantly across countries and clinical settings, likely due to organizational factors (e.g., research vs. non-academic hospitals), expertise and availability of resources (e.g., financial and human). Two international guidelines, one from the European Academy of Neurology (EAN) and one from the American Academy of Neurology (AAN) in collaboration with the American Congress of Rehabilitation Medicine (ACRM) and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR), were developed to facilitate consistent practice among professionals working with this challenging patient population. While the recommendations of both guidelines agree in principle, it remains an open issue how to implement them into clinical practice in the care pathway for patients with pDoCs. We conducted an online survey to

explore health professional clinical practices related to the management of patients with pDoCs, and compare said practices with selected recommendations from both the guidelines. The survey revealed that while some recommendations are being followed, others are not and/or may require more honing/specificity to enhance their clinical utility. Particular attention should be given to the implementation of a multimodal assessment of residual consciousness, to the detection and treatment of pain, and to the impact of restrictions imposed by COVID-19 pandemics on the involvement of patients' families/representatives.

Stalh, B., Eke, D. The Ethics of ChatGPT - Exploring the Ethical Issues of an Emerging Technology. International Journal of Information Management (2023) doi: 10.1016/j.ijinfomgt.2023.102700

Abstract: This article explores ethical issues raised by generative conversational AI systems like ChatGPT. It applies established approaches for analysing ethics of emerging technologies to undertake a systematic review of possible benefits and concerns. The methodology combines ethical issues identified by Anticipatory Technology Ethics, Ethical Impact Assessment, and Ethical Issues of Emerging ICT Applications with AI-specific issues from the literature. These are applied to analyse ChatGPT's capabilities to produce humanlike text and interact seamlessly. The analysis finds ChatGPT could provide high-level societal and ethical benefits. However, it also raises significant ethical concerns across social justice, individual autonomy, cultural identity, and environmental issues. Key high-impact concerns include responsibility, inclusion, social cohesion, autonomy, safety, bias, accountability, and environmental impacts. While the current discourse focuses narrowly on specific issues such as authorship, this analysis systematically uncovers a broader, more balanced range of ethical issues worthy of attention. Findings are consistent with emerging research and industry priorities on ethics of generative AI. Implications include the need for diverse stakeholder engagement, considering benefits and risks holistically when developing applications, and multi-level policy interventions to promote positive outcomes. Overall, the analysis demonstrates that applying established ethics of technology methodologies can produce a rigorous, comprehensive foundation to guide discourse and action around impactful emerging technologies like ChatGPT. The paper advocates sustaining this broad, balanced ethics perspective as use cases unfold to realize benefits while addressing ethical downsides.

Chen, G., Scherr, F., Maass, W. Data-based large-scale models provide a window into the organization of cortical computations, bioRxiv 2023.04.28.538662; (2023) doi: [10.1101/2023.04.28.538662](10.1101/2023.04.28.538662)

Abstract: Understanding how the brain solves demanding computational tasks is one of the most exciting scientific challenges of our times. So far, recurrently connected artificial neural network models (RANNs) were primarily used to reverse-engineer brain

computations. We show that it is now also feasible to reverse-engineer computations of detailed data-based large-scale models of cortical microcircuits. Furthermore, results of these analyses produce hypotheses that can readily be tested in biological experiments since they clarify from which neurons one needs to record and what type of information can be expected at specific time points during a trial. We apply this approach to a demanding visual processing task that has often been used in mouse experiments. Both the cortical microcircuit model and RANNs can solve this task as well as the mouse. But the resulting network dynamics matches only for the cortical microcircuit model experimental data on the sparseness of network activity and the impact of individual neurons on the network decision. Reverse-engineering of the computation in the cortical microcircuit model suggests that a particular subset of neurons causes a bifurcation of the network dynamics that triggers the network decision. Altogether, our results introduce a new type of neural network model for brain computations.

Farisco, M., Baldassarre, G., Cartoni, E., Leach, A., Petrovici, M.A., Rosemann, A., Salles, A., Stahl, B., van Albada, S. J. A method for the ethical analysis of brain-inspired AI, Artificial Intelligence Review, (2023) arXiv:2305.10938 doi: 10.48550/arXiv.2305.10938

Abstract: Despite its successes, to date Artificial Intelligence (AI) is still characterized by a number of shortcomings with regards to different application domains and goals. These limitations are arguably both conceptual (e.g., related to underlying theoretical models, such as symbolic vs. connectionist), and operational (e.g., related to robustness and ability to generalize). Biologically inspired AI, and more specifically brain-inspired AI, promises to provide further biological aspects beyond those that are already traditionally included in AI, making it possible to assess and possibly overcome some of its present shortcomings. This article examines some conceptual, technical, and ethical issues raised by the development and use of brain-inspired AI. Against this background, the paper asks whether there is anything ethically unique about brain-inspired AI. The aim of the paper is to introduce a method that has a heuristic nature and that can be applied to identify and address the ethical issues arising from brain-inspired AI. The conclusion resulting from the application of this method is that, compared to traditional AI, brain-inspired AI raises new foundational ethical issues and some new practical ethical issues, and exacerbates some of the issues raised by traditional AI.

Peer Reviewed Articles

Abstract: Scientific research relies mainly on multimodal, multidimensional big data generated from both animal and human organisms as well as technical data. However, unlike human data that is increasingly regulated at national, regional and international levels, regulatory frameworks that can govern the sharing and reuse of non-human animal data are yet to be established. Whereas the legal and ethical principles that shape animal data generation in many countries and regions differ, the generated data are shared beyond boundaries without any governance mechanism. This paper, through perspectives from neuroscience, shows conceptually and empirically that there is a need for animal data governance that is informed by ethical concerns. There is a plurality of ethical views on the use of animals in scientific research that data governance mechanisms need to consider.

Abstract: Despite poor treatment results, a family-oriented approach and the securing of residency have been deemed essential to recovery from resignation syndrome (RS). In a retrospective cohort study, we evaluated an alternative method involving environmental therapy, with patients separated from their parents, while actively abstaining from involving the asylum process in treatment. We examined medical records, social services acts, and residential care home acts from 13 individuals treated at Solsidan residential care home between 2005 and 2020. Severity and outcome were assessed with Clinical Global Impression, Severity and Improvement subscales. Thirteen participants were included and out of these nine (69%) recovered, i.e. they very much or much improved. Out of the eight that were separated, all recovered, also, one non-separated recovered. The difference in outcome between subjects separated and not was significant ($p = 0.007$). Moreover, out of the five which received a residency permit during treatment, one recovered whereas four did not. The difference in outcome between subjects granted residency and not was significant ($p = 0.007$). The data revealed three (23%) cases of simulation where parents were suspected to have instigated symptoms. Our evaluation suggests that separation from parents and abstaining from invoking residency permit could be essential components when treating RS. Relying on a family-oriented approach, and residency could even be detrimental to recovery. The examined intervention was successful also in cases of probable malingering by proxy.

Farisco, M. The Ethical Spectrum of Consciousness, AJOB Neuroscience, 14:2, 55-57, (2023) doi: [10.1080/21507740.2023.2188312](10.1080/21507740.2023.2188312)

Abstract: This special issue of *AJOB Neuroscience* explores a number of ethical questions emerging from some of the most recent results of research on consciousness, including its moral interpretation, its technological manipulation, its artificial replication, its pharmacological alteration, and its possible attribution to engineered brain cells.

Ogoh, G., Akintoye, S., Eke, D., Farisco, M., Fernow, J., Grasenick, K., Guerrero, M., Rosemann, A., Salles, A., Ulnicane, I. Developing Capabilities for Responsible Research and Innovation (RRI), Journal of Responsible Technology, 2023, 100065, ISSN 2666-6596, doi: [10.1016/j.jrt.2023.100065](10.1016/j.jrt.2023.100065)

Abstract: This paper examines the notion of capability development through the lens of Responsible Research and Innovation (RRI). It describes how RRI capabilities can be promoted and developed through capacity-building activities by drawing on a reflection of the experiences of some members of the Ethics and Society team of the EU-funded Human Brain Project (HBP) who have contributed to the development of an extensive RRI capacity building programme for HBP researchers, EBRAINS research infrastructure developers and users, as well as interested external audiences. It also highlights some of the outcomes of the RRI capacity-building exercise and the challenge of assessing the impact of such activities.

Ulnicane, I., & Aden, A. (2023). Power and politics in framing bias in Artificial Intelligence policy. Review of Policy Research, 40(5). doi: 10.1111/ropr.12567

Abstract: Bias is a key issue in expert and public discussions about Artificial Intelligence (AI). While some hope that AI will help to eliminate human bias, others are concerned that AI will exacerbate it. To highlight political and power aspects of bias in AI, this contribution examines so far largely overlooked topic of framing of bias in AI policy. Among diverse approaches of diagnosing problems and suggesting prescriptions, we can distinguish two stylized framings of bias in AI policy—one more technical, another more social. Powerful technical framing suggests that AI can be a solution to human bias and can help to detect and eliminate it. It is challenged by an alternative social framing, which emphasizes the importance of social contexts, balance of power and structural inequalities. Technological frame sees simple technological fix as a way to deal with bias in AI. For the social frame, we suggest to approach bias in AI as a complex wicked problem, for which a broader strategy is needed involving diverse stakeholders and actions. The social framing of bias in AI considerably expands the legitimate understanding of bias and the scope of potential actions beyond technological fix. We argue that, in the context of AI policy, intersectional bias should not be perceived as a niche issue but rather be seen as a key to radically reimagine AI governance, power and politics in more participatory and inclusive ways.

Stahl, B. C., Bitsch, L. (2023). Building a responsible innovation toolkit as project legacy, Front. Res. Metr. Anal., 13 March 2023, Sec. Research Policy and Strategic Management, Volume 8. doi: 10.3389/frma.2023.1112106

Abstract: This article explores whether and in what way it is possible to employ toolkits for responsible research and innovation (RRI toolkits) as mechanisms for ensuring the legacy of RRI in research projects. Based on a review of the concept of responsible research and innovation as well as existing toolkits in the area, the article offers an account of the development of an RRI toolkit in the context of the EU- funded Human Brain Project. This toolkit is designed to integrate insights and practices of responsible research and innovation developed over a 10 year period into the project legacy, the EBRAINS research infrastructure. The article suggests that toolkits have the potential to contribute to ensuring a long- lasting legacy of work undertaken in responsible research and innovation, but that this potential requires further support from institutions and the broader research environment to become realized.

Farisco, M., Changeux, J.-P. About the compatibility between the perturbational complexity index and the global neuronal workspace theory of consciousness, Neuroscience of Consciousness, Volume 2023, Issue 1, 2023, niad016, doi: 10.1093/nc/niad016

Abstract: This paper investigates the compatibility between the theoretical framework of the global neuronal workspace theory (GNWT) of conscious processing and the perturbational complexity index (PCI). Even if it has been introduced within the framework of a concurrent theory (i.e. Integrated Information Theory), PCI appears, in principle, compatible with the main tenet of GNWT, which is a conscious process that depends on a long-range connection between different cortical regions, more specifically on the amplification, global propagation, and integration of brain signals. Notwithstanding this basic compatibility, a number of limited compatibilities and apparent differences emerge. This paper starts from the description of brain complexity, a notion that is crucial for PCI, to then summary of the main features of PCI and the main tenets of GNWT. Against this background, the text explores the compatibility between PCI and GNWT. It concludes that GNWT and PCI are fundamentally compatible, even though there are some partial disagreements and some points to further examine.

Stahl, B.C. Embedding responsibility in intelligent systems: from AI ethics to responsible AI ecosystems. *Sci Rep* 13, 7586 (2023). doi: 10.1038/s41598-023-34622-w

Abstract: Intelligent systems that are capable of making autonomous decisions based on input from their environment have great potential to do good, but they also raise significant social and ethical concerns. The discourse on ethics and artificial intelligence (AI) has covered these concerns in depth and developed an array of possible ways of addressing them. This

article argues that a shortcoming of this discourse is that it concentrates on specific issues and their mitigation but neglects the nature of intelligent systems as socio-technical systems of systems that are often described as ecosystems. Building on the discussion of ethics and AI, the article suggests that it would be beneficial to come to an understanding of what would constitute responsible AI ecosystems. By introducing the concept of meta-responsibility or higher-level responsibility, the article proposes characteristics that an ecosystem would have to fulfil, in order to be considered a responsible ecosystem. This perspective is theoretically interesting because it extends the current AI ethics discourse. It furthermore offers a novel perspective for researchers and developers of intelligent system and helps them reflect on the way they relate to ethical issues.

Stahl, B.C., Brooks, L., Hatzakis, T., Santiago, N., Wright. D. Building a responsible (2023) Exploring ethics and human rights in artificial intelligence – A Delphi study, Technological Forecasting and Social Change, 2023;191. doi: 10.1016/j.techfore.2023.122502

Abstract: Ethical and human rights issues of artificial intelligence (AI) are a prominent topic of research and innovation policy as well as societal and scientific debate. It is broadly recognised that AI-related technologies have properties that can give rise to ethical and human rights concerns, such as privacy, bias and discrimination, safety and security, economic distribution, political participation or the changing nature of warfare. Numerous ways of addressing these issues have been suggested. In light of the complexity of this discussion, we undertook a Delphi study with experts in the field to determine the most pressing issues and prioritise appropriate mitigation strategies. The results of the study demonstrate the difficulty of defining clear priorities. Our findings suggest that the debate around ethics and human rights of AI would benefit from being reframed and more strongly emphasising the systems nature of AI ecosystems.

Stahl, B.C., Antoniou, J., Bhalla, N., L, Brooks., P, Jansen., B, Lindqvist., A, Kirichenko., S, Marchal., R, Rodrigues., N, Santiago., Z, Warso., D, Wright. (2023) A systematic review of artificial intelligence impact assessments. Artif Intell Rev (2023). doi: https://doi.org/10.1007/s10462-023-10420-8

Abstract: Artificial intelligence (AI) is producing highly beneficial impacts in many domains, from transport to healthcare, from energy distribution to marketing, but it also raises concerns about undesirable ethical and social consequences. AI impact assessments (AI-IAs) are a way of identifying positive and negative impacts early on to safeguard AI's benefits and avoid its downsides. This article describes the first systematic review of these AI-IAs. Working with a population of 181 documents, the authors identified 38 actual AI-IAs and subjected them to a rigorous qualitative analysis with regard to their purpose, scope, organisational context, expected issues, timeframe, process and methods, transparency and challenges. The review demonstrates some convergence between AI-IAs. It also shows that the field is not yet at the point of full agreement on content, structure and

implementation. The article suggests that AI-IAs are best understood as means to stimulate reflection and discussion concerning the social and ethical consequences of AI ecosystems. Based on the analysis of existing AI-IAs, the authors describe a baseline process of implementing AI-IAs that can be implemented by AI developers and vendors and that can be used as a critical yardstick by regulators and external observers to evaluate organisations' approaches to AI.

Ulnicane, I., Mahfoud, T., Salles, A. (2023). Experimentation, learning, and dialogue: an RRI-inspired approach to dual-use of concern. Journal of Responsible Innovation 10 (1), doi:

Abstract: Responsible Research and Innovation is promoted by research funders and scientific communities as a way to place societal needs and values at the centre of research and innovation. In practice, however, legal compliance still tends to dominate the RRI agenda. In order to move beyond the dominance of legal compliance and address a broader societal agenda, this article argues that RRI requires: (1) a productive intertwining of research and practice; (2) the integration of anticipation, reflection, engagement, and action (AREA) in a non-linear process; and (3) an experimental approach. Based on this framework, this article draws on our experience of developing and institutionalizing an RRI-inspired approach to address dual-use and misuse issues in the EU-funded Human Brain Project. Our experience suggests that the four dimensions of the AREA framework work better not as separate stages but rather being flexibly intertwined to enable experimentation, learning, and dialogue.

Farisco, M., Pennartz, C., Annen, J. et al. (2022). Indicators and criteria of consciousness: ethical implications for the care of behaviourally unresponsive patients. BMC Med Ethics 23, 30. doi:

Abstract: Assessing consciousness in other subjects, particularly in non-verbal and behaviourally disabled subjects (e.g., patients with disorders of consciousness), is notoriously challenging but increasingly urgent. The high rate of misdiagnosis among disorders of consciousness raises the need for new perspectives in order to inspire new technical and clinical approaches. We take as a starting point a recently introduced list of operational indicators of consciousness that facilitates its recognition in challenging cases like non-human animals and Artificial Intelligence to explore their relevance to disorders of consciousness and their potential ethical impact on the diagnosis and healthcare of relevant patients. Indicators of consciousness mean particular capacities that can be deduced from observing the behaviour or cognitive performance of the subject in question (or from neural correlates of such performance) and that do not define a hard threshold in deciding about the presence of consciousness, but can be used to infer a graded measure based on the consistency amongst the different indicators. The indicators of consciousness under

consideration offer a potential useful strategy for identifying and assessing residual consciousness in patients with disorders of consciousness, setting the theoretical stage for an operationalization and quantification of relevant brain activity. Our heuristic analysis supports the conclusion that the application of the identified indicators of consciousness to its disorders will likely inspire new strategies for assessing three very urgent issues: the misdiagnosis of disorders of consciousness; the need for a gold standard in detecting consciousness and diagnosing its disorders; and the need for a refined taxonomy of disorders of consciousness.

Farisco, M., Evers, K., Salles, A. (2022). On the Contribution of Neuroethics to the Ethics and Regulation of Artificial Intelligence. Neuroethics 15, 4. doi: [10.1007/s12152-022-09484-0](10.1007/s12152-022-09484-0)

Abstract: Contemporary ethical analysis of Artificial Intelligence (AI) is growing rapidly. One of its most recognizable outcomes is the publication of a number of ethics guidelines that, intended to guide governmental policy, address issues raised by AI design, development, and implementation and generally present a set of recommendations. Here we propose two things: first, regarding content, since some of the applied issues raised by AI are related to fundamental questions about topics like intelligence, consciousness, and the ontological and ethical status of humans, among others, the treatment of these issues would benefit from interfacing with neuroethics that has been addressing those same issues in the context of brain research. Second, the identification and management of some of the practical ethical challenges raised by AI would be enriched by embracing the methodological resources used in neuroethics. In particular, we focus on the methodological distinction between conceptual and action-oriented neuroethical approaches. We argue that the normative (often principles-oriented) discussion about AI will benefit from further integration of conceptual analysis, including analysis of some operative assumptions, their meaning in different contexts, and their mutual relevance in order to avoid misplaced or disproportionate concerns and achieve a more realistic and useful approach to identifying and managing the emerging ethical issues.

Farisco, M., Zecchino, I., Capasso, G. CONNECT Consortium, The need for a multi-disciplinary reflection about frailty and cognitive impairment in chronic kidney disease, Nephrology Dialysis Transplantation. Nephrology Dialysis Transplantation, Volume 38, Issue 5, (2022), Pages 1064–1066, doi: 10.1093/ndt/gfac334

Abstract: This paper presents a call for multi-disciplinary colaboraiton, including euroscientific research on the molecular and cognitive dimensions of CI, neurological and neuro-psychological assessments of related physiological and behavioral factors, nephrological research on the relevant pato-pshysiological mechanisms, and ethico-legal reflection on the emerging ethical and regulatory issues; in the framework of the COST Action CONNECT (Cognitive Decline in Nephro-Neurology.

Schnakers, C., Bauer, C., Formisano, R., Noé, E., Llorens, R., Lejeune, N., Farisco, M., Teixeira, L., Morrissey, A-M., De Marco, S., Veeramuthu, V., Ilina, K., Edlow, BL., Gosseries, O., Zandalasini, M., De Bellis, F., Thibaut, A., Estraneo, A. (2022) What names for covert awareness? A systematic review. Front. Hum. Neurosci. 16:971315.

Abstract: With the emergence of Brain Computer Interfaces (BCI), clinicians have been facing a new group of patients with severe acquired brain injury who are unable to show any behavioral sign of consciousness but respond to active neuroimaging or electrophysiological paradigms. However, even though well documented, there is still no consensus regarding the nomenclature for this clinical entity. This systematic review aims to 1) identify the terms used to indicate the presence of this entity through the years, and 2) promote an informed discussion regarding the rationale for these names and the best candidates to name this fascinating disorder. The Disorders of Consciousness Special Interest Group (DoC SIG) of the International Brain Injury Association (IBIA) launched a search on Pubmed and Google scholar following PRISMA guidelines to collect peer-reviewed articles and reviews on human adults (>18 years) published in English between 2006 and 2021. The search launched in January 2021 identified 4,089 potentially relevant titles. After screening, 1,126 abstracts were found relevant. Finally, 161 manuscripts were included in our analyses. Only 58% of the manuscripts used a specific name to discuss this clinical entity, among which 32% used several names interchangeably throughout the text. We found 25 different names given to this entity. The five following names were the ones the most frequently used: covert awareness, cognitive motor dissociation, functional locked-in, non-behavioral MCS (MCS*) and higher-order cortex motor dissociation. Since 2006, there has been no agreement regarding the taxonomy to use for unresponsive patients who are able to respond to active neuroimaging or electrophysiological paradigms. Developing a standard taxonomy is an important goal for future research studies and clinical translation. We recommend a Delphi study in order to build such a consensus.

Das, J., Forlini, C., Porcello, D.M., Rommelfanger, K.S., Salles, A., Global Neuroethics Summit Delegates. (2022). Neuroscience is ready for neuroethics engagement. Frontiers in Communication, Vol. 7. doi: 10.3389/fcomm.2022.909964

Abstract: Neuroscience research has been expanding, providing new insights into brain and nervous system function and potentially transformative technological applications. In recent years, there has been a flurry of prominent international scientific academies and intergovernmental organizations calling for engagement with different publics on social, ethical, and regulatory issues related to neuroscience and neurotechnology advances. Neuroscientific activities and outputs are value-laden; they reflect the cultural, ethical, and political values that are prioritized in different societies at a given time and impact a variety of publics beyond the laboratory. The focus on engagement in neuroscience recognizes the breadth and significance of current neuroscience research whilst acknowledging the need

for a neuroethical approach that explores the epistemic and moral values influencing the neuroscientific agenda. The field of neuroethics is characterized by its focus on the social, legal, and philosophical implications of neuroscience including its impact on cultural assumptions about the cognitive experience, identity, consciousness, and decision-making. Here, we outline a proposal for neuroethics engagement that reflects an enhanced and evolving understanding of public engagement with neuroethical issues to create opportunities to share ideation, decision-making, and collaboration in neuroscience endeavors for the benefit of society. We demonstrate the synergies between public engagement and neuroethics scholarship and activities that can guide neuroethics engagement.

Grasenick, K., Beranek, S., Godfroy, A.-S., Reidl, S., Romero, P. F., Schiffbänker, H., Stadlbauer, J., Trattnig, J., Wolffram, A., Wächter, C. (2022). The Gendered Analog-Digital Divide in Virtual Academia, Critical Issues in Science, Technology, and Society Studies. doi: [10.3217/978-3-85125-932-2-06](10.3217/978-3-85125-932-2-06)

Abstract: During the Covid-19 pandemic, in times of worldwide lockdowns, academic careers were impacted in a gendered way (Gabster et al., 2020): Existing gender inequalities have increased (Oleschuk, 2020) and female academics, especially early career researchers, have conducted less research compared to men (Viglione, 2020; Amano-Patiño et al., 2020) dropped radically. One might argue that the pandemic is over, and that academic life is back to normal, leaving us with the benefits of improved tools and practices for virtual collaboration. However, virtual academia risks increasing inequalities, an effect we will refer to as the analogue-digital divide. These risks affect especially researchers at the beginning of their career, and those who contribute to a greater extent to academic and family care work, which are mainly women. In this paper, we reflect and consolidate the findings of six projects with partners across Europe and two universities which analysed the lasting effects of the pandemic from the perspectives of researchers at different career stages, service staff, as well as decision makers in boards and juries. We conclude that strategies and measures developed before the Covid-19 pandemic do not consider virtual academia sufficiently. We thus suggest how to counteract the analog-digital divide with requests of funding organisations and implemented by research processing organisations in their gender equality plans (GEPs).

Stahl, B.C., Leach, T. (2022). Assessing the ethical and social concerns of artificial intelligence in neuroinformatics research: an empirical test of the European Union Assessment List for Trustworthy AI (ALTAI). AI Ethics. doi: [10.1007/s43681-022-00201-4](10.1007/s43681-022-00201-4)

Abstract: Ethical and social concerns are a key obstacle to the adoption of artificial intelligence (AI) in the life sciences and beyond. The discussion of these issues has intensified in recent years and led to a number of approaches, tools and initiatives. Key amongst them is the idea of ex-ante impact assessments that aim to identify issues at the early stages of

development. One prominent example of such ex-ante impact assessment is the European Union's (EU) Assessment list for Trustworthy AI (ALTAI). This article uses the findings of a large-scale application of the ALTAI to a large neuro-informatics project as an exemplar to demonstrate the effectiveness and limitations of the ALTAI in practice. The article shows that ex-ante impact assessment has the potential to help identify and address ethical and social issues. However, they need to be understood as part of a broader socio-technical ecosystem of AI. For ALTAI and related approaches to be useful in bio-medical research, they should be interpreted from a systems theory perspective which allows for their integration into the rich set of tools, legislation and approaches. The paper argues that ex-ante impact assessments have the best chance of being successful if seen applied in conjunction with other approaches in the context of the overall AI ecosystem.

Ulnicane, I. (2022). Emerging technology for economic competitiveness or societal challenges? Framing purpose in Artificial Intelligence policy. Global Public Policy and Governance 2(3): 326-345. doi: 10.1007/s43508-022-00049-8

Abstract: This study examines the policy discussions surrounding the purpose of the development and use of an emerging technology. It applies the two stylized technology policy frames of economic growth and societal challenges to analyse framing of one of the key emerging technologies today—Artificial Intelligence (AI). It demonstrates that recent AI policy documents include both—economic growth as well as societal challenges—frames. While AI is a novel technology, its recent policy builds on traditional ideas about the role of technology in facilitating economic growth and competitiveness supported by well-known measures such as investment in research and highly skilled workforce. Additionally, AI policy draws on more recent frame on the contribution of technology to addressing societal challenges and the Sustainable Development Goals but presents AI as a technological solution to complex societal issues. While some interest in addressing both economic and social objectives in AI policy can be observed, the policy documents remain silent about their compatibility.

Pedrocchi, A. (2022). The Cultural and Social Challenge of Promoting the Professional Value of Motherhood and Fatherhood. Frontiers in Neuroscience, Vol. 16.
doi: 10.3389/fnins.2022.853329

Abstract: Speaking of gender balance in professional fields is a point of honor and maturation of our culture. It is now standard, and we are honored that, in all the most organized contexts, especially in science technology engineering mathematics (STEM) disciplines, there is a strong commitment to gender balance. This commitment is usually and mainly expressed by facilitating women's presence, with roles and perspectives similar to those of male colleagues. Huang et al. (2020) thoroughly analyzed the gender gap in STEM academia. Men still outnumber women 2 to 1 in the scientific workforce and, on average, have more productive careers and accumulate more impact in every STEM discipline and most

geographic regions. Interestingly, those authors found that men and women publish a comparable number of papers per year and have an equivalent career-wise impact for the same total number of publications. They highlighted the issue of gender-specific sustainability of publication effort throughout an entire academic career, suggesting significant consequences for institutions and policymakers.

Matar, A., Silverman, H.J. (2022). Ethical Analysis of Egypt's Law Regulating Clinical Research. Journal of Empirical Research on Human Research Ethics, Volume 17, Issue 4, doi: 10.1177/15562646221096188

Abstract: Lately, there has been increased research performed in Egypt. In response, the Egyptian Parliament published its first clinical research law in December 2020. The official version of the law was translated to English from Arabic and back by an accredited translation service. We performed an ethical analysis of the law based on the seven ethical requirements for clinical research proposed by Emanuel et al. and compared it with other regulations in the Arab region. The law contains provisions that fulfill all requirements for ethical research to varying degree. Provisions necessitating the sharing of participants' data and biospecimens by the Central Intelligence Agency requires further specifications to ensure privacy protection. Also, the law poses problematic liabilities that could hamper medical research. Egypt's law compares favorably with other laws in the region. Potential items that require further specification can be addressed in the executive regulations currently being drafted for the law.

Farisco, M., Salles, A., (2022). American and European Guidelines on Disorders of Consciousness: Ethical Challenges of Implementation. Journal of Head Trauma Rehabilitation 37(4):p 258-262, doi: 10.1097/htr.0000000000000776

Abstract: The recently published Guidelines on Disorders of Consciousness (DoCs) by the European Academy of Neurology (EAN)1 and by the American Academy of Neurology (AAN) in collaboration with the American Congress of Rehabilitation Medicine (ACRM) and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR)2 stand as the most ambitious international attempts to provide clear and standardized recommendations to clinicians working with patients with DoCs. They offer an updated, timely, and wide-ranging list of recommendations for the diagnosis, prognosis, and clinical care of affected patients.

However, while commendable, the guidelines pose a number of questions,3,4 including some related to the practical implementation of their recommendations. For example, both documents rightly consider that the integration of behavioral and instrumental assessments is the best strategy to improve diagnostic accuracy and quality of care. Yet, the operationalization of this recommendation in actual clinical settings, where the necessary

technology, for example, functional magnetic resonance image (fMRI), might not be available or convenient to use, remains an open issue. The potential unavailability of the required technology or of the necessary expertise for use in some contexts might undermine the reliability, practical value, and ethical impact of some recommendations. These are not minor concerns because the practical inapplicability of specific recommendations has clear ethical implications: it might directly impact patients' well-being, their right to the best possible care, the communication between clinicians and family members, and overall shared decision-making and unintentionally lead to unequal and unfair treatment of some patients.

In this commentary, we propose a responsibility-oriented strategy to address some of the practical and normative barriers to implementation. We begin by summarizing the main points of the 2 guidelines. Next, we outline a Distributed Responsibility Model (DRM) based on a distributed multilevel understanding of responsibility as a means to better understand and address barriers to the implementation of the recommendations. While we do not aim at elaborating an in-depth analysis of all the relevant factors, we hope to set the stage for a more inclusive and comprehensive discussion that involves diverse relevant stakeholders (ie, clinicians, researchers, and hospital managers, among others).

Ulnicane, I. Against the new space race: global AI competition and cooperation for people. *AI & Soc* 38, 681–683 (2023). doi: [10.1007/s00146-022-01423-0](#)

Abstract: This Open Forum contribution critically interrogates the use of space race rhetoric in current discussions about artificial intelligence (AI). According to this rhetoric, similar to the space race of the twentieth century, AI development is portrayed as a rivalry among superpowers where one country will win and reap major benefits, while others will be left behind. Using this rhetoric to frame AI development tends to prioritize narrow and short-term economic interests over broader and longer-term societal needs. Three particularly problematic aspects of using space race narrative in AI discussions are highlighted: first, perceiving technology development globally as a zero-sum game; second, pressuring governments to regulate less; and third, prioritising high technologies over social issues. This contribution invites to challenge underlying political assumptions of the AI space race discourse and to open up conversations about alternative futures.

Rommelfanger, K. S., Pustilnik, A., Salles, A. (2022). Mind the Gap: Lessons Learned from Neurorights, Science & Diplomacy.

https://www.sciencediplomacy.org/article/2022/mind-gap-lessons-learned-neurorights

Abstract: Transnational organizations, think tanks, and scholars are currently debating the implications of new and powerful neurotechnologies, including whether there is need for regulations and new "neurorights" to protect individual cognitive liberty. The Organisation for Economic Co-operation and Development (OECD), an international policy organization

with 38 member countries, generated the Recommendation on Responsible Innovation in Neurotechnology, and, together with the Council of Europe, recently organized a roundtable, "Neurotechnologies and Human Rights Framework: Do We Need New Rights?" Neurorights is also on the Organization of American States (OAS) Inter-American Juridical Committee (CJI)'s agenda, and the UN considered neurotechnology development in its 2021 report to the Secretary-General. Numerous civil society organizations are currently evaluating legal and ethical frameworks for neurotechnologies, notably the Neurorights Foundation, whose mission to develop "new human rights for the age of neurotechnology." At the national level, Chile has enshrined a right to neuroprotection at the highest possible legal level—its constitution—and is developing legislation for neuroprotection as well.

Amunts, K., DeFelipe, J., Pennartz, C., Destexhe, A., Migliore, M., Ryvlin, P., . . . Jirsa, V. (2022). Linking brain structure, activity, and cognitive function through computation. Eneuro, 9(2). doi: 10.1523/eneuro.0316-21.2022

Abstract: Understanding the human brain is a "Grand Challenge" for 21st century research. Computational approaches enable large and complex datasets to be addressed efficiently, supported by artificial neural networks, modeling and simulation. Dynamic generative multiscale models, which enable the investigation of causation across scales and are guided by principles and theories of brain function, are instrumental for linking brain structure and function. An example of a resource enabling such an integrated approach to neuroscientific discovery is the BigBrain, which spatially anchors tissue models and data across different scales and ensures that multiscale models are supported by the data, making the bridge to both basic neuroscience and medicine. Research at the intersection of neuroscience, computing and robotics has the potential to advance neuro-inspired technologies by taking advantage of a growing body of insights into perception, plasticity and learning. To render data, tools and methods, theories, basic principles and concepts interoperable, the Human Brain Project (HBP) has launched EBRAINS, a digital neuroscience research infrastructure, which brings together a transdisciplinary community of researchers united by the quest to understand the brain, with fascinating insights and perspectives for societal benefits.

Schirner, M., Domide, L., Perdikis, D., Triebkorn, P., Stefanovski, L., Pai, R., . . . Ritter, P. (2022). Brain simulation as a cloud service: The Virtual Brain on EBRAINS. NeuroImage, 251, 118973. doi: 10.1016/j.neuroimage.2022.118973

Abstract: The Virtual Brain (TVB) is now available as open-source services on the cloud research platform EBRAINS (ebrains.eu). It offers software for constructing, simulating and analysing brain network models including the TVB simulator; magnetic resonance imaging (MRI) processing pipelines to extract structural and functional brain networks; combined simulation of large-scale brain networks with small-scale spiking networks; automatic conversion of user-specified model equations into fast simulation code; simulation-ready brain models of patients and healthy volunteers; Bayesian parameter optimization in

epilepsy patient models; data and software for mouse brain simulation; and extensive educational material. TVB cloud services facilitate reproducible online collaboration and discovery of data assets, models, and software embedded in scalable and secure workflows, a precondition for research on large cohort data sets, better generalizability, and clinical translation.

Ulnicane, I., Knight, W., Leach, T., Stahl, B.C., Wanjiku, W. (2021) Framing governance for a contested emerging technology:insights from AI policy, Policy and Society, 40:2, 158-177, doi: 10.1080/14494035.2020.1855800

Abstract: This paper examines how the governance in AI policy documents have been framed as way to resolve public controversies surrounding AI. It draws on the studies of governance of emerging technologies, the concept of policy framing, and analysis of 49 recent policy documents dedicated to AI which have been prepared in the context of technological hype expecting fast advances of AI that will fundamentally change economy and society. The hype about AI is accompanied by major public controversy about positive and negative effects of AI. Against the backdrop of this policy controversy, governance emerges as one of the frames that diagnoses the problems and offers prescriptions. Accordingly, the current governance characterized by oligopoly of a small number of large companies is indicated as one of the reasons for problems such as lack of consideration of societal needs and concerns. To address these problems, governance frame in AI policy documents assigns more active and collaborative roles to the state and society. Amid public controversies, the state is assigned the roles of promoting and facilitating AI development while at the same time being a guarantor of risk mitigation and enabler of societal engagement. High expectations are assigned to public engagement with multiple publics as a way to increase diversity, representation and equality in AI development and use. While this governance frame might have a normative appeal, it is not specific about addressing some well-known challenges of the proposed governance mode such as risks of capture by vested interests or difficulties to achieve consensus.

Evers, K., Salles, A. (2021). Epistemic Challenges of Digital Twins & Virtual Brains: Perspectives from Fundamental Neuroethics. SCIO: Journal of Philosophy, (21), 27–53, doi: 10.46583/scio_2021.21.846

Abstract: In this article, we present and analyse the concept of Digital Twin (DT) linked to distinct types of objects (artefacts, natural, inanimate or living) and examine the challenges involved in creating them from a fundamental neuroethics approach that emphasises conceptual analyses. We begin by providing a brief description of DTs and their initial development as models of artefacts and physical inanimate objects, identifying core challenges in building these tools and noting their intended benefits. Next, we describe attempts to build DTs of model living entities, such as hearts, highlighting the novel challenges raised by this shift from DTs of inanimate to DTs of living objects. Against that

background, we give an account of contemporary research aiming to develop DTs of the human brain by building "virtual brains", e.g. the simulation engine The Virtual Brain (TVB) as it is carried out in the European Human Brain Project. Since the brain is structurally and functionally the most complex organ in the human body, and our integrated knowledge of its functional architecture remains limited in spite of recent neuroscientific advances, the attempts to create virtual copies of the human brain are correspondingly challenging. We suggest that a clear scientific theoretical structure, conceptual clarity and transparency regarding the methods and goals of this technological development are necessary prerequisites in order to make the project of constructing virtual brains a theoretically promising and socially beneficial scientific, technological and philosophical enterprise.

Stahl, B. C. (2021). Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence. International Journal of Information Management, Vol. 62, doi: 10.1016/j.ijinfomgt.2021.102441

Abstract: The concept of innovation ecosystems has become prominent due to its explanatory power. It offers a convincing account of innovation, explaining how and why innovation pathways change and evolve. It has been adopted to explain, predict, and steer innovation. The increasing importance of innovation for most aspects of human life calls for the inclusion of ethical and social rights aspects into the innovation ecosystems discourse. The current innovation ecosystems literature does not provide guidance on how the integration of ethical and social concerns into innovation ecosystems can be realised. One way to achieve this is to draw on the discussion of responsible research and innovation (RRI). This paper applies RRI to the innovation ecosystems discourse and proposes the concept of responsible innovation systems. It draws on the discussion of the ethics of artificial intelligence (AI) to explore how responsible AI innovation ecosystems can be shaped and realised.

Eke, D., Aasebø, I. E.J., Akintoye, S., Knight, W., Karakasidis, A., Mikulan, E., Ochang, P., Ogoh, G., Oostenveld, R., Pigorini, A., Stahl, B.C., White, T., Zehl L. (2021). Pseudonymisation of neuroimages and data protection: Increasing access to data while retaining scientific utility, Neuroimage: Reports, Vol. 1, No. 4, doi: 10.1016/j.ynirp.2021.100053

Abstract: For a number of years, facial features removal techniques such as 'defacing', 'skull stripping' and 'face masking/blurring', were considered adequate privacy preserving tools to openly share brain images. Scientifically, these measures were already a compromise between data protection requirements and research impact of such data. Now, recent advances in machine learning and deep learning that indicate an increased possibility of re-identifiability from defaced neuroimages, have increased the tension between open science and data protection requirements. Researchers are left pondering how best to comply with the different jurisdictional requirements of anonymization, pseudonymisation or de-identification without compromising the scientific utility of neuroimages even further. In this

paper, we present perspectives intended to clarify the meaning and scope of these concepts and highlight the privacy limitations of available pseudonymisation and de-identification techniques. We also discuss possible technical and organizational measures and safeguards that can facilitate sharing of pseudonymised neuroimages without causing further reductions to the utility of the data.

Stahl, B. C., Akintoye, S., Bitsch, L., Bringedal, B., Eke, D., Farisco, M., Grasenick, K., Guerrero, M., Knight, W., Leach, T., Nyholm, S., Ogoh, G., Rosemann, A., Salles, A., Trattnig, J. & Ulnicane, I. (2021). From Responsible Research and Innovation to responsibility by design. Journal of Responsible Innovation, 8 (2): 175-198, doi: 10.1080/23299460.2021.1955613

Abstract: Drawing on more than eight years working to implement Responsible Research and Innovation (RRI) in the Human Brain Project, a large EU-funded research project that brings together neuroscience, computing, social sciences, and the humanities, and one of the largest investments in RRI in one project, this article offers insights on RRI and explores its possible future. We focus on the question of how RRI can   have   long-lasting impact and persist beyond the time horizon of funded projects. For this purpose, we suggest the concept of 'responsibility by design' which is intended to encapsulate the idea of embedding RRI in research and innovation in a way that makes it part of the fabric of the resulting outcomes, in our case, a distributed European Research Infrastructure.

Stahl, B. C. (2021). From computer ethics and the ethics of AI towards an ethics of digital ecosystems. AI and Ethics, 1-13, doi 10.1007/s43681-021-00080-1

Abstract: Ethical, social and human rights aspects of computing technologies have been discussed since the inception of these technologies. In the 1980s, this led to the development of a discourse often referred to as computer ethics. More recently, since the middle of the 2010s, a highly visible discourse on the ethics of artificial intelligence (AI) has developed. This paper discusses the relationship between these two discourses and compares their scopes, the topics and issues they cover, their theoretical basis and reference disciplines, the solutions and mitigations options they propose and their societal impact. The paper argues that an understanding of the similarities and differences of the discourses can benefit the respective discourses individually. More importantly, by reviewing them, one can draw conclusions about relevant features of the next discourse, the one we can reasonably expect to follow after the ethics of AI. The paper suggests that instead of focusing on a technical artefact such as computers or AI, one should focus on the fact that ethical and related issues arise in the context of socio-technical systems. Drawing on the metaphor of ecosystems which is widely applied to digital technologies, it suggests preparing for a discussion of the ethics of digital ecosystems. Such a discussion can build on and benefit from a more detailed understanding of its predecessors in computer ethics and the ethics of AI.

Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., & Stahl, B. (2021). Research and Practice of AI Ethics: A Case Study Approach Juxtaposing Academic Discourse with Organisational Reality. Science and Engineering Ethics, 27(2), 16, doi: 10.1007/s11948-021-00293-x

Abstract This study investigates the ethical use of Big Data and Artificial Intelligence (AI) technologies (BD+AI)—using an empirical approach. The paper categorises the current literature and presents a multi-case study of 'on-the- ground' ethical issues that uses qualitative tools to analyse findings from ten targeted case-studies from a range of domains. The analysis coalesces identified singular ethical issues, (from the literature), into clusters to offer a comparison with the proposed classification in the literature. The results show that despite the variety of different social domains, fields, and applications of AI, there is overlap and correlation between the organisations' ethical concerns. This more detailed understanding of ethics in AI+BD is required to ensure that the multitude of suggested ways of addressing them can be targeted and succeed in mitigating the pertinent ethical issues that are often discussed in the literature.

Ulnicane, I., Eke, D. O., Knight, W., Ogoh, G., & Stahl, B. C. (2021). Good governance as a response to discontents? Déjà vu, or lessons for AI from other emerging technologies. Interdisciplinary Science Reviews, 46(1-2), 71-93, doi: 10.1080/03080188.2020.1840220

Abstract: Recent advances in Artificial Intelligence (AI) have led to intense debates about benefits and concerns associated with this powerful technology. These concerns and debates have similarities with developments in other emerging technologies characterized by prominent impacts and uncertainties. Against this background, this paper asks, What can AI governance, policy and ethics learn from other emerging technologies to address concerns and ensure that AI develops in a socially beneficial way? From recent literature on governance, policy and ethics of emerging technologies, six lessons are derived focusing on inclusive governance with balanced and transparent involvement of government, civil society and private sector; diverse roles of the state including mitigating risks, enabling public participation and mediating diverse interests; objectives of technology development prioritizing societal benefits; international collaboration supported by science diplomacy, as well as learning from computing ethics and Responsible Innovation.

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2021). Organisational responses to the ethical issues of artificial intelligence. AI & Society, 1-15, doi: 10.1007/s00146-021-01148-6

Abstract: The ethics of artifial intelligence (AI) is a widely discussed topic. There are numerous initiatives that aim to develop the principles and guidance to ensure that the development, deployment and use of AI are ethically acceptable. What is generally unclear is how organisations that make use of AI understand and address these ethical issues in

practice. While there is an abundance of conceptual work on AI ethics, empirical insights are rare and often anecdotal. This paper fills the gap in our current understanding of how organisations deal with AI ethics by presenting empirical findings collected using a set of ten case studies and providing an account of the cross-case analysis. The paper reviews the discussion of ethical issues of AI as well as mitigation strategies that have been proposed in the literature. Using this background, the cross-case analysis categorises the organisational responses that were observed in practice. The discussion shows that organisations are highly aware of the AI ethics debate and keen to engage with ethical issues proactively. However, they make use of only a relatively small subsection of the mitigation strategies proposed in the literature. These insights are of importance to organisations deploying or using AI, to the academic AI ethics debate, but maybe most valuable to policymakers involved in the current debate about suitable policy developments to address the ethical issues raised by AI.

Stahl, B. C., Andreou, A., Brey, P., Hatzakis, T., Kirichenko, A., Macnish, K., Shaelou, S. L., Patel, A., Ryan, M. & Wright, D. (2021). Artificial intelligence for human flourishing–Beyond principles for machine learning. Journal of Business Research, 124, 374-388, doi: 10.1016/j.jbusres.2020.11.030

Abstract The technical and economic benefits of artificial intelligence (AI) are counterbalanced by legal, social and ethical issues. It is challenging to conceptually capture and empirically measure both benefits and downsides. We therefore provide an account of the findings and implications of a multi-dimensional study of AI, comprising 10 case studies, five scenarios, an ethical impact analysis of AI, a human rights analysis of AI and a technical analysis of known and potential threats and vulnerabilities. Based on our findings, we separate AI ethics discourse into three streams: (1) specific issues related to the application of machine learning, (2) social and political questions arising in a digitally enabled society and (3) metaphysical questions about the nature of reality and humanity. Human rights principles and legislation have a key role to play in addressing the ethics of AI. This work helps to steer AI to contribute to human flourishing.

Prescott, T. J., Robillard, J. M. (2021). Are friends electric? The benefits and risks of human-robot relationships. iScience, Vol. 24, No. 1, doi: 10.1016/j.isci.2020.101993

Abstract Social robots that can interact and communicate with people are growing in popularity for use at home and in customer-service, education, and healthcare settings. Although growing evidence suggests that co-operative and emotionally aligned social robots could benefit users across the lifespan, controversy continues about the ethical implications of these devices and their potential harms. In this perspective, we explore this balance between benefit and risk through the lens of human-robot relationships. We review the definitions and purposes of social robots, explore their philosophical and psychological status, and relate research on human-human and human-animal relationships to the emerging literature on human-robot relationships. Advocating a relational rather than

essentialist view, we consider the balance of benefits and harms that can arise from different types of relationship with social robots and conclude by considering the role of researchers in understanding the ethical and societal impacts of social robotics.

Kleinberger-Pierer, M., Pohn-Weidinger, S., Grasenick, K. (2020). Fair projects-bad data? Evaluating the gender balance in science projects. fteval Journal for Research and Technology Policy Evaluation 50. 60-71, doi: 10.22163/fteval.2020.471

Abstract: Researchers are more and more frequently required to report the gender balance of their teams in order to receive funding for their research projects. In Europe for instance, Horizon 2020 guidelines determine that applications with a balanced, 50/50 representation of women and men will be given preference, and each project must at least justify the composition of their teams and leadership positions. In scientific fields where women are still a minority, like robotics or artificial intelligence, or many STEMM fields in general, this approach has left the applicants with the issue of how to justify the ratio of women on their teams. For individual projects and project leaders, realistic objectives are required that take into account specific framework conditions in different scientific fields. This paper examines approaches to measure the disciplinary background and career development of women and men in science by assessing a range of available data sources. It provides insights on how to derive figures allowing science projects to evaluate their gender ratios against a possible underrepresen-tation of women.

Grasenick, K., Guerrero, M. (2020). Responsible Research and Innovation & Digital Inclusiveness during Covid-19 Crisis in the Human Brain Project (HBP). Journal of Responsible Technology, 1, 100001, doi: 10.1016/j.jrt.2020.06.001

Abstract: Covid-19 changes the lives for all of us: Institutions and other places are closed; it is not possible to see friends and family personally and keeping distance is the topmost commandment. Therefore, most of us are working from home and digitalisation is on the way up in many aspects of life. The HBP has a long-lasting experience of interdisciplinary collaboration by virtually bridging distances because its involved partners are not only complex but also spatially remote. In these challenging times of the pandemic, the HBP's Diversity and Equal Opportunities Committee together with the Ethics Rapporteur Programme has started "I-include", an Initiative for Inclusive Digital Engagement to make sure that no one is left behind virtually and that diversity matters in digital collaborations. It offers recommendations based on practical experiences of HBP members. Considering this new framework during the current situation is a way to ensure that our digitally distributed work becomes a valuable and successful experience corresponding to the standards of Responsible Research and Innovation (RRI). RRI is a dynamic, iterative process in which all stakeholders in research and innovation become mutually responsive and share responsibility for both the process and its outcomes. Even and particularly in difficult times.

Salles, A., & Farisco, M. (2020). Of ethical frameworks and neuroethics in big neuroscience projects: A view from the HBP. AJOB neuroscience, 11(3), 167-175, doi: 10.1080/21507740.2020.1778116

Abstract: The recently published BRAIN 2.0 Neuroethics Report offers a very helpful overview of the possible ethical, social, philosophical, and legal issues raised by neuroscience in the context of BRAIN's research priorities thus contributing to the attempt to develop ethically sound neuroscience. In this article, we turn to a running theme of the document: the need for an ethical framework for the BRAIN Initiative and for further integration of neuroethics and neuroscience. We assess some of the issues raised and provide an explanation of how we have addressed them in the Human Brain Project. We offer our experience in the HBP as a potential contribution to the international debate about neuroethics in the big brain initiatives. Our hope is that among other things, the type of exchange proposed by this AJOB special issue will prove productive in further identifying and discussing the issues and in inspiring appropriate solutions.

Aicardi, C., Akintoye, S., Fothergill, B. T., Guerrero, M., Klinker, G., Knight, W., Klüver, L., Morel, Y., Morin, F. O., Stahl, B. C. & Ulnicane, I. (2020). Ethical and social aspects of neurorobotics. Science and Engineering Ethics, 26(5), 2533-2546, doi 10.1007/s11948-020-00248-8

Abstract: The interdisciplinary field of neurorobotics looks to neuroscience to overcome the limitations of modern robotics technology, to robotics to advance our understanding of the neural system's inner workings, and to information technology to develop tools that support those complementary endeavours. The development of these technologies is still at an early stage, which makes them an ideal candidate for proactive and anticipatory ethical reflection. This article explains the current state of neurorobotics development within the Human Brain Project, originating from a close collaboration between the scientific and technical experts who drive neurorobotics innovation, and the humanities and social sciences scholars who provide contextualising and reflective capabilities. This article discusses some of the ethical issues which can reasonably be expected. On this basis, the article explores possible gaps identified within this collaborative, ethical reflection that calls for attention to ensure that the development of neurorobotics is ethically sound and socially acceptable and desirable.

Northoff, G, Wainio-Theberge, S, Evers, K, Spatiotemporal neuroscience – what is it and why we need it. (2020) Physics of Life Reviews; 33, 78-87, doi: 10.1016/j.plrev.2020.06.005

Abstract: The excellent commentaries to our target paper hint upon three main issues, (i) spatiotemporal neuroscience; (ii) neuro-mental relationship; and (iii) mind, brain, and world relationship. (i) We therefore discuss briefly the history of Spatiotemporal Neuroscience.

Distinguishing it from Cognitive Neuroscience and related branches (like Affective, Social, etc. Neuroscience), Spatiotemporal Neuroscience can be characterized by focus on brain activity (rather than brain function), spatiotemporal relationship (rather than input-cognition-output relationship), and structure (rather than stimuli/contents). (ii) Taken in this sense, Spatiotemporal Neuroscience allows one to conceive the neuro-mental relationship in dynamic spatiotemporal terms that complement and extend (rather than contradict) their cognitive characterization. (iii) Finally, more philosophical issues like the need to dissolve the mind-body problem (and replace it by the world-brain relation) and the question for different levels of time including their nestedness are discussed.

Northoff, G, Wainio-Theberge, S, Evers, K, Is temporo-spatial dynamics the "common currency" of brain and mind? In Quest of "Spatiotemporal Neuroscience" (2020). Physics of Life Reviews; 33, 34-54, doi: 10.1016/j.plrev.2019.05.002

Abstract: Neuroscience has made considerable progress in unraveling the neural correlates of mental phenomena like self, consciousness, and perception. However, the "common currency" shared between neuronal and mental activity, brain and mind, remains yet unclear. In this article, we propose that the dynamics of time and space provides a "common currency" that connects neuronal and mental features. Time and space are here understood in a dynamic context (as in contemporary physics): that is, in terms of the way the brain's spontaneous activity constructs its spatial and temporal relationships, for instance in terms of functional connectivity and different frequencies of fluctuations. Recruiting recent empirical evidence, we show that the different ways in which the spontaneous activity constructs its "inner time and space" are manifested in distinct mental features. Specifically, we demonstrate how spatiotemporal mechanisms like spatiotemporal repertoire, integration, and speed yield mental features like consciousness, self, and time speed perception. The focus on the brain's spatiotemporal mechanisms entails what we describe as "Spatiotemporal Neuroscience". Spatiotemporal Neuroscience conceives neuronal activity in terms of its temporo-spatial dynamics rather than its various functions (e.g., cognitive, affective, social, etc.) as in other branches of neuroscience (as distinguished from Cognitive, Affective, Cultural, Social, etc. Neuroscience). That allows Spatiotemporal Neuroscience to take into view the so-called 'spatio-temporality' of mental features including their non-causal, intrinsic and transformative relationship with neuronal features. In conclusion, Spatiotemporal Neuroscience opens the door to investigate and ultimately reveal the brain's own temporo-spatial dynamics as the hitherto missing "common currency" of neuronal and mental features.

Evers, K. (2020). The Culture-Bound Brain: Epigenetic Proaction Revisited. Theoria, 86(6), 783-800, doi: 10.1111/theo.12264

Abstract: Progress in neuroscience – notably, on the dynamic functions of neural networks – has deepened our understanding of decision-making, acquisition of character and

temperament, and the development of moral dispositions. The evolution of our cerebral architecture is both genetic and epigenetic: the nervous system develops in continuous interaction with the immediate physical and socio-cultural environments. Each individual has a unique cerebral identity even in the relative absence of genetic distinction, and the development of this identity is strongly influenced by social and cultural environments leaving major traces in the connectivity of the brain. This interaction introduces important elements of variability and plasticity. Synaptic epigenetic theories of cultural and social imprinting on our brain architecture suggest the possibility that we can be "epigenetically proactive" and adapt our social structures, in both the short and the long term, to benefit and constructively interact with the ever-developing neuronal architecture of our brains. Epigenetic proaction can be described as a way to socially steer evolution by influencing the cultural imprints to be stored in our brains. The purpose of this article is to present, in a historical and ideological light, the idea that we may culturally shape the developing brain, and discuss some key risks and benefits of this endeavour. The article is a contribution to neuroethics.

Farisco, M., Evers, K., & Salles, A. (2020). Towards establishing criteria for the ethical analysis of artificial intelligence. Science and Engineering Ethics, 26(5), 2413-2425, doi:

Abstract: Ethical reflection on Artificial Intelligence (AI) has become a priority. In this article, we propose a methodological model for a comprehensive ethical analysis of some uses of AI, notably as a replacement of human actors in specific activities. We emphasize the need for conceptual clarification of relevant key terms (e.g., intelligence) in order to undertake such reflection. Against that background, we distinguish two levels of ethical analysis, one practical and one theoretical. Focusing on the state of AI at present, we suggest that regardless of the presence of intelligence, the lack of morally relevant features calls for caution when considering the role of AI in some specific human activities.

Rainey, S., McGillibray, K., Akintoye, S., Fothergill, T., Bublitz, C., & Stahl, B. C. (2020). Is European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology? Journal of Law and the Biosciences, 7(1), lsaa051, doi:

Abstract: Research-driven technology development in th fields of the neurosciences presents interesting and potentially complicated issues around data in general and brain data specifically. The data produced from brain recordings are unlike names and addresses in that it may result from the processing of largely involuntarily brain activity, it can be processed and reprocessed for different aims, and it is highly sensitive. Consenting for brain recordings of a specific type, or for a specific purpose, is complicated by these factors. Brain data collection, retention, processing, storage, and destruction are each of high ethical importance. This leads us to ask: Is the present European Data Protection Regulation

sufficient to deal with emerging data concerns relating to neurotechnology? This is pressing especially in a context of rapid advancement in the fields of brain computer interfaces (BCIs), where devices that can function via recorded brain signals are expanding from research labs, through medical treatments, and beyond into consumer markets for recreational uses. One notion we develop herein is that there may be no trivial data collection when it comes to brain recording, especially where algorithmic processing is involved. This article provides analysis and discussion of some specific data protection questions related to neurotechnology, especially BCIs. In particular, whether and how brain data used in BCI-driven applications might count as personal data in a way relevant to data protection regulations. It also investigates how the nature of BCI data, as it appears in various applications, may require different interpretations of data protection concepts. Importantly, we consider brain recordings to raise questions about data sensitivity, regardless of the purpose for which they were recorded. This has data protection implications.

Salles, A., Evers, K., Farisco, M., (2020) Anthropomorphism in AI, AJOB Neuroscience, 11:2, 88-95, doi: 10.1080/21507740.2020.1740350

Abstract: AI research is growing rapidly raising various ethical issues related to safety, risks, and other effects widely discussed in the literature. We believe that in order to adequately address those issues and engage in a productive normative discussion it is necessary to examine key concepts and categories. One such category is anthropomorphism. It is a well-known fact that AI's functionalities and innovations are often anthropomorphized (i.e., described and conceived as characterized by human traits). The general public's anthropomorphic attitudes and some of their ethical consequences (particularly in the context of social robots and their interaction with humans) have been widely discussed in the literature. However, how anthropomorphism permeates AI research itself (i.e., in the very language of computer scientists, designers, and programmers), and what the epistemological and ethical consequences of this might be have received less attention. In this paper we explore this issue. We first set the methodological/theoretical stage, making a distinction between a normative and a conceptual approach to the issues. Next, after a brief analysis of anthropomorphism and its manifestations in the public, we explore its presence within AI research with a particular focus on brain-inspired AI. Finally, on the basis of our analysis, we identify some potential epistemological and ethical consequences of the use of anthropomorphic language and discourse within the AI research community, thus reinforcing the need of complementing the practical with a conceptual analysis.

Fothergill, T., Knight, W., Stahl, B.C. and Ulnicane, I. (2019) Intersectional observations of the Human Brain Project's approach to sex and gender. Journal of Information, Communication and Ethics in Society 17(2) 128-144, doi: 10.1108/JICES-11-2018-0091

Abstract: This paper aims to critically assess approaches to sex and gender in the Human Brain Project (HBP) as a large ICT project case study using intersectionality. The strategy of

the HBP is contextualised within the wider context of the representation of women in ICT, and critically reflected upon from an intersectional standpoint. The policy underpinning the approach deployed by the HBP in response to these issues parallels Horizon 2020 wording and emphasises economic outcomes, productivity and value, which aligns with other "equality" initiatives influenced by neoliberalised versions of feminism. The frameworks underpinning the HBP approach to sex and gender issues present risks with regard to the further entrenchment of present disparities in the ICT sector, may fail to acknowledge systemic inequalities and biases and ignore the importance of intersectionality. Shortcomings of the approach employed by the HBP up to March, 2018 included aspects of each of these risks, and replicated problematic understandings of sex, gender and diversity. This paper is the first to use an intersectional approach to issues of sex and gender in the context of large-scale ICT research. Its value lies in raising awareness, opening a discursive space and presenting opportunities to consider and reflect upon potential, contextualised intersectional solutions to such issues.

Abstract: Current discussions of the ethical aspects of big data are shaped by concerns regarding the social consequences of both the widespread adoption of machine learning and the ways in which biases in data can be replicated and perpetuated. We instead focus here on the ethical issues arising from the use of big data in international neuroscience collaborations. Neuroscience innovation relies upon neuroinformatics, large-scale data collection and analysis enabled by novel and emergent technologies. Each step of this work involves aspects of ethics, ranging from concerns for adherence to informed consent or animal protection principles and issues of data re-use at the stage of data collection, to data protection and privacy during data processing and analysis, and issues of attribution and intellectual property at the data-sharing and publication stages. Significant dilemmas and challenges with far-reaching implications are also inherent, including reconciling the ethical imperative for openness and validation with data protection compliance and considering future innovation trajectories or the potential for misuse of research results. Furthermore, these issues are subject to local interpretations within different ethical cultures applying diverse legal systems emphasising different aspects. Neuroscience big data require a concerted approach to research across boundaries, wherein ethical aspects are integrated within a transparent, dialogical data governance process. We address this by developing the concept of "responsible data governance," applying the principles of Responsible Research and Innovation (RRI) to the challenges presented by the governance of neuroscience big data in the Human Brain Project (HBP).

Larrivee, D., Farisco, M. (2019) Realigning the Neural Paradigm for Death. Journal of Bioethical Inquiry, doi: [10.1007/s11673-019-09915-3](#)

Abstract: Whole brain failure constitutes the diagnostic criterion for death determination in most clinical settings across the globe. Yet the conceptual foundation for its adoption was slow to emerge, has evoked extensive scientific debate since inception, underwent policy revision, and remains contentious in praxis even today. Complications result from the need to relate a unitary construal of the death event with an adequate account of what is meant by living, and what distinguishes the human state in particular. Advances in the neuroscience of higher human faculties, such as the self, personal identity, and consciousness, and dynamical philosophy of science accounts, however, are yielding a portrait of higher order global integration shared between body and brain. Such conceptual models of integration challenge a praxis relying exclusively on a neurological criterion for death.

Northoff, G., Tumati, S. (2019) "Average is good, extremes are bad"-Non-linear inverted U-shaped relationship between neural mechanisms and functionality of mental features. Neuroscience and Behavioural Reviews, 104, doi: [10.1016/j.neubiorev.2019.06.030](#)

Abstract: Traditionally, studies emphasize differences in neural measures between pathological and healthy groups, assuming a binary distinction between the groups, and a linear relationship between neural measures and symptoms. Here, we present four examples that show a continuous relation across the divide of normal and pathological states between neural measures and mental functions. This relation can be characterized by a non-linear inverted-U shaped curve. Along this curve, mid-range or average expression of a neural measure is as- sociated with optimal function of a mental feature (in healthy states), whereas extreme expression, either high or low, is associated with sub-optimal function, and occurs in different neural disorders. Neural expression between the optimal or intermediate and pathological or extreme values is associated with sub-optimal function and at- risk mental states. Thus, this model of neuro-mental relationship can be summarized as "average is good, ex- tremes are bad". By focussing on neuro-mental relationships, this model can facilitate the transition of psychiatry from a categorical to a dimensional and individualized approach needed in the era of precision medicine.

Northoff, G. Wainio-Theberge, S., Evers, K. (2019). Is temporo-spatial dynamics the "common currency" of brain and mind? In Quest of "Spatiotemporal Neuroscience". Physics of Life Review, doi: [10.1016/j.plrev.2019.05.002](#)

Abstract: Neuroscience has made considerable progress in unravelling the neural correlates of mental phenomena like self, consciousness, and perception. However, the "common currency" shared between neuronal and mental activity, brain and mind, remains yet unclear. In this article, we propose that the dynamics of time and space provides a "common currency" that connects neuronal and mental features. Time and space are here understood in a dynamic context (as in contemporary physics): that is, in terms of the way the brain's

spontaneous activity constructs its spatial and temporal relationships, for instance in terms of functional connectivity and different
frequencies of fluctuations. Recruiting recent empirical evidence, we show that the different ways in which the spontaneous activity constructs its "inner time and space" are manifested in distinct mental features. Specifically, we demonstrate how spatiotemporal mechanisms like spatiotemporal repertoire, integration, and speed yield mental features like consciousness, self, and time speed perception. The focus on the brain's spatiotemporal mechanisms entails what we describe as "Spatiotemporal Neuroscience". Spatiotemporal Neuroscience conceives neuronal activity in terms of its temporo-spatial dynamics rather than its various functions (e.g., cognitive, affective, social, etc.) as in other branches of neuroscience (as distinguished from Cognitive, Affective, Cultural, Social, etc. Neuroscience). That allows Spatiotemporal Neuroscience to take into view the so-called 'spatio-temporality' of mental features including their non-causal, intrinsic and transformative relationship with neuronal features. In conclusion, Spatiotemporal Neuroscience opens the door to investigate and ultimately reveal the brain's own temporo-spatial dynamics as the hitherto missing "common currency" of neuronal and mental features.

Pennartz, C., Farisco, M., Evers, K. (2019) Indicators and criteria of consciousness in animals and intelligent machines: an inside-out approach. Frontiers in Systems Neuroscience. 25, doi: 10.3389/fnsys.2019.00025

Abstract: In today's society it becomes increasingly important to assess which non-human and non-verbal beings possess consciousness. This review aims to delineate criteria for consciousness especially in animals, while also taking into account intelligent artifacts. First, we circumscribe what we mean with "consciousness" and describe key features of subjective experience: qualitative richness, situatedness, intentionality and interpretation, integration and the combination of dynamic and stabilizing properties. We argue that consciousness has a biological function, which is to present the subject with a multimodal, situational survey of the surrounding world and body, subserving complex decision-making and goal-directed behavior. This survey reflects the brain's capacity for internal modelling of external events underlying changes in sensory state. Next, we follow an inside-out approach: how can the features of conscious experience, correlating to mechanisms inside the brain, be logically coupled to externally observable ('outside') properties? Instead of proposing criteria that would each define a "hard" threshold for consciousness, we outline six indicators: (i) goal-directed behavior and model-based learning, (ii) anatomic and physiological substrates for generating integrative multimodal representations, (iii) psychometrics and meta-cognition, (iv) episodic memory, (v) susceptibility to illusions and multistable perception, and (vi) specific visuospatial behaviors. Rather than emphasizing a particular indicator as being decisive, we propose that the consistency amongst these indicators can serve to assess consciousness in particular species. The integration of scores on the various indicators yields an overall, graded criterion for consciousness, somewhat comparable to the Glasgow Coma Scale for unresponsive patients.

When considering theoretically derived measures of consciousness, it is argued that their validity should not be assessed on the basis of a single quantifiable measure, but requires cross-examination across multiple pieces of evidence, including the indicators proposed here. Current intelligent machines, including deep learning neural networks and agile robots, are not indicated to be conscious yet. Instead of assessing machine consciousness by a brief Turing-type of test, evidence for it may gradually accumulate when we study machines ethologically and across time, considering multiple behaviors that require flexibility, improvisation, spontaneous problem-solving and the situational conspectus typically associated with conscious experience.

Salles, A., Bjaalie, J., Evers, K., Farisco, M., Fothergill, T. Guerrero, M., Maslen, H., Muller, J., Prescott, T., Stahl, B. C., Walter, H., Zilles, K., Amunts, K. (2019) The Human Brain Project: Responsible Brain Research for the Benefit of Society. Neuron; 101:3; doi: 10.1016/j.neuron.2019.01.005

Abstract: Recognizing that its research may raise various ethical, social, and philosophical issues, the HBP has made the identification, examination, and management of those issues a top priority. The Ethics and Society subproject is part of the core research project.

Salles, A., Evers, K., Farisco, M. (2019) The need for a conceptual expansion of neuroethics. AJOB Neuroscience; 10:3, doi: 10.1080/21507740.2019.1632972

Abstract: In "Neuroethics at 15: The Current and Future Environment for Neuroethics" the Emerging Issues Task Force of the INS provides an overview of the current and future topics for neuroethics and the foreseeable challenges that the field will face. The authors note that these challenges, emerging both at individual, societal, and often global levels, are importantly connected to increasing knowledge of the brain and neurotechnical capabilities, to increasing awareness of value diversity and of the need to attend to a global landscape, and to novel applications (commercial, military, governmental) of neuroscientific findings. The overarching theme, the authors note, is expansion. In this commentary we focus on the fourth needed expansion: an expansion in how neuroethics and its methodologies are conceived and how neuroethical issues should be approached. Accordingly, we explore the key role that conceptual analysis plays in normative discussions, in refining our empirical knowledge, and in fostering a clearer and more reliable vision on how to respond the many philosophical issues raised by neuroscientific knowledge and neurotechnologies.

Salles, A., Evers, K., Farisco, M. (2019) Neuroethics and Philosophy in Responsible Research and Innovation: The Case of the Human Brain Project. Neuroethics, 12, 201–211, doi: 10.1007/s12152-018-9372-9

Abstract: Responsible Research and Innovation (RRI) is an important ethical, legal, and political theme for the European Commission. Although variously defined, it is generally understood as an interactive process that engages social actors, researchers, and innovators

who must be mutually responsive and work towards the ethical permissibility of the relevant research and its products. The framework of RRI calls for contextually addressing not just research and innovation impact but also the background research process, specially the societal visions underlying it and the norms and priorities that shape scientific agendas. This requires the integration of anticipatory, inclusive, and responsive dimensions, and the nurturing of a certain type of reflexivity among a variety of stakeholders, from scientists to funders. In this paper, we do not address potential limitations but focus on the potential contribution of philosophical reflection to RRI in the context of the Ethics and Society subproject of the Human Brain Project (HBP). We show how the type of conceptual analysis provided by philosophically oriented approaches theoretically and ethically broadens research and innovation within the HBP. We further suggest that overt inclusion of philosophical reflection can promote the aims and objectives of RRI.

Stahl, B. C., Akintoye, S., Fothergill, B. T., Guerrero, M., Knight, W. and Ulnicane, I. (2019) Beyond Research Ethics: Dialogues in Neuro-ICT Research. Front. Hum. Neurosci. 13:105. doi: 10.3389/fnhum.2019.00105

Abstract: The increasing use of information and communication technologies (ICTs) to help facilitate neuroscience adds a new level of complexity to the question of how ethical issues of such research can be identified and addressed. Current research ethics practice, based on ethics reviews by institutional review boards (IRB) and underpinned by ethical principlism, has been widely criticized. In this article, we develop an alternative way of approaching ethics in neuro-ICT research, based on discourse ethics, which implements Responsible Research and Innovation (RRI) through dialogues. We draw on our work in Ethics Support, using the Human Brain Project (HBP) as empirical evidence of the viability of this approach.

Scalabrini, A., Ebisch, S. J. H., Huang, Z., Plinio, S. D., Perrucci, M. G., Romani, G. L., Mucci, C., Northoff, G. (2019) Spontaneous Brain Activity Predicts Task-Evoked Activity During Animate Versus Inanimate Touch Cerebral Cortex; bhy340, doi: 10.1093/cercor/bhy340

Abstract: The spontaneous activity of the brain is characterized by an elaborate temporal structure with scale-free properties as indexed by the power law exponent (PLE). We test the hypothesis that spontaneous brain activity modulates task-evoked activity during interactions with animate versus inanimate stimuli. For this purpose, we developed a paradigm requiring participants to actively touch either animate (real hand) or inanimate (mannequin hand) stimuli. Behaviorally, participants perceived the animate target as closer in space, temporally more synchronous with their own self, and more personally relevant, compared with the inanimate. Neuronally, we observed a modulation of task-evoked activity by animate versus inanimate interactions in posterior insula, in medial prefrontal cortex, comprising anterior cingulate cortex, and in medial superior frontal gyrus. Among these regions, an increased functional connectivity was shown between posterior insula and perigenual anterior cingulate cortex (PACC) during animate compared with inanimate

interactions and during resting state. Importantly, PLE during spontaneous brain activity in PACC correlated positively with PACC task-evoked activity during animate versus inanimate stimuli. In conclusion, we demonstrate that brain spontaneous activity in PACC can be related to the distinction between animate and inanimate stimuli and thus might be specifically tuned to align our brain with its animate environment.

Wolff, A., Yao, L. Gomez-Pilar, J., Shoaran, M., Jiang, N., Northoff G. (2019) Neural variability quenching during decision-making: Neural individuality and its prestimulus complexity. Neuroimage; 192. doi: 10.1016/j.neuroimage.2019.02.070

Abstract: The spontaneous activity of the brain interacts with stimulus-induced activity which is manifested in event-related amplitude and its trial-to-trial variability (TTV). TTV describes the variability in the amplitude of the stimulus-evoked response across trials, and it is generally observed to be reduced, or quenched. While such TTV quenching has been observed on both the cellular and regional levels, its exact behavioral relevance and neuronal basis remains unclear. Applying a novel paradigm for testing neural markers of individuality in internally-guided decision-making, we here investigated whether TTV (i) represents an individually specific response by comparing individualized vs shared stimuli; and (ii) is mediated by the complexity of prestimulus activity as measured by the Lempel-Ziv Complexity index (LZC). We observed that TTV - and other electrophysiological markers such as ERP, ERSP, and ITC - showed first significant differences between individualized and shared stimuli (while controlling for task-related effects) specifically in the alpha and beta frequency bands, and secondly that TTV in the beta band correlated significantly with reaction time and eLORETA activity. Moreover, we demonstrate that the complexity (LZC) of neuronal activity is higher in the prestimulus period while it decreases during the poststimulus period, with the former also correlating specifically with poststimulus individualized TTV in alpha (but not with shared TTV). Together, our results show that the TTV represents a marker of 'neural individualization' which, being related to internal processes on both neural and psychological levels, is mediated by the information complexity of prestimulus activity. More generally, our results inform the pre-post-stimulus dynamics of rest-stimulus interaction, which is a basic and ubiquitous neural phenomenon in the brain and highly relevant for mental features including their individuality.

Wolff, A., Di Giovanni, D. A., Gómez-Pilar, J., Nakao, T., Huang, Z., Longtin, A., Northoff, G. (2019) The temporal signature of self: Temporal measures of resting-state EEG predict self-consciousness. Hum Brain Mapp.; 40(3):789-803, doi: 10.1002/hbm.24412

Abstract: The spontaneous activity of the brain interacts with stimulus-induced activity which is manifested in event-related amplitude and its trial-to-trial variability (TTV). TTV describes the variability in the amplitude of the stimulus-evoked response across trials, and it is generally observed to be reduced, or quenched. While such TTV quenching has been observed on both the cellular and regional levels, its exact behavioral relevance and neuronal

basis remains unclear. Applying a novel paradigm for testing neural markers of individuality in internally-guided decision-making, we here investigated whether TTV (i) represents an individually specific response by comparing individualized vs shared stimuli; and (ii) is mediated by the complexity of prestimulus activity as measured by the Lempel-Ziv Complexity index (LZC). We observed that TTV - and other electrophysiological markers such as ERP, ERSP, and ITC – showed first significant differences between individualized and shared stimuli (while controlling for task-related effects) specifically in the alpha and beta frequency bands, and secondly that TTV in the beta band correlated significantly with reaction time and eLORETA activity. Moreover, we demonstrate that the complexity (LZC) of neuronal activity is higher in the prestimulus period while it decreases during the poststimulus period, with the former also correlating specifically with poststimulus individualized TTV in alpha (but not with shared TTV). Together, our results show that the TTV represents a marker of 'neural individualization' which, being related to internal processes on both neural and psychological levels, is mediated by the information complexity of prestimulus activity. More generally, our results inform the pre-post-stimulus dynamics of rest-stimulus interaction, which is a basic and ubiquitous neural phenomenon in the brain and highly relevant for mental features including their individuality.

Cahn-Fuller, K., Shook, J., Giordano, J. (2018) Moral mentation: What neurocognitive studies of psychopathy may really offer the internalism/externalism debate. J Cogn Neuroethics; 5(2): 1-20.

Abstract: There is ongoing debate in moral philosophy and psychology about whether moral judgments necessarily motivate an agent's actions (what is known as internalism), or if moral judgments do not inherently motivate an agent to perform moral actions (what is known as externalism). Investigations focusing upon brain structures and functions that are involved in moral cognition (and which constitute an aspect of the discipline of neuroethics) have been leveraged in attempt to resolve this debate. In this way, neuroethics – as the neuroscientific studies of moral thought and actions - is important to informing both philosophical notions of morality, as well as the practices of forensic psychiatry and the legal system, which increasingly look to empirical data about psychopathy to understand and mitigate criminal behavior. However, brain science is unlikely to provide definitive answers to the conceptual questions that drive the current internalism/externalism debate. Thus, moving forward, it will be necessary to carefully define the questions that neuroscience is employed to address and answer, and equally vital to ensure that empirical findings are not distorted to support preconceived theoretical assumptions. In this way, neuroscientific investigations – and neuroethics - can be employed in a conciliatory way. Not only to balance views of processes operative in moral cognition, but to bring together the sciences and humanities to both address questions about human morality, and iteratively raise ethical,

legal and social questions about what research findings actually mean, and what medicine – and societies – will effect through the use of such information and meanings.

Conio, B., Magioncalda, P., Martino, M., Tumati, S., Capobianco, L., Escelsior, A., Adavastro, G., Russo, D., Amore, M., Inglese, M., Northoff G. (2018) Opposing patterns of neuronal variability in the sensorimotor network mediate cyclothymic and depressive temperaments. Hum Brain Mapp., doi: 10.1002/hbm.24453

Abstract: Affective temperaments have been described since the early 20th century and may play a central role in psychiatric illnesses, such as bipolar disorder (BD). However, the neuronal basis of temperament is still unclear. We investigated the relationship of temperament with neuronal variability in the resting state signal—measured by fractional standard deviation (fSD) of BloodOxygen-Level Dependent signal—of the different large-scale networks, that is, sensorimotor network (SMN), along with default-mode, salience and central executive networks, in standard frequency band (SFB) and its sub-frequencies slow4 and slow5, in a large sample of healthy subject (HC, n = 109), as well as in the various temperamental subgroups (i.e., cyclothymic, hyperthymic, depressive, and irritable). A replication study on an independent dataset of 121 HC was then performed. SMN fSD positively correlated with cyclothymic z-score and was significantly increased in the cyclothymic temperament compared to the depressive temperament subgroups, in both SFB and slow4. We replicated our findings in the independent dataset. A relationship between cyclothymic temperament and neuronal variability, an index of intrinsic neuronal activity, in the SMN was found. Cyclothymic and depressive temperaments were associated with opposite changes in the SMN variability, resembling changes previously described in manic and depressive phases of BD. These findings shed a novel light on the neural basis of affective temperament and also carry important implications for the understanding of a potential dimensional continuum between affective temperaments and BD, on both psychological and neuronal levels.

Datta, S. (2018) An endogenous explanation of growth: direct-to-consumer stem cell therapies in PR China, India and the USA. Regenerative Medicine; 13, no. 5559-579. doi: 10.2217/rme-2017-0144

Abstract: The recent expansion of direct-to-consumer stem cell therapies (DSCTs) across nations where medical mal- practice laws are the strongest globally challenges the causal assumption that low regulatory standards in developing countries bolster DSCTs. Drawing on firm-level data of existing biopharmaceuticals, approved stem cell therapies (SCTs) and DSCT clinics across the USA, PR China and India, this paper provides an innovation studies perspective of the ways in which the paradigmatic shift in fundamental knowledge production – from in vitro to in vivo stem cells – is transforming SCT discovery and delivery. It argues that the endogenous and inherent disruptive attributes of SCTs, rather than exogenous conditions like regulations, provide a substantive explanation for the recent

expansion of DSCTs and urges regulatory adaptation to endogenous imperatives for effective governance of SCTs.

DiEuliis, D., Lutes, C., Giordano, J. (2018) Biodata risks and synthetic biology: A critical juncture. J Bioterrorism Biodef; 9(1): 2-14. doi: [10.4172/2157-2526.1000159](10.4172/2157-2526.1000159).

Abstract: The tools of synthetic biology and the life sciences are rapidly advancing, as the ability to apply classical engineering to biological systems creates increasing possibilities for innovations in health and medicine, materials science, energy and agriculture. Intrinsic to these capabilities is the mounting 'digitization of biology', as the genetic code and its related metadata (including translated proteins, associated functions, herein referred to as "biodata") are amassed in order to engineer biology for specific purposes. The full spectrum of risks associated with the compilation and use of a wide range of biodata has not been fully identified or comprehensively understood. Further, divergences in traditional attitudes about security among disciplines, namely, biological sciences, engineering, information technology, and data science, complicate discussions on approaches to risk mitigation. To provide a more unified perspective and clarity, we propose that there are unique risks associated with the digitization of biology, represented by overlapping concerns of biosecurity and privacy. We discuss these in three categories of risk: 1) pathogen risks; 2) manufacturing risks, and 3) risks to individual privacy that can allow human harms. Further, we note that there is insufficient address or treatment of these risks in the formulation of ethics, policy and governance. Mitigation of risks will require characterization of all three spheres of risk, acknowledgement that they may require different solutions, and engagement of divergent disciplines and stakeholders to design solutions.

Farisco, M., Evers, K., Changeux, J.-P. (2018) Drug addiction: from neuroscience to ethics. Frontiers in Psychiatry; 9: 595, doi: 10.3389/fpsyt.2018.00595

Abstract: In the present paper, we suggest a potential new ethical analysis of addiction focusing on the relationship between aware and unaware processing in the brain. We take the case of the opioids epidemics to argue that a consideration of both aware and unaware processing provides a more comprehensive ethical framework to discuss the ethical issues raised by addiction. Finally, our hypothesis is that in addition to identified Central Nervous System's neuronal/neurochemical factors contributing to addictive dynamics, the socio-economic status plays a causal role through epigenetic processes, originating the need for additional reward in the brain. This provides a strong base for a socio-political form of responsibility for preventing and managing addiction crisis.

Farisco, M., Hellgren, Kotaleski J., Evers, K. (2018) Large-scale brain simulation and disorders of consciousness: Mapping technical and conceptual issues. Frontiers in Psychology; 9: 585 doi:

Abstract: Modeling and simulations have gained a leading position in contemporary attempts to describe, explain, and quantitatively predict the human brain's operations. Computer models are highly sophisticated tools developed to achieve an integrated knowledge of the brain with the aim of overcoming the actual fragmentation resulting from different neuroscientific approaches. In this paper we investigate the plausibility of simulation technologies for emulation of consciousness and the potential clinical impact of large- scale brain simulation on the assessment and care of disorders of consciousness (DOCs), e.g., Coma, Vegetative State/Unresponsive Wakefulness Syndrome, Minimally Conscious State. Notwithstanding their technical limitations, we suggest that simulation technologies may offer new solutions to old practical problems, particularly in clinical contexts. We take DOCs as an illustrative case, arguing that the simulation of neural correlates of consciousness is potentially useful for improving treatments of patients with DOCs.

Farisco, M., Salles, A., Evers, K. (2018) Neuroethics: A Conceptual Approach. Cambridge Quarterly of Healthcare Ethics; 27(4): 717-727, doi:

Abstract: In this paper, we start by identifying three main neuroethical approaches: neurobioethics, empirical neuroethics, and conceptual neuroethics. We focus on conceptual approaches which generally emphasize the need to develop and use a methodological *modus operandi* for fruitfully linking scientific (i.e., neuroscience) and philosophical (i.e., ethics) interpretations. We explain and assess the value of conceptual neuroethics approaches and explain and defend one such approach that we believe is particularly fruitful to address the various issues raised by neuroscience: fundamental neuroethics.

Giordano, J. and Evers, K. (2018), "Dual Use in Neuroscientific and Neurotechnological Research: A Need for Ethical Address and Guidance", Ethics and Integrity in Health and Life Sciences Research (Advances in Research Ethics and Integrity, Vol. 4), Emerald Publishing Limited, pp. 129-145. ISBN: 978-1-78743-572-8, eISBN: 978-1-78743-571-1, ISSN: 2398-6018, doi:

Abstract: Extant and newly developing techniques and technologies generated by research in brain sciences are characteristically employed in clinical medicine. However, the increasing capabilities conferred by these approaches to access, assess and affect cognition, emotion and behavior render them viable and attractive for engagement beyond the clinical realm, in what are referred to as "dual-use" applications. Definitions of what constitutes dual-use research and applications can vary so as to include utilization in the public sector for lifestyle or wellness purposes – with growing participation of a do-it-yourself (i.e., biohacking) community, and an iterative interest and use in military and

warfare operations. Such uses can pose risks to public safety, and challenge research ethics' principled imperative for non-harm (although while complete avoidance of any harm may be in reality impossible, certainly any/all harms incurred should be minimized). Thus, it is important to both clarify the construct of dual-use brain research and address the ethical issues that such research fosters. This chapter provides a review and clarification of the concept of dual-use brain science, and describes how current and emerging tools and techniques of brain research are actually or potentially employed in settings that threaten public health and incur ethical concerns. Key ethical issues are addressed, and recommendations for ethical guidance of potentially dual-use research are proposed.

Global Neuroethics Summit Delegates, Rommelfanger, K., Jeong, Sung-Jin, Ema A., Fukushi T., Kasai K., Ramos, K. Salles, A., Singh, I. (2018) Neuroethics Questions to Guide Ethical Research in the International Brain Initiatives. Neuron; 100, doi: 10.1016/j.neuron.2018.09.021

Abstract: Increasingly, national governments across the globe are prioritizing investments in neuroscience. Currently, seven active or in-development national-level brain research initiatives exist, spanning four continents. Engaging with the underlying values and ethical concerns that drive brain research across cultural and continental divides is critical to future research. Culture influences what kinds of science are supported and where science can be conducted through ethical frameworks and evaluations of risk. Neuroscientists and philosophers alike have found themselves together encountering perennial questions; these questions are engaged by the field of neuroethics, related to understanding of the nature of the self and identity, the existence and meaning of free will, defining the role of reason in human behavior, and more. With this Perspective article, we aim to prioritize and advance to the foreground a list of neuroethics questions for neuroscientists operating in the context of these international brain initiatives.

Stahl, B. C. and Wright, D. (2018) Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation. IEEE Security and Privacy 16(3): 26-33. doi: [10.1109/MSP.2018.2701164](10.1109/MSP.2018.2701164)

Abstract: Emerging combinations of artificial intelligence, big data, and the applications these enable are receiving significant media and policy attention. Much of the attention concerns privacy and other ethical issues. In our article, we suggest that what is needed now is a way to comprehensively understand these issues and find mechanisms of addressing them that involve stakeholders, including civil society, to ensure that these technologies' benefits outweigh their disadvantages. We suggest that the concept of responsible research and innovation (RRI) can provide the framing required to act with a view to ensuring that the technologies are socially acceptable, desirable, and sustainable. We draw from our work on

the Human Brain Project, one potential driver for the next generation of these technologies, to discuss how RRI can be put in practice.

Stahl, B. C., Rainey, S., Harris, E. and Fothergill, B. T. (2018) The Role of Ethics in the Data Governance of Large Neuro-ICT Projects. Journal of the American Medical Informatics Association, 25 (8), pp. 1099-1107 doi: 10.1093/jamia/ocy040

Abstract: We describe current practices of ethics-related data governance in large neuro-ICT projects, identify gaps in current practice, and put forward recommendations on how to collaborate ethically in complex regulatory and normative contexts. Materials and Methods We undertake a survey of published principles of data governance of large neuro-ICT projects. This grounds an approach to a normative analysis of current data governance approaches. Results Several ethical issues are well covered in the data governance policies of neuro-ICT projects, notably data protection and attribution of work. Projects use a set of similar policies to ensure users behave appropriately. However, many ethical issues are not covered at all. Implementation and enforcement of policies remain vague. Conclusion The data governance policies we investigated indicate that the neuro-ICT research community is currently close-knit and that shared assumptions are reflected in infrastructural aspects. This explains why many ethical issues are not explicitly included in data governance policies at present. With neuro-ICT research growing in scale, scope, and international involvement, these shared assumptions should be made explicit and reflected in data governance.

Suskin, Z. D., Giordano, J. J. (2018) Body -to-head transplant; a "caputal" crime? Examining the corpus of ethical and legal issues. Philos Ethics Humanit Med 13(1), doi: 10.1186/s13010-018-0063-2

Abstract Neurosurgeon Sergio Canavero proposed the HEAVEN procedure - i.e. head anastomosis venture - several years ago and has recently received approval from the relevant regulatory bodies to perform this body-head transplant (BHT) in China. The BHT procedure involves attaching the donor body (D) to the head of the recipient (R) and discarding the body of R and head of D. Canavero's proposed procedure will be incredibly difficult from a medical standpoint. Aside from medical doubt, the BHT has been met with great resistance from many, if not most bio- and neuroethicists.Given both the known challenges and unknown outcomes of HEAVEN, several important neuroethical and legal questions have emerged should Canavero be successful, including: (1) What are the implications for transplantology in the U.S., inclusive of issues of expense, distributive justice, organizational procedures, and the cost(s) of novel insight(s)? (2) How do bioethical and neuroethical principles, and legal regulations of human subject research apply? (3) What are the legal consequences for Canavero (or any other surgeon) performing a BHT? (4) What are the tentative implications for the metaphysical and legal identity of R should they survive post-BHT? These questions are analyzed, issues are identified, and several solutions are proposed in an attempt to re-configure HEAVEN into a safe, clinically effective, and thus

(more) realistically viable procedure.Notably, the permissibility of conducting the BHT in China fosters additional, important questions, focal to (1) whether Western ethics and professional norms be used to guide the BHT - or any neuroscientific research and its use - in non-Western countries, such as China; (2) if the models of responsible conduct of research are identical, similar, or applicable to the intent and conduct of research in China; and (3) what economic and political implications (for China and other countries) are fostered if/when such avant garde techniques are successful.These questions are discussed as a further impetus to develop a globally applicable neuroethical framework that would enable both local articulation and cosmopolitan inquiry and oversight of those methods and approaches deemed problematic, if and when rendered in more international settings.

Weinberger, A. B., Cortes, R. A., Green, A. E., Giordano, J. (2018) Neuroethical and social implications of using transcranial electrical stimulation to augment creative cognition. Creativity Res J; 30(3): 249-255, doi: 10.1080/10400419.2018.1488199

Abstract: Recent research indicates that transcranial electrical stimulation (tES) of specific brain regions can successfully improve various forms of creative cognition. Although the endeavor to increase human creative capacity is intriguing from a neuroscientific perspective, and of interest to the general public, it raises numerous neuroethico-legal and social issues (NELSI). This review explores these issues by considering (a) whether using brain stimulation to improve creative cognition qualifies as a 'treatment' or an 'enhancement,' (b) how direct-to-consumer (DTC) and do-it-yourself (DIY) use of tES should be regarded and regulated, and (c) what the developing landscape of creativity-related neurostimulation could (and should) become.

Zamora, A. R., Giordano, J., Gunduz, A., et al. (2018) Evolving applications, technical challenges and future opportunities in neuromodulation. Front Neurosci; 11:734, doi: 10.3389/fnins.2017.00734

Abstract: The annual Deep Brain Stimulation (DBS) Think Tank provides a focal opportunity for a multidisciplinary ensemble of experts in the field of neuromodulation to discuss advancements and forthcoming opportunities and challenges in the field. The proceedings of the fifth Think Tank summarize progress in neuromodulation neurotechnology and techniques for the treatment of a range of neuropsychiatric conditions including Parkinson's disease, dystonia, essential tremor, Tourette syndrome, obsessive compulsive disorder, epilepsy and cognitive, and motor disorders. Each section of this overview of the meeting provides insight to the critical elements of discussion, current challenges, and identified future directions of scientific and technological development and application. The report addresses key issues in developing, and emphasizes major innovations that have occurred during the past year. Specifically, this year's meeting focused on technical developments in DBS, design considerations for DBS electrodes, improved sensors, neuronal signal processing, advancements in development and uses of responsive DBS (closed-loop

systems), updates on National Institutes of Health and DARPA DBS programs of the BRAIN initiative, and neuroethical and policy issues arising in and from DBS research and applications in practice.

Akram, F., Giordano, J., (2017) Research domain criteria as psychiatric nosology: Conceptual, practical and neuroethical implications. Camb Q Health Care Ethics- Clin Neuroethics; 26(4): 592-601, doi: 10.1017/S096318011700010X

Abstract: Diagnostic classification systems in psychiatry have continued to rely on clinical phenomenology despite limitations inherent to that approach. In view of these limitations and recent progress in neuroscience, the National Institute of Mental Health (NIMH) has initiated the Research Domain Criteria (RDoC) project in order to develop a more neuroscientifically-based system of characterizing and classifying psychiatric disorders. The RDoC initiative aims to transform psychiatry into an integrative science of psychopathology in which mental illnesses will be defined as involving putative dysfunctions in neural nodes and networks. However, conceptual, methodological and Neuroethical and social issues inherent to and/or derived from the use of RDoC need to be addressed before any attempt at implementing use in clinical psychiatry. This essay describes current progress in RDoC, defines key technical, neuroethical and social issues generated by RDoC adoption and use, and posits key questions that must be addressed and resolved if RDoC are to be employed for psychiatric diagnoses and therapeutics. Specifically, we posit that objectivization of complex mental phenomena may raise ethical questions about autonomy, the value of subjective experience, what constitutes normality, a disorder, and what represents a treatment, enablement and/or enhancement. Ethical issues may also arise from the (mis)use of biomarkers and phenotypes in predicting and treating mental disorders, and what such definitions, predictions and interventions portend for concepts and views of sickness, criminality, professional competency and social functioning. Given these issues, we offer that a preparatory neuroethical framework is required to define and guide the ways in which RDoC- oriented research can – and arguably should - be utilized in clinical psychiatry, and perhaps more broadly, in the social sphere.

Evers, K., Changeux, J.P., (2017) Proactive epigenesis and ethical innovation: A neuronal hypothesis for the genesis of ethical rules. Response by the authors, EMBO reports; 18(8): 1272, doi: 10.15252/embr.201642783

Abstract: In her commentary to our article on proactive epigenesis, Arleen Salles constructively criticizes aspects of our approach that, she suggests, merit further analyses and justification. We welcome her contributions that we consider both pertinent and useful to the development and application of our ideas. In particular, as Salles argues, core concepts should be more clearly defined, and normative claims must be justified.

Evers K., (2017) The contribution of neuroethics to international brain research initiatives, Nature Reviews Neuroscience; 18:1-2, doi: 10.1038/nrn.2016.143

Abstract: Neuroethics research can contribute a level of conceptual clarity to international brain research initiatives that is essential for their ethics management as well as for the interpretations, applications and management of their emerging neuroscientific findings.

Evers, K., Giordano, J. (2017) The utility – and use – of neurotechnology to recover consciousness: Technical and neuroethical considerations in approaching the 'hard question' of neuroscience. Front Hum Neurosci 11(564): 2-5, doi: 10.3389/fnhum.2017.00564

Abstract: In any attempt to recover a loss, it becomes important to assess what is lost, and to what extent. In clinical medicine, evaluation of loss of particular function(s) is fundamental to both diagnosis and planning possible interventions that may be restorative. The term diagnosis is etymologically derived from the Greek *diagignoskein*, "to discern." The root of the word, *gignőskein (γιγνώσκειν)* means "to learn"; the implication being that one must gain knowledge about those aspects of the thing(s) that are important for it to be distinguished, identified and characterized. If, however, the focus of discernment is the clinical assessment of consciousness in a patient who is unable to respond and/or communicate through speech or overt behavior, then the act of discernment becomes complicated, given that the cardinal characteristics of consciousness are subjectivity and self-transparency, and to that extent not viable for direct objective evaluation.

Farisco, M., Evers, K., Salles, A. (2017) The Computational Shift in Neuroscience: A Multifaceted Neuroethical Analysis, AJOB Neuroscience,; 8(1):W4-W5, doi: 10.1080/21507740.2017.1293191

Abstract: Lately it has been argued that a paradigm shift in neuroscience is necessary, a turn from a hypothesis-led approach grounded on empirical observation, to a data-led simulation modelling grounded on the computational analysis and manipulation of big data repositories. This novel approach is known as "predictive biology." It is notably exemplified by the European Human Brain Project, which aims at overcoming the existing limitations and fragmentation of contemporary neurosciences by digitalizing the massive empirical data available and developing highly sophisticated tools, such as computer models and simulations, to achieve an integrated knowledge of the brain. The development of predictive computational neuroscience is usually justified for both epistemological (i.e., the intrinsic limitation of the empirical methodology) and practical (i.e., the need for new therapeutic and clinical applications) reasons. We think that the proposed computational shift raises the need for a specific neuroethical reflection.

Lipina, S., Evers, K. (2017) Neuroscience of Childhood Poverty: Evidence of Impacts and Mechanisms as Vehicles of Dialog With Ethics, Frontiers in Psychology; 8: 61, doi: 10.3389/fpsyg.2017.00061

Abstract: Several studies have identified associations between poverty and development of self-regulation during childhood, which is broadly defined as those skills involved in cognitive, emotional, and stress self-regulation. These skills are influenced by different individual and contextual factors at multiple levels of analysis (i.e., individual, family, social, and cultural). Available evidence suggests that the influences of those biological, psychosocial, and sociocultural factors on emotional and cognitive development can vary according to the type, number, accumulation of risks, and co-occurrence of adverse circumstances that are related to poverty, the time in which these factors exert their influences, and the individual susceptibility to them. Complementary, during the past three decades, several experimental interventions that were aimed at optimizing development of self-regulation of children who live in poverty have been designed, implemented, and evaluated. Their results suggest that it is possible to optimize different aspects of cognitive performance and that it would be possible to transfer some aspects of these gains to other cognitive domains and academic achievement. We suggest that it is an important task for ethics, notably but not exclusively neuroethics, to engage in this interdisciplinary research domain to contribute analyses of key concepts, arguments, and interpretations. The specific evidence that neuroscience brings to the analyses of poverty and its implications needs to be spelled out in detail and clarified conceptually, notably in terms of causes of and attitudes toward poverty, implications of poverty for brain development, and for the possibilities to reduce and reverse these effects.

Palchik, G., Chen, C., Giordano, J. (2017) Monkey business? Development, influence and ethics of potentially dual-use brain science on the world stage, Neuroethics, 2017; 10: 1-4, doi: 10.1007/s12152-017-9308-9

Abstract: At the recent annual meeting of the International Neuroethics Society, Dr. Mu-Ming Poo, of the China Brain Science Project, provided an overview of the proposed future directions and goals of neuroscientific research in China. Through concentrated efforts of the Institute of Neuroscience of the Chinese Academy of Science (CAS), China is rapidly developing increasingly greater scientific and technological capability – and global prominence – in brain research. Poo emphasized the strong translational focus of such research in light of recognized epidemiological trends in neurological disorders. What effect might this have on a broader, international scale? This raises concerns about tacit capabilities, and to yoking NHP studies and findings to military agendas under programs of dual- or direct-use. Such endeavors are not unique to China, as many countries, including the United States, engage in brain research that is applicable to, and employed for national security In this paper, we posit that it is important to examine methods, (explicit, implicit, and possible) uses, and consequences that can be incurred from brain science. Therefore, if neuroethical address and analyses are to authentically inform international guidelines and

policies that direct the sound conduct and use of brain science, then veracity, veridicality, and clarity in the discourses and activities of neuroscience *and* neuroethics will be ever more essential- regardless of where they occur.

Racine, E., Dubljevic, V., Jox, R., Baertschi, B., Christensen, J., Farisco, M., Jotterand, F., Kahane, G., Muller, S. (2017) Can Neuroscience Contribute to Practical Ethics? A Critical Review and Discussion of the Methodological and Translational Challenges of the Neuroscience of Ethics, Bioethics; 31(5): 328-337, doi: 10.1111/bioe.12357

Abstract: Neuroethics is an interdisciplinary field that arose in response to novel ethical challenges posed by advances in neuroscience. Historically, neuroethics has provided an opportunity to synergize different disciplines, notably proposing a two-way dialogue between an 'ethics of neuroscience' and a 'neuroscience of ethics'. However, questions surface as to whether a 'neuroscience of ethics' is a useful and unified branch of research and whether it can actually inform or lead to theoretical insights and transferable practical knowledge to help resolve ethical questions. In this article, we examine why the neuroscience of ethics is a promising area of research and summarize what we have learned so far regarding its most promising goals and contributions. We then review some of the key methodological challenges which may have hindered the use of results generated thus far by the neuroscience of ethics. Strategies are suggested to address these challenges and improve the quality of research and increase neuroscience's usefulness for applied ethics and society at large. Finally, we reflect on potential outcomes of a neuroscience of ethics and discuss the different strategies that could be used to support knowledge transfer to help different stakeholders integrate knowledge from the neuroscience of ethics.

Salles, A. (2017) Proactive Epigenesis and Ethics. EMBO reports; 18(8): 1271, doi: 10.15252/embr.201744697

Abstract: A recent article by Kathinka Evers and Jean Pierre Changeux offers a new approach to the issue of moral change. They propose proactive epigenesis as a tool to communicate and establish social and ethical norms in education and upbringing so as to build better societies. In this short commentary I explain their view and then identify and explain some of the normative issues raised by their proposal. In particular, I focus on some moral claims they make that raise deep questions about justification and frameworks and thus require further discussion. Based on my analysis, I propose that the authors themselves further develop their views and elaborate on the specifically moral issues raised by their proposal and hope that their joint work on this issue inspires empirical and theoretical research from disciplines such as moral philosophy, pedagogy, and social science to further examine proactive epigenesis and the possibilities it opens for addressing moral improvement.

Shook, J. R., Giordano, J. (2017) Ethics transplants? Addressing the risks and benefits of guiding international biomedicine. AJOB-Neuroscience; 8(4): 230-232, doi: 10.1080/21507740.2017.1392377

Abstract: In this paper, we address the planned body-head transplant (B-H-T) proposed by Canavero, to occur in China later this year. We speak to a current report by Canavero and /Ren, and to a rebuttal by Wolpe, and we argue that research plans such as those exemplified by Canavero's B-H-T experiment look more like opportunism at present. All the same, the controversial issues regarding intra- and cross-cultural ethical norms and conduct do create a timely opportunity to develop a broader and deeper conception of a globally relevant neuroethics. We strongly urge that neuroethics embrace this intercultural vantage point. We assert that neuroethics need not – and should not – dictate social priorities or impose a regulatory code based on a single country's lessons learned from hard experience. Nuanced neuroethical evaluations can develop practical guidelines by asking contextual questions first and delivering recommendations second. This approach would not retard progress; it would be a positive resource to enable the highest-quality professional conduct and scientific credibility. Neurothical consultations should be cooperative, deliberative, future-oriented, and multi-national as any neuroscientific project. Neuroethical engagement should avoid inflexible absolutism as well as relativism, appreciating instead a consensus view of what constitutes good science and sound medical practice. There is no need to transplant western ethics or law into non-western countries, or vice versa. Neuroethics as an international enterprise should guide innovative brain science with sensitivity to both the exigencies of particular cultures and the contingencies of inter-cultural engagement.

Herrera-Ferra, K., Giordano, J. (2017) Recurrent violent behavior: Revised classification and implications for global psychiatry, Front Psychiatry; 8 (151): 1-4, doi: 10.3389/fpsyt.2017.00151

Abstract: In this paper, we propose that instantiating recurrent violent behavior (RVB) as a psychiatric classifier may leverage medical and psychosocial interventions in order to improve both mental health care and public safety needs, as consistent with Article 25 of the UN Declaration of Human Rights., and the WHO Mental Health Action Plan. We acknowledge that instituting RVB as a psychiatric classifier does not necessarily guarantee proper –if any- medical assessment or care, especially in developing and non-developed countries. We address distinctions between developed and non-developed countries mental health services, and seek to align our proposed use of RVB as a psychiatric classifier to be in accordance with the Global Mental Health Initiative. We address technical, as well as neuroethico-legal and social issues that such a proposal may foster, and posit approaches toward their address and possible resolution.

Kraft, C., Giordano, J. (2017) Integrating brain science and law: Neuroscientific evidence and legal perspectives on protecting individual liberties. Front Neurosci 11: 1-10, doi: 10.3389/fnins.2017.00621

Abstract: Advances in neuroscientific techniques have found increasingly broader applications, including in legal neuroscience (or "neurolaw"), where experts in the brain sciences are called to testify in the courtroom. But does the incursion of neuroscience into the legal sphere constitute a threat to individual liberties? And what legal protections are there against such threats? In this paper, we outline individual rights as they interact with neuroscientific methods. We then proceed to examine the current uses of neuroscientific evidence, and ultimately determine whether the rights of the individual are endangered by such approaches. Based on our analysis, we conclude that while federal evidence rules constitute a substantial hurdle for the use of neuroscientific evidence, more ethical safeguards are needed to protect against future violations of fundamental rights. Finally, we assert that it will be increasingly imperative for the legal and neuroscientific communities to work together to better define the limits, capabilities, and intended direction of neuroscientific methods applicable for use in law.

Raynor, S., Giordano, J. (2017) Treating Alzheimer's dementia with CT-induced low dose ionizing radiation: Problematic, yet potential for more precise inquiry. Dose Response; 15(3): 1-4, doi: 10.1177/1559325817729247

Abstract: This commentary evaluates a recent single-case study by Cuttler et al that posits that a series of computerized tomographic (CT) scans ameliorated symptoms and signs of advanced Alzheimer's dementia in an elderly female patient. The report proposes that CT scanning delivered low-dose ionizing radiation (LDIR) that activated adaptive mechanisms in the brain to induce the effects observed and reported. However, the report evidenced methodologic problems that threaten the validity and value of its approach, stated results, and conclusions. We provide discussion of these issues, with view and intent toward developing more precise investigations of the potential mechanisms and utility of LDIR in treating Alzheimer's dementia and possibly other neurodegenerative disorders.

Shook, J., Giordano, J. (2017) Moral Bioenhancement for Social Welfare: Are Civic Institutions Ready? Front. Sociol., doi: 10.3389/fsoc.2017.00021

Abstract: Positive assessments of moral enhancement too often isolate intuitive notions about its benefits apart from the relevance of surrounding society or civic institutions. If moral bioenhancement should benefit both oneself and others, it cannot be conducted apart from the enhancement of local social conditions, or the preparedness of civic institutions. Neither of those considerations has been adequately incorporated into typical neuroethical assessments of ambitious plans for moral bioenhancement. Enhancing a person to be far less aggressive and violent than an average person, what we label as "civil enhancement,"

seems to be quite moral, yet its real-world social consequences are hardly predictable. A hypothetical case about how the criminal justice system would treat an offender who already received civil enhancement serves to illustrate how civic institutions are unprepared for moral enhancement.

Aicardi, C., Del Savio, L., Dove, E.S., Lucivero, F., Tempini, N. and Prainsack, B. (2016) Emerging ethical issues regarding digital health data. On the World Medical Association Draft Declaration on Ethical Considerations Regarding Health Databases and Biobanks, Croatian Medical Journal 57, no. 2: 207-213, doi: 10.3325%2Fcmj.2016.57.207

Abstract: With the growth of data-intensive approaches in biomedical research, a rich discussion on the saliency of moral considerations pertaining to the management of personal data in health databases and biobanks has unfolded. Regional, national, and international institutions that make policy on biomedical research ethics recognize this development and are revising their position on the ethics of data-driven biomedical research to keep pace with these transformative developments. In March 2015, for example, the WMA opened a public consultation on their "Draft on ethical considerations regarding health databases and biobanks." The Declaration was available for public consultation until June 2015. While we support the codification of ethical principles for use of data in health data-bases and human biological material in biobanks, particularly given the pervasive use of digital health data, we find that the remit of the draft Declaration is unduly narrow and fails to offer meaningful advancement of the ethical principles stated in the Declaration of Helsinki. Changing practices in the collection and use of digital data require a revised framework and nomenclature regarding the norms, rules, and principles governing biomedical research. In the remainder of this article, we discuss five areas that ought to be taken into consideration in this process.

Aicardi, C. (2016) Francis Crick, cross-worlds influencer: A narrative model to historicize big bioscience. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 55: 83-95, doi: 10.1016/j.shpsc.2015.08.003

Abstract: The essay is an empirical case study of famed British scientist Francis Crick. Viewing him as a 'cross-worlds influencer' who was moreover dedicated to a cause, I have tried to understand how these two characteristics influenced the trajectory of his long career and how they shaped his contributions to the diverse research fields in which he was active and concluded that these characteristics reconfigure Crick's career into a coherent whole. First, I identify a major thread running through Crick's career: helping organise 'un-disciplined' new research fields and show that his successive choices were not serendipitous but motivated by what he construed as a crusade against 'vitalism': anti-vitalism was a defining driver of his career. I then examine how Crick put his skills as a crossworlds influencer to the service of his cause, by helping organise his chosen fields of intervention. I argue that his activities as a cross-worlds influencer were an integral part of his way of 'doing science' and

that his contributions to science, neuroscience in particular, should be re-evaluated in this light. This leads me to advance a possible strategy for historians to investigate big bioscience fields. Following Abir-Am, I propose to trace their genealogies back to the fluctuating semi-institutional gatherings and the institutional structures that sustained them. My research on Crick supports the view that such studies can bring insights into the question of why the contours of contemporary big bioscience endeavours have come to be shaped the way they are. Further, the essay provides a heuristic device for approaching these enquiries: 'follow the cross-worlds influencers' who worked to build and organise these semi-institutional gatherings and institutional structures.

Aicardi, C. and García-Sancho, M. (eds.) (2016) Towards future archives and historiographies of 'big biology'. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 55, doi: 10.1016/j.shpsc.2015.09.009

Abstract: The 20th century has seen the progressive rise of 'big science', especially after 1945. In the last seventy years, research has increasingly been conducted by large, interdisciplinary teams spanning different countries and institutions, and attracting generous funding by both public and private actors (Galison and Hevly, 1992). Recent biological research presents historically specific contours that require attention at both the level of sources and of scholarly narration (Hilgartner, 2013; Aronova, Baker and Oreskes, 2010; Davies, Frow and Leonelli, 2013). This edited volume explores whether it is desirable to transform biomedical history into a data-driven endeavour and the alternatives to this course of action. In their essays, the authors set out to explore the connections and interplay between historiographical and archival issues raised by the contemporary transformation of the life sciences into big science enterprises. Reflecting on narrative models, the nature and availability of sources, and the construction of archives, they challenge overly simplistic 'big data' strategies and propose a number of alternative methods for navigating 'big biology'.

Evers, K. (2016) Neurotechnological assessment of consciousness disorders: five ethical imperatives, Dialogues in Clinical Neuroscience;18(2): 155-162, doi: 10.31887/DCNS.2016.18.2/kevers

Abstract: Disorders of Consciousness (DOCs) cause great human suffering and material costs for society. Understanding of these disorders has advanced remarkably in recent years, but uncertainty remains with respect to the diagnostic criteria and standards for care. One of the most serious problems concerns misdiagnoses, their impact on medical decision-making, and on patients' wellbeing. Recent studies of DOC patients use neurotechnology to assess residual consciousness in DOC patients that traditional behavioural diagnostic criteria are unable to detect. The results show an urgent need to strengthen the development of new diagnostic tools and more refined diagnostic criteria. If residual consciousness may be inferred from robust and repeated results from neurotechnological communication with DOC patients, this also raises ethical challenges. With reference to the moral notions of

beneficence and fundamental rights, five ethical imperatives are here suggested in terms of diagnosis, communication, interpretation of subjective states, adaptation of living conditions and care.

Evers, K., Changeux, J-P. (2016) Proactive epigenesis and ethical innovation. A neuronal hypothesis for the genesis of ethical rules, EMBO Reports; 17 (10): 1361-1364, doi: 10.15252/embr.201642783

Abstract: During the long period of postnatal development in humans, the cerebral cortex undergoes intense synaptogenesis, which persists into adulthood. The steady interaction with the physical, social, and cultural environment drives an epigenetic selection of neuronal networks to internalize, in particular, the common cultural and ethical rules of the society to which the child and her/his family belong. Based on this knowledge, we propose the idea of proactive epigenesis to develop new ethical rules and educational approaches to influence, and constructively interact with the developing neuronal architecture of the human brain.

Farisco, M., Evers. K., Salles, A. (2016) Big Science, Brain Simulation and Neuroethics, AJOB Neuroscience; 7(1): 28–30, doi: [10.1080/21507740.2015.1135834](10.1080/21507740.2015.1135834)

Abstract: The ethical analysis of brain simulation is a relatively new field of research, which is gaining an increasing multidisciplinary interest. However, at present the debate revolves mainly around the practical concerns raised by emerging neuro-technology, concerns that are not fundamentally different from those raised by emerging technologies in general. The authors suggest that while valuable, this common approach is not enough to fully capture the issues at stake and propose the examination of conceptual understandings of the brain and of simulation in order to better grasp the ethical implications of simulation technology in particular.

Fitzgerald, D., Ilina, S. and Rose, N., (2016) Revitalizing sociology: urban life and mental illness between history and the present, British Journal of Sociology 67, no. 1: 138-160, doi: [10.1111/1468-4446.12188](10.1111/1468-4446.12188)

Abstract: This paper proposes a re-thinking of the relationship between sociology and the biological sciences. Tracing lines of connection between the history of sociology and the contemporary landscape of biology, the paper argues for a reconfiguration of this relationship beyond popular rhetorics of 'biologization' or 'medicalization'. At the heart of the paper is a claim that, today, there are some potent new frames for re-imagining the traffic between sociological and biological research - even for 'revitalizing' the sociological enterprise as such. The paper threads this argument through one empirical case: the relationship between urban life and mental illness. In its first section, it shows how this relationship enlivened both early psychiatric epidemiology, and some forms of the new discipline of sociology; it then traces the historical division of these sciences, as the

sociological investment in psychiatric questions waned, and 'the social' become marginalized within an increasingly 'biological' psychiatry. In its third section, however, the paper shows how this relationship has lately been revivified, but now by a nuanced epigenetic and neurobiological attention to the links between mental health and urban life. What role can sociology play here? In its final section, the paper shows how this older sociology, with its lively interest in the psychiatric and neurobiological vicissitudes of urban social life, can be our guide in helping to identify intersections between sociological and biological attention. With a new century now underway, the paper concludes by suggesting that the relationship between urban life and mental illness may prove a core testing-ground for a 'revitalized' sociology.

Evers, K. (2016). Can we be epigenetically proactive? In T. Metzinger & J. M. Windt (2016) (Eds). Open Mind: Philosophy and the mind sciences in the 21st century, MIT Press, Cambridge, pp. 497-518, doi: 10.15502/9783958570238

Abstract: The human brain is an essentially evaluative organ endowed with reward systems engaged in learning and memory as well as in higher evaluative tendencies. Our innate species-specific, neuronally-based identity disposes us to develop universal evaluative tendencies, such as self-interest, control-orientation, dissociation, selective sympathy, empathy, and xenophobia. The combination of these tendencies may place us in a predicament. Our neuronal identity makes us social, but also individualistic and self-projective, with an emotional and intellectual engagement that is far more narrowly focused in space and time than the effects of our actions. However, synaptic epigenesis theories of cultural and social imprinting on our brain architecture suggest that there is a possibility of culturally influencing these predispositions. In an analysis of epigenesis by selective stabilisation of synapses, I discuss the relationships between genotype and brain phenotype: the paradox of non-linear evolution between genome and brain complexity; the selection of cultural circuits in the brain during development; and the genesis and epigenetic transmission of cultural imprints. I proceed to discuss the combinatorial explosion of brain representations, and the channelling of behaviour through "epigenetic rules" and top-down control of decision-making. In neurobiological terms, these "rules" are viewed as acquired patterns of connections (scaffoldings), hypothetically stored in frontal cortex long-term memory, which frame the genesis of novel representations and regulate decision-making in a top-down manner. Against that background I propose the possibility of being epigenetically proactive, and adapting our social structures, in both the short and the long term, to benefit, influence, and constructively interact with the ever-developing neuronal architecture of our brains.

Rose, N. (2016) Reading the Human Brain How the Mind Became Legible. Body & Society, doi: 10.1177/1357034X15623363

Abstract: The human body was made legible long ago. But what of the human mind? Is it possible to 'read' the mind, for one human being to know what another is thinking or feeling, their beliefs and intentions. And if I can read your mind, how about others – could our authorities, in the criminal justice system or the security services? Some developments in contemporary neuroscience suggest the answer to this question is 'yes'. While philosophers continue to debate the mind-brain problem, a range of novel technologies of brain imaging have been used to argue that specific mental states, and even specific thoughts, can be identified by characteristic patterns of brain activation; this has led some to propose their use in practices ranging from lie detection and security screening to the assessment of brain activity in persons in persistent vegetative states. This article reviews the history of these developments, sketches their scientific and technical bases, considers some of the epistemological and ontological mutations involved, explores the ecological niches where they have found a hospitable environment, and considers some implications of this materialization of the readable, knowable, transparent mind.

Sallin, K., Lagercrantz, H., Evers, K., Engström, I., Hjern, A., Petrovic, P. (2016) Resignation Syndrome: Catatonia? Culture-Bound?, Front. Behav. Neurosci, doi: 10.3389/fnbeh.2016.00007

Abstract: Resignation syndrome (RS) designates a long-standing disorder predominately affecting psychologically traumatized children and adolescents in the midst of a strenuous and lengthy migration process. Typically, a depressive onset is followed by gradual withdrawal progressing via stupor into a state that prompts tube feeding and is characterized by failure to respond even to painful stimuli. The patient is seemingly unconscious. Recovery ensues within months to years and is claimed to be dependent on the restoration of hope to the family. Descriptions of disorders resembling RS can be found in the literature and the condition is unlikely novel. Nevertheless, the magnitude and geographical distribution stand out. Several hundred cases have been reported exclusively in Sweden in the past decade prompting the Swedish National Board of Health and Welfare to recognize RS as a separate diagnostic entity. The currently prevailing stress hypothesis fails to account for the regional distribution and contributes little to treatment. Consequently, a re-evaluation of diagnostics and treatment is required. Psychogenic catatonia is proposed to supply the best fit with the clinical presentation. Treatment response altered brain metabolism or preserved awareness would support this hypothesis. Epidemiological data suggests culture-bound beliefs and expectations to generate and direct symptom expression and we argue that culture-bound psychogenesis can accommodate the endemic distribution. Last, we review recent models of predictive coding indicating how expectation processes are crucially involved in the placebo and nocebo effect, delusions and conversion disorders. Building on this theoretical framework we propose a neurobiological model of RS in which the impact of overwhelming negative expectations are directly

causative of the down-regulation of higher order and lower order behavioral systems in particularly vulnerable individuals.

Salter, B., Zhou, Y., Datta, S., and Salter, C. (2016) Bioinformatics and the politics of innovation in the life sciences: science and the state in the UK, China, and India. Science, Technology and Human Values 41, no.5: 793–826, doi: 10.1177/0162243916631022

Abstract: The governments of China, India, and the United Kingdom are unanimous in their belief that bioinformatics should supply the link between basic life sciences research and its translation into health benefits for the population and the economy. Yet at the same time, as ambitious states vying for position in the future global bioeconomy they differ considerably in the strategies adopted in pursuit of this goal. At the heart of these differences lies the interaction between epistemic change within the scientific community itself and the apparatus of the state. Drawing on desk-based research and thirty-two interviews with scientists and policy makers in the three countries, this article analyzes the politics that shape this interaction. From this analysis emerges an understanding of the variable capacities of different kinds of states and political systems to work with sciencein harnessing the potential of new epistemic territories in global life sciences innovation.

Stahl, B. C., Timmermans, J., and Mittelstadt, B. D. (2016). The Ethics of Computing: A Survey of the Computing-Oriented Literature. ACM Computing Surveys 48(4), 55:1–55:38, doi: 10.1145/2871196

Abstract: Computing technologies and artifacts are increasingly integrated into most aspects of our professional, social, and private lives. One consequence of this growing ubiquity of computing is that it can have significant ethical implications that computing professionals need to be aware of. The relationship between ethics and computing has long been discussed. However, this is the first comprehensive survey of the mainstream academic literature of the topic. Based on a detailed qualitative analysis of the literature, the article discusses ethical issues, technologies that they are related to, and ethical theories, as well as the methodologies that the literature employs, its academic contribution, and resulting recommendations. The article discusses general trends and argues that the time has come for a transition to responsible research and innovation to ensure that ethical reflection of computing has practical and manifest consequences.

Stahl, B. C., Rainey, S., and Shaw, M. (2016) Managing Ethics in the HBP: A Reflective and Dialogical Approach. AJOB Neuroscience 7(1): 20–24, doi: 10.1080/21507740.2016.1138155

Abstract: This comment introduces ethics management practices and their development in the HBP. The HBP has been aware from the outset that it was likely to raise ethical and social concerns. In order to react to these appropriately, it has created a program of research and

other activities around responsible research and innovation (RRI). This society and ethics program, which constitutes one of the subprojects of the HBP, was inspired by the European drive to integrate RRI into all research. The comment provides and overview of ethical issues included in the HBP Ethics Map and introduces theoretical underpinnings of discourse ethics approach that form basis for ethics management within the HBP.

Farisco, M., Laureys, S., Evers, K. (2015) Externalization of Consciousness. Scientific Possibilities and Clinical Implications, in: Geyer, M., Ellenbroek, B., Marsden, C. (eds), Current Topics in Behavioural Neurosciences; 19: 205-222, doi: 10.1007/7854_2014_338

Abstract: The paper starts by analyzing recent advancements in neurotechnological assessment of residualconsciousness in patients with disorders of consciousness and in neurotechnology-mediated communication with them. Ethical issues arising from these developments are described, with particular focus on informed consent. Against this background, we argue for the necessity of further scientific efforts and ethical reflection in neurotechnological assessment of consciousness and 'cerebral communication' with verbally non-communicative patients.

Rose, N. (2015) Neuroscience and the future for mental health? Epidemiology and psychiatric sciences 25, no. 2: 95-100, doi: 10.1017/S2045796015000621

Abstract: Psychiatry is in one of its regular crises. It is a crisis of its diagnostic systems despite – perhaps because – of the recurrent claims about the extent of diagnosable 'brain disorders'. It is a crisis of its explanatory systems despite – perhaps because – of its current wager on the brain as the ultimate locus for explanations of mental disorders. It is a crisis of its therapeutic capacities despite – perhaps because – more and more people are making use of its primary mode of intervention focussed on the brain – psychiatric drugs. In this editorial, I will suggest that this triple crisis of diagnosis, explanation and therapeutics arises from the dominant reductionist approaches to the role of neurobiology in psychiatry that priorities the analysis of brain mechanisms, at the expense of an understanding of the whole living organism in its milieu, and the processes which social experience shapes neurobiology from the moment of conception if not before. I shall suggest a different approach that starts from the experience of persons coping with adversity in their forms of life. This approach does not require giving up on our search for plausible explanations of mental health problems that engage neurobiological mechanisms, but it begins from a commitment to understanding, and hence intervening in, the ways in which social adversity shapes and blights the lives of so many of our fellow citizens.

Aicardi, C. (2014) Of the Helmholtz Club, South-Californian seedbed for visual and cognitive neuroscience, and its patron Francis Crick. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 45:1-11, doi: 10.10162Fj.shpsc.2013.11.012

Abstract: Taking up the view that semi-institutional gatherings such as clubs, societies, research schools, have been instrumental in creating sheltered spaces from which many a 20th-century project-driven interdisciplinary research programme could develop and become established within the institutions of science, the paper explores the history of one such gathering from its inception in the early 1980s into the 2000s, the Helmholtz Club, which brought together scientists from such various research fields as neuroanatomy, neurophysiology, psychophysics, computer science and engineering, who all had an interest in the study of the visual system and of higher cognitive functions relying on visual perception such as visual consciousness. It argues that British molecular biologist turned South Californian neuroscientist Francis Crick had an early and lasting influence over the Helmholtz Club of which he was a founding pillar, and that from its inception, the club served as a constitutive element in his long-term plans for a neuroscience of vision and of cognition. Further, it argues that in this role, the Helmholtz Club served many purposes, the primary of which was to be a social forum for interdisciplinary discussion, where 'discussion' was not mere talk but was imbued with an epistemic value and as such, carefully cultivated. Finally, it questions what counts as 'doing science' and in turn, definitions of success and failure— and provides some material evidence towards re-appraising the successfulness of Crick's contribution to the neurosciences.

Dudai, Y., Evers, K., (2014) To Simulate or not to Simulate: What are the questions? Neuron; 84(2):254-261, doi: 10.1016/j.neuron.2014.09.031

Abstract: Simulation is a powerful method in science and engineering. However, simulation is an umbrella term, and its meaning and goals differ among disciplines. Rapid advances in neuroscience and computing draw increasing attention to large-scale brain simulations. What is the meaning of simulation, and what should the method expect to achieve? We discuss the concept of simulation from an integrated scientific and philosophical vantage point and pinpoint selected issues that are specific to brain simulation.

Farisco, M., Evers, K. and Petrini, C. (2014) Biomedical research involving patients with disorders of consciousness: ethical and legal dimensions, Ann Ist Super Sanità;50(3):221-228, doi: 10.4415/ANN_14_03_04

Abstract: The directive 2001/20/UE and the research involving patients with docs. Research involving patients with disorders of consciousness (DOCs) deserves special ethical and legal attention because of its Janus-faced nature. On the one hand, it raises concerns about the risk to expose the involved subjects to disproportionate risks not respecting their individual

dignity, particularly their right to be cared for; on the other hand, research is an essential tool in order to improve the clinical condition of patients with DOCs. The present paper concerns the ethical and legal dimensions of biomedical research involving patients with disorders of consciousness. In particular, it focuses on informed consent to experimental treatments, which is a challenging issue both from an ethical and legal point of view. The first part reads the Directive 2001/20/EU in the light of the experimentation of patients with DOCs, and suggests a revision in order to better assess the issue of informed consent. The particular case of informed consent for observational studies of non-communicative patients. The second part presents an informed consent form for studies through video-recording of patients unable to communicate their own consent. This form has been elaborated by the bioethics unit of the project "Review of the nosography of vegetative states: application of methods of behavioral analysis to individuals in coma or vegetative state" developed at the Italian National Institute of Health. Relevance of the suggested form. The paper describes the conceptual framework of the form for informed consent to studies through video-recoding, which is a relevant example of what issues should be included in an informed consent for any type of studies through video-recording of patients unable to express their own consent. The article has been sent on November the 7th 2013, before the adoption of the Regulation (EU) no. 536/2014 (and consequent abrogation of the Directive 2001/20/EU) and the release of the new edition of the Italian Code of Medical Ethics.

Farisco, M., Petrini, C. (2014) Misdiagnosis as an ethical and scientific challenge, Ann Ist Super Sanità; 50(3):229-233, doi: 10.4415/ANN_14_03_05

Abstract: Difficulties of behavioral assessment of consciousness. An astonishingly high rate of misdiagnosis between vegetative state/unresponsive wakefulness syndrome and minimally conscious state has been detected. This raises the issue of the adequacy of the consciousness' assessment in patients with disorders of consciousness. The behavioral assessment of consciousness could be not able to detect covert awareness, which is increasingly identified by the instrumental assessment. Ancillary methods. Neurotechnology, particularly neuroimaging, provides relevant data concerning the neurological underlying condition of patients with DOCs, but the instrumental approach has still to assess some technical issues. Ethical considerations. A correct diagnosis of a DOC is not only an instrumental issue, but also an ethically relevant demand to the scientific community. Finally, an integration between behavioral and instrumental assessments seems to be the most adequate strategy in order to decrease the rate of misdiagnosis.

Gutierrez-Arenas, O. (2014) Handling and analyzing meshed rendering of segmented structures from 3D image stacks in Blender, Neuroinformatics; 12(4):3-6, doi: 10.1007/s12021-014-9250-5

Abstract: Volume electron microscopy (EM) is an aggregate of techniques (e.g. FIB-SEM, SBF-SEM) which are evolving toward easing the trade-off between resolution and field of view

size so that increasingly large volumes of tissue can be imaged at just a few nanometer resolution. These technical developments have been driven by the necessity of characterizing spatially extended structures with functionally relevant supramolecular elements that are several orders of magnitude smaller. Neuronal microcircuits with their intracellular organelles and synaptic specializations stand out among these demanding biological objects. With the increasing image throughput, the bottleneck in the pipeline is the segmentation of both the neurites (axon and dendrites) and the organelles and synaptic structures from the collected 3D EM image stack. In both cases, iterations between human curation and machine learning procedures are the norm

Gutierrez-Arenas, O., Eriksson, O., and Hellgren Kotaleski, J. (2014) Segregation and crosstalk of D1 receptor-mediated activation of ERK in striatal medium spiny neurons upon acute administration of psychostimulants, PLoS Computational Biology;10(1): e1003445, doi: 10.1371/journal.pcbi.1003445

Abstract: Development and regeneration of the nervous system requires the precise formation of axons and dendrites. Kinases and phosphatases are pervasive regulators of cellular function and have been implicated in controlling axodendritic development and regeneration. We undertook a gain-of-function analysis to determine the functions of kinases and phosphatases in the regulation of neuron morphology. Over 300 kinases and 124 esterases and phosphatases were studied by high-content analysis of rat hippocampal neurons. Proteins previously implicated in neurite growth, such as ERK1, GSK3, EphA8, FGFR, PI3K, PKC, p38, and PP1a, were confirmed to have effects in our functional assays. We also identified novel positive and negative neurite growth regulators. These include neuronal-developmentally regulated kinases such as the activin receptor, interferon regulatory factor 6 (IRF6) and neural leucine-rich repeat 1 (LRRN1). The protein kinase N2 (PKN2) and choline kinase alpha (CHKA) kinases, and the phosphatases PPEF2 and SMPD1, have little or no established functions in neuronal function, but were sufficient to promote neurite growth. In addition, pathway analysis revealed that members of signaling pathways involved in cancer progression and axis formation enhanced neurite outgrowth, whereas cytokine-related pathways significantly inhibited neurite formation.

Nair, A. G., Gutierrez-Arenas, O., Eriksson, O., Jauhiainen, A., Blackwell, K. T. and Hellgren Kotaleski, J. (2014) Modeling intracellular signaling underlying striatal function in health and disease. In "Computational Molecular Neurosciences: Novel approaches to understanding diseases of the nervous systems";123: 277-304, doi: 10.1016/B978-0-12-397897-4.00013-9

Abstract: Striatum, which is the input nucleus of the basal ganglia, integrates cortical and thalamic glutamatergic inputs with dopaminergic afferents from the substantia nigra pars compacta. The combination of dopamine and glutamate strongly modulates molecular and cellular properties of striatal neurons and the strength of corticostriatal synapses. These actions are performed via intracellular signaling networks, containing several intertwined

feedback loops. Understanding the role of dopamine and other neuromodulators requires the development of quantitative dynamical models for describing the intracellular signaling, in order to provide precise unambiguous descriptions and quantitative predictions. Building such models requires integration of data from multiple data sources containing information regarding the molecular interactions, the strength of these interactions, and the subcellular localization of the molecules. Due to the uncertainty, variability, and sparseness of these data, parameter estimation techniques are critical for inferring or constraining the unknown parameters, and sensitivity analysis evaluates which parameters are most critical for a given observed macroscopic behavior. Here, we briefly review the modeling approaches and tools that have been used to investigate biochemical signaling in the striatum, along with some of the models built around striatum. We also suggest a future direction for the development of such models from the, now becoming abundant, high-throughput data.

Rose, N., and Abi-Rached, J. (2014) Governing through the Brain: Neuropolitics, Neuroscience and Subjectivity, Cambridge Anthropology 32, no. 1: 3-23, doi: 10.3167/ca.2014.320102

Abstract: This article considers how the brain has become an object and target for governing human beings. How, and to what extent, has governing the conduct of human beings come to require, presuppose and utilize a knowledge of the human brain? How, and with what consequences, are so many aspects of human existence coming to be problematized in terms of the brain? And what role are these new 'cerebral knowledges' and technologies coming to play in our contemporary forms of subjectification, and our ways of governing ourselves? After a brief historical excursus, we delineate four pathways through which neuroscience has left the lab and became entangled with the government of the living: psychopharmacology, brain imaging, neuroplasticity and genomics. We conclude by asking whether the 'psychological complex' of the twentieth century is giving way to a 'neurobiological complex' in the twenty-first, and, if so, how the social and human sciences should respond.

Rose, N. (2014) The Human Sciences in a Biological Age (Portuguese), Polis e Psique 4, no. 2: 3-4, doi: 10.1177/0263276412456569

Abstract: We live, according to some, in the century of biology, where we now understand ourselves in radically new ways as the insights of genomics and neuroscience have opened up the workings of our bodies and our minds to new kinds of knowledge and intervention. Is a new figure of the human, and of the social, taking shape in the 21st century? With what consequences for the politics of life today? And with what implications, if any, for the social, cultural and human sciences? These are the issues that are discussed in this article, which argues that a new relation is required with the life sciences, beyond commentary and critique, if the social and human sciences are to revitalize themselves for the 21st century.

Rose, N. (2013) The Human Sciences in a Biological Age. Theory, Culture and Society 30, no. 1 (2013): 3-34, 10.1177/0263276412456569

Abstract: We live, according to some, in the century of biology, where we now understand ourselves in radically new ways as the insights of genomics and neuroscience have opened up the workings of our bodies and our minds to new kinds of knowledge and intervention. Is a new figure of the human, and of the social, taking shape in the 21st century? With what consequences for the politics of life today? And with what implications, if any, for the social, cultural and human sciences? These are the issues that are discussed in this article, which argues that a new relation is required with the life sciences, beyond commentary and critique, if the social and human sciences are to revitalize themselves for the 21st century.

Rose, N. (2013) Democracy in the contemporary life sciences, BioSocieties 7, no. 4: 459-472, doi: 10.1057/biosoc.2012.26

Abstract: In this article, I reflect on the contemporary arguments for democratisation of science, in light of the work of the historian of the life sciences Ludwik Fleck. I explore some possible reasons for the current demands for 'responsibility' among scientific researchers, and briefly consider this in the context of the various arguments that have made a link between democracy and science, or considered the role of science in a democratic society. I conclude by considering some recent proposals for opening up the secluded spaces of scientific research and truth finding, and suggest that, far from destabilising scientific truth, such developments might actually address the well known failures of 'translation from bench to bedside', and make scientific truth claims in the life sciences more robust when they leave the lab and enter the world of everyday life.

Buchser, W. J., Slepak, T.I., Gutierrez-Arenas, O., Bixby, J.L., Lemmon, V.P., (2010) Kinase/phosphatase overexpression reveals pathways regulating hippocampal neuron morphology, Mol Syst Biol., 6:391, doi: 10.1038/msb.2010.52

Abstract: The convergence of corticostriatal glutamate and dopamine from the midbrain in the striatal medium spiny neurons (MSN) triggers synaptic plasticity that underlies reinforcement learning and pathological conditions such as psychostimulant addiction. The increase in striatal dopamine produced by the acute administration of psychostimulants has been found to activate not only effectors of the AC5/cAMP/PKA signalling cascade such as GluR1, but also effectors of the NMDAR/Ca(2+)/RAS cascade such as ERK. The dopamine-triggered effects on both these cascades are mediated by D1R coupled to Golf but while the phosphorylation of GluR1 is affected by reductions in the available amount of Golf but not of D1R, the activation of ERK follows the opposite pattern. This segregation is puzzling considering that D1R-induced Golf activation monotonically increases with DA and that there is crosstalk from the AC5/cAMP/PKA cascade to the NMDAR/Ca(2+)/RAS cascade via a

STEP (a tyrosine phosphatase). In this work, we developed a signalling model which accounts for this segregation based on the assumption that a common pool of D1R and Golf is distributed in two D1R/Golf signalling compartments. This model integrates a relatively large amount of experimental data for neurons in vivo and in vitro. We used it to explore the crosstalk topologies under which the sensitivities of the AC5/cAMP/PKA signalling cascade to reductions in D1R or Golf are transferred or not to the activation of ERK. We found that the sequestration of STEP by its substrate ERK together with the insensitivity of STEP activity on targets upstream of ERK (i.e. Fyn and NR2B) to PKA phosphorylation are able to explain the experimentally observed segregation. This model provides a quantitative framework for simulation based experiments to study signalling required for long term potentiation in MSNs.

Books and Book Chapters

M. Farisco, The ethical implications of indicators of consciousness in artificial system, in G. Stark, M. Ienca (eds.), Brains and Machines: Towards a unified Ethics of AI and Neuroscience, Elsevier (In Press).

Abstract: The prospect of artificial consciousness raises both theoretical, technical and ethical challenges which converge in the core issue of how to eventually identify and characterize it. In order to provide an answer to this question, I propose to start from a theoretical reflection about the meaning and main characteristics of consciousness. On the basis of this conceptual clarification it is then possible to think about relevant empirical indicators (i.e., features that facilitate the attribution of consciousness to the system considered) and identify key ethical implications that arise.

In this chapter, I further elaborate previous work on the topic, presenting a list of candidate indicators of consciousness in artificial systems and introducing an ethical reflection about their potential implications. Specifically I focus on two main ethical issues: the conditions for considering an artificial system as a moral subject; the need for a non-anthropocentric approach in reflecting about the science and the ethics of artificial consciousness.

M. Farisco, The ethics in the management of patients with disorders of consciousness, in C. Schnakers-S. Laureys (eds.), Coma and Disorders of Consciousness, Springer Publishing Company (In press)

Abstract: Ethical reflection can play an important role in optimizing the management of patients with Disorders of Consciousness, for instance revealing both good and bad practices. In fact, several issues arise, both from research and from clinics, and they deserve specific attention. It is urgent to move beyond a defensive type of ethics focused on safe-guarding from potential risks towards an ethics that is also pro-active, and to think about a methodology and a model for translating ethical thinking into the actual clinical treatment of affected patients. In this chapter, I first propose the distinction between fundamental and practical ethical issues as a methodology for the ethics in the management of patients with Disorders of Consciousness, then I provide relevant illustrations of both kinds of ethical issues, and finally I describe the Distributed Responsibility Model for the clinical operationalization of ethics.

Salles, A. (2023). Some reflections on the neurorights debate. In: The risks and challenges of neurotechnologies for human rights. Eds Sosa Navarro, M., Salvador Dura-Bernal, C. M. G. (UNESCO, University of Milan-Bicocca, SUNY Downstate). ISBN 978-92-3-100567-1.

https://unesdoc.unesco.org/ark:/48223/pf0000384185

Abstract: The development and use of a large variety of invasive and non-invasive technologies that interface with the brain show promise in a number of contexts. At the same time, insofar as they record brain activity and/or modulate its function, they prompt questions and concerns about their potential impact on our belief system and our humanness itself. Lately, some scientists and neuro-ethicists have proposed special rights, known as "neurorights." Such a proposal has generated a lively debate. Concerns range from whether the articulation of a new set of rights might undermine existing human rights and the possibility of rights inflation, to how to address the relevant conceptual ambiguities and practical challenges in the proposal. Here I point to two related issues that would benefit from more reflection: framing and engagement.

Stahl, B. C., Scroeder, D., Rodrigues, R., Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges, SpringerBriefs in Research and Innovation Governance, eBook ISBN: 978-3-031-17040-9. doi: [10.1007/978-3-031-17040-9](10.1007/978-3-031-17040-9)

Abstract: This open access collection of AI ethics case studies is the first book to present real-life case studies combined with commentaries and strategies for overcoming ethical challenges. Case studies are one of the best ways to learn about ethical dilemmas and to achieve insights into various complexities and stakeholder perspectives. Given the omnipresence of AI ethics in academic, policy and media debates, the book will be suitable for a wide range of audiences, from scholars of different disciplines (e.g. AI science, ethics, politics, philosophy, economics) to policy-makers, lobbying NGOs, teachers and the educated public.

M. Farisco (ed.), *Neuroethics and cultural diversity*, ISTE Ltd, London 2022

Abstract: There is a growing discussion about the relation between neuroethical reflection and cultural diversity, which is among the most impactful factors shaping neuroethics, both as a scientific discipline and as a social enterprise. The impact of culture on science and on its public perception are particularly relevant to neuroethics, which aims to facilitate the creation of an interface between neuroscience and society at large. Time is ripe for neuroethics to go over the influence of the culturally specific contexts (i.e., North America and Western Europe) it originated from and to include also voices from other cultures in the discussion.

This book illustrates that a convergent approach among different cultures in identifying the main issues raised by neuroscience and emerging technologies and getting a consensus about the priorities to be assessed is possible. This should be taken as the starting point for advancing in the search for shared solutions, if not definitive, at least sufficiently reliable to be translated in democratic deliberative processes.

Ulnicane, I., Knight, W., Leach, T., Stahl, B. C., Wanjiku. W-G. (2022). Governance of Artificial Intelligence: Emerging international trends and policy frames. In: The Global Politics of Artificial Intelligence, Chapman and Hall/CRC: 29-55, eBook ISBN: 9780429446726, doi: 10.1201/9780429446726-2

Abstract: In recent years, national governments, international organizations and stakeholders have launched numerous Artificial Intelligence (AI) strategies and reports. Recent research has mostly focused on AI ethics, while topics of AI policy and governance have received less attention. To address this research gap, this chapter addresses two main questions: what is driving fast-developing AI policies around the world and what are the main frames of emerging AI policies. To make sense of recent AI policy developments, this chapter draws on literature on emerging technologies, in particular on studies of performative function of hypes and expectations as well as of collaboration and competition dynamics in emerging fields. The analysis demonstrates that the fast-development of AI policy is largely driven, firstly, by a wide range of impacts of AI, and, secondly, by international assemblies such as the World Economic Forum and the Organization for Economic Co-operation and Development as well as by cross-national policy learning. However, AI policy developments are unevenly distributed around the world and are predominantly concentrated in the most developed regions. This chapter identifies three main AI policy frames: first, framing AI as revolutionary, transformative and disruptive technology; second, closely interconnected global competition and collaboration in the field of AI; and thirdly, a three-pillar approach of realising opportunities, mitigating risks and ensuring responsible AI. The chapter highlights that AI policy developments influenced by perceptions of hype, positive and negative expectations as well as global competition and collaboration can have not only positive but also problematic effects on resource allocation and political prioritization.

Ulnicane, I. (2022). Artificial intelligence in the European Union: Policy, ethics and regulation. In: The Routledge Handbook of European Integrations, Routledge: 254-269, eBook ISBN: 9780429262081, doi: 10.4324/9780429262081-15

Abstract: In the context of recent advances in Artificial Intelligence (AI) and launch of AI policy documents and ethics guidelines around the world, since 2017, the European Union (EU) is developing its approach to AI. This chapter examines the key EU AI policy documents and the way the EU is positioning its AI policy vis-à-vis other global players. It draws on 'Europe as a power' debate and in particular on the concepts of Normative Power Europe and Market Power Europe to examine if in its approach to AI the EU prioritizes its norms or rather its market power. In the EU policy documents the elements of both normative and market power are closely intertwined. The EU attempts to project itself globally as a Normative Power Europe promoting its value-based and human-centric approach based and ethics guidelines for Trustworthy AI. These are closely connected with the EU's ambitions to be a Market Power with appropriate regulation and investments which could also support implementation of its values and norms.

Ulnicane, I. (2022). Introduction – technologies and European integrations In: The Routledge Handbook of European Integrations, Routledge (1st ed.). eBook ISBN: 9780429262081 doi: 10.4324/9780429262081

Abstract: This introductory chapter to the technology section provides an overview of complex relationships between technologies and European integrations. First, it highlights that in European integration literature the term 'technology' is often in multiple ways typically without explicit reflection on its meaning. While the term technology is frequently used in instrumental ways, this chapter emphasizes political, social, and cultural aspects of technology. Second, the chapter outlines three modes of European integration relevant for technology: first, 'hidden integration' of technology as infrastructure; second, integration outside the treaty framework, and, third, integration of technology policy within the EU. Finally, the chapter concludes with some questions for future research on mutual shaping of technologies and European integrations.

Farisco, M. Filosofía de las neurociencias: Cerebro, mente, persona (Spanish Edition) Ediciones Universidad Católica de Salta; 2nd edition (2021)

Abstract: Frente a la gran difusión e impacto de las neurociencias, presentamos este texto de Michele Farisco al mundo de la lengua española. Consideramos de vital importancia cultivar y difundir los resultados de estudios serios que se encuentran apartados tanto de la postura biologicista, que hace uso de las neurociencias para extender sus conclusiones de modo totalmente arbitrario y falso, como de aquellas posiciones que mantienen una actitud de condena frente a ellas.

El autor presenta los desafíos que las denominadas neurociencias plantean al saber filosófico, sobre todo en el dominio ético-antropológico, y se opone a la posición antropológica reductiva, en que la ideología prevalece sobre la ciencia, por cuanto formula afirmaciones que escapan totalmente al dominio de las neurociencias en particular, y de la ciencia en general.

La Universidad Católica de Salta presenta al mundo de lengua española este valioso aporte de Michele Farisco, considerando de importancia —frente a la difusión e impacto de las neurociencias—, cultivar y difundir los resultados de estudios serios.

Stahl, B. C. (2021). Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies, doi: 10.1007/978-3-030-69978-9

Abstract: This open access book proposes a novel approach to Artificial Intelligence (AI) ethics. AI offers many advantages: better and faster medical diagnoses, improved business processes and efficiency, and the automation of boring work. But undesirable and ethically problematic consequences are possible too: biases and discrimination, breaches of privacy and security, and societal distortions such as unemployment, economic exploitation and

weakened democratic processes. There is even a prospect, ultimately, of super-intelligent machines replacing humans. The key question, then, is: how can we benefit from AI while addressing its ethical problems? This book presents an innovative answer to the question by presenting a different perspective on AI and its ethical consequences. Instead of looking at individual AI techniques, applications or ethical issues, we can understand AI as a system of ecosystems, consisting of numerous interdependent technologies, applications and stakeholders. Developing this idea, the book explores how AI ecosystems can be shaped to foster human flourishing. Drawing on rich empirical insights and detailed conceptual analysis, it suggests practical measures to ensure that AI is used to make the world a better place.

Salles, A. (2021). Humanness: Some neuroethical reflections. In: Developments in Neuroethics and Bioethics, Academic Press, ISSN 2589-2959, doi: [10.1016/bs.dnb.2021.03.002](10.1016/bs.dnb.2021.03.002)

Abstract: Neuroscience research is expected to further our understanding of humans. Such knowledge might make an impact on some ingrained beliefs about what humans are and invite reconsideration of ethical and legal categories that tend to draw a sharp line between humans and living and non-living non-humans and the legal protections each deserves. Moreover, the convergence of neuroscience and technology leads to the development and applications of neurotechnology to alleviate diseases and even enhance the human brain. Some neurotechnological applications, it is often suggested, could have an impact on humanness in general and on the identity and personhood of specific individuals in particular and that this calls for legal responses. A lively debate regarding the ethical implications of neurotechnologies on personal identity and authenticity has taken a prominent place within neuroethics. Less has been written on what is the "humanness" that some people suggest might be altered by some neurotechnological applications.

In this chapter, I focus on this issue. First, I provide some conceptual distinctions and outline how "humanness" has typically been addressed. Next, I outline current neuroscientific research that gives support to the view that the constant and multifactorial human bio-cultural interplay that enables high level behavioral and cognitive features might be paramount in what humans are. Finally, I advance some implications for the ethical and legal discussion.

Ulnicane, I. (2020). Ever-Changing Big Science and Research Infrastructures: Evolving European Union Policy. In: Big Science and Research Infrastructures in Europe. Edward Elgar Publishing: 76-100, doi: [10.4337/9781839100017.00010](10.4337/9781839100017.00010)

Abstract: This chapter analyzes ongoing European Union (EU) policy changes that affect Big Science and Research Infrastructures. These changes are part of a broader EU political agenda to strengthen the EU's global position, facilitate growth and enhance coordination among EU member states. Important initiatives for Big Science and Research Infrastructures include the launch of large-scale research projects, support for e-infrastructures and the

development of EU policy towards Research Infrastructures in the framework of the European Research Area initiative. To illustrate these EU policy changes, the chapter uses the example of the Human Brain Project – a large-scale research project that is turned into a research infrastructure. The chapter demonstrates that EU policy for Research Infrastructures supports a model of differentiated integration where a number of EU member states and non-members collaborate. While this model has the advantage of bringing together countries interested in particular Research Infrastructures, it can also present challenges when supranational projects are turned into Research Infrastructures.

Ulnicane, I. (2020). Governance of Dual Use Research in the EU: The Case of Neuroscience. In: Emerging Security Technologies and EU Governance. Taylor & Francis. doi: [10.4324/9780429351846-12](10.4324/9780429351846-12)

Abstract: Recently, major research powers around the world have made unprecedented investments in neuroscience as one of the most promising technologies of the twenty-first century. However, advances in neuroscience can be used for beneficial as well as harmful purposes. This chapter analyses the governance of dual-use research in the EU by focusing on the main EU-funded neuroscience project and one of the large-scale international brain initiatives – the Human Brain Project. It shows how this project has developed its governance approach that goes beyond the European Union (EU) Framework Programme definition of dual-use and incorporates concepts of dual-use research of concern, Responsible Research and Innovation as well as political, security, intelligence and military issues. The resulting approach supports the governance of dual-use research that is based on anticipation, reflection, engagement and action as well as involving contributions from researchers, stakeholders and citizens. The chapter demonstrates how the Human Brain Project is overcoming challenges of limitations of the dual-use definition used in the EU Framework Programme, issues of awareness and education and questions of global collaboration. Lessons learned and good practices developed can be relevant for other international brain initiatives as well as research projects in other disciplines.

Ogoh, G., Stahl, B., Eke, D., Akintoye, S., Knight, W., Ulnicane, I. (2019) Data Governance in International Neuroscience Research. In: George C, Whitehouse D, Duquenoy P, eds. Proceedings of the 2019 Health IT Workshop. Middlesex University, London, UK: 15 -19, ISBN 978-1-64713-330-6

https://dora.dmu.ac.uk/bitstream/handle/2086/18941/Position_Paper_Data%20Governance_in_International_Neuroscience_Research.pdf?sequence=1

Abstract: Medical research is governed by a number of universal principles like those laid out in the 1964 Declaration of Helsinki, which stipulates them as 'ethical principles for medical research involving human subjects, including research on identifiable human material and data'. However, the details of its implementation vary from country to country. One reason for this is differences in legislation and agency policy which have an

impact on the conduct of research and level of protection accorded research subjects. For example, since 2008 the U.S. Food and Drug Administration (FDA) only abides by the 1989 version of the Declaration rather than the more recent 2013 version and some have suggested this allows U.S companies to cut ethical corners when working abroad. In the arena of international collaboration in medical research, such differences raise issues for data governance because they affect how data is shared and used, what data is shared, and with whom data can be shared.

With an ever-growing appetite for collaborative research, one of the areas where issues relating to data governance can easily arise is in the field of neuroscience. Neuroscientists have come to realise that the complexity of the human brain and nervous system mean that, only by working collaboratively together, they can in good time hope to successfully unravel the mysteries of the brain for the benefit of humankind. However, it is not yet clear what rules will govern neuroscientific research collaborations particularly when it spans across national borders and what level of protection will be in place for research subjects when their data is shared across multiple geographic regions. In this era of big neuroscience data and large brain projects, this type of collaboration raises serious concerns as the principles governing data collection, sharing, and use vary from country to country. This position paper therefore highlights how growing collaborations in neuroscience projects may raise important questions for data governance that needs to be addressed.

Rainey, S., Stahl, B., Shaw., M., Reinsborough, M. (2019) Ethics Management and Responsible Research and Innovation in the Human Brain Project. In: R von Schomberg and J. Hankins (Eds.), International Handbook on Responsible Innovation: A Global Resource. Edward Elgar Publishing Ltd: 379-392, doi: 10.4337/9781784718862.00035

Abstract: Responsible research and innovation (RRI) is a key concept in current discourses concerning research governance and policy. The practice of Ethics Management in the European Union (EU) Future and Emerging Technologies (FET) Flagship Human Brain Project (HBP) utilises a concept of 'meta-responsibility' in order to further RRI. This chapter will explain the theory and practice of meta-responsibility to demonstrate RRI in practice in the HBP. As a Flagship EC research project, the HBP represents a particular opportunity to espouse the best aspirations of the European research area. In this chapter, particular focus is given to responsible research and innovation as it is theorised and implemented within the HBP. This article focuses specifically on the role and practice of ethics management in the RRI efforts of the HBP. As such, it presents a truncated and incomplete view. This is unavoidable, given the complex nature of the area – the map is not the territory. Other perspectives are possible, from which other aspects of RRI, and of the HBP overall, might gain or lose emphasis. Nevertheless, here is presented an ethics management perspective on, and role in, RRI so far in the HBP.

Rose, N. and Wahlberg, A. (2019) Gouvernmentaliseringen af livet: beregningen af den globale sundhed, in Mads Karlsen and Kaspar Villadsen, eds. Sundhed Og Magt: 59-106, Hans Reitzels Forlag.

https://research.ku.dk/search/result/?pure=en%2Fpublications%2Fguvernementaliseringen-af-livet(42aa7dd7-ae50-4020-aebd-2999b4f87bda).html

Abstract: The contemporary global health agenda has shifted emphasis from mapping disease patterns to calculating disease burden in efforts to gauge 'the state of world health'. In this paper, we account for this shift by showing how a novel epidemiological style of thought emerged in the closing decades of the twentieth century. As is well known, the compilation and tabulation of vital statistics – death-rates, birth-rates, morbidity rates – contributed to the birth of the 'population' in the eighteenth and nineteenth centuries. The population is reformatted from the middle of the twentieth century by 'modified life tables' made up of disability weightings, health state valuations, quality of life scores, disease burden estimates, etc. The problem of morbid death gives way to that of morbid *living*, made calculable through a metrics of 'severity', 'disability' and 'impairment'. A series of new indices and scales (e.g. the QALY and DALY) has contributed to a governmentalization of living, in the course of which the social and consequences of living with disease come to be an object of political concern, and made knowable, calculable and thereby amenable to various strategies of intervention. We conclude by showing how this style of epidemiological thought has generated a new global visibility for brain disorders as their impact on individuals, health care systems and nations are calculated in novel ways.

Stahl, B. C., Timmermans, J., Rainey, S., and Shaw, M. (2019) Ethics in Innovation Management as Meta-Responsibility: The Practice of Responsible Research and Innovation in Human Brain Simulation. In: J. Chen, A. Brem, E. Viardot, and P. K. Wong (Eds.), The Routledge Companion to Innovation Management New York: Routledge: 435-454.

https://www.taylorfrancis.com/chapters/edit/10.4324/9781315276670-22/ethics-innovation-management-meta-responsibility-bernd-carsten-stahl-jos-timmermans-stephen-rainey-mark-shaw

Abstract: RRI is a key concept of R&I governance and policy. The discourse around RRI is rich and offers numerous competing definitions and components. One aspect that is currently not clear is how RRI goes beyond the numerous well-established activities in science and research governance, such as technology assessment, foresight or science and technology studies. Similarly, on the European policy level it is not clear whether the six pillars (ethics, engagement, science education, gender, open access and governance) are comprehensive or whether addressing them would render research automatically responsible. In this chapter we propose a different view of RRI as a meta-responsibility that aims to shape, maintain, develop, coordinate and align existing and novel R&I-related processes, actors and responsibilities, with a view to ensuring desirable and acceptable research outcomes. This

proposal assigns a fundamentally different role to RRI that encompasses the various components and yet goes beyond them. We believe that this idea contributes to the RRI discourse by providing a focus and offering ways of rendering RRI practically relevant. The theoretical contribution of the chapter is thus a development of the RRI discourse based on long-established theories of responsibility. The idea of networks of responsibility provides the basis for the re-conceptualisation of RRI as a meta-responsibility.

Salles, A. (2018) Neurociencia y cambio moral: expectativas y límites. In García Marzá, Domingo; Lozano Aguilar, José Félix; Martínez Navarro, Emilio; Siurana Aparisi, Juan Carlos (eds.), Ética y filosofía política. Homenaje a Adela Cortina, Tecnos, Madrid.

Abstract: In this chapter, I focus on the proactive epigenesis proposal (Evers, Changeux) to unveil some of the deep questions about justification and frameworks that it raises. I propose that the authors further develop their views and elaborate on the specifically moral issues raised by their proposal and hope that their joint work on this issue inspires empirical and theoretical research from disciplines such as moral philosophy, pedagogy, and social science to further examine proactive epigenesis and the possibilities it opens for addressing moral improvement.

Evers, K., Salles, A., Farisco, M. (2017) Theoretical framing of neuroethics: the need for a conceptual approach. In: Racine E, Aspler J (eds.), Debates About Neuroethics. Perspectives on Its Development, Focus, and Future, Springer International Publishing, Dordrecht: 89-107, doi: 10.1007/978-3-319-54651-3_7

Abstract: There are different dominant perspectives, theories, and methodologies within neuroethics, each importantly shaping the identification, understanding, and discussion of the relevant ethical, social, philosophical and scientific issues. In this chapter, we first provide a brief overview of current neuroethical approaches calling attention to a common tendency to underestimate the role and value of conceptual analysis. Against that background, we present and develop the theoretical framework of fundamental neuroethics. Next, we suggest that neuroethics should be built on the sound scientific and philosophical foundations of informed materialism. Finally, we apply the proposed theoretical framework to the neuroethical discussion of brain simulation.

Salles, A. (2017) Neuroethics in context: the development of the discipline in Argentina. In eds Karen Rommelfanger and Syd Johnson Routledge Handbook of Neuroethics, Routledge, New York, doi: 10.4324/9781315708652-32

Abstract: In this chapter, I make explicit some of the salient topics and challenges shaping the development of neuroethics in Argentina. I begin with a brief description of some of the research carried out in the country. Next, I focus on prevalent cultural and socio-political considerations that play a role in how neuroethical issues are identified, perceived, and approached in Argentina. Finally, I briefly explain some of the neuroethical concerns that attract more attention locally or are deemed to be particularly relevant.

Salles, A., Evers, K. (2017) Social Neuroscience and Neuroethics: A Fruitful Synergy. In: Ibanez A, Sedeno L, Garcia A (eds.), Social Neuroscience and Social Science: The Missing Link: Springer International Publishing, Dordrecht: 531-546, doi: 10.1007/978-3-319-68421-5_22

Abstract: Social neuroscience is shedding new light on the relationship between the brain and its environments. In the process, and despite some criticisms from the social sciences, the field is contributing to the discussion of long standing controversies concerning, for example, the "nature-nurture" distinction, and the relationships between social and neurobiological structures. In this article, we argue that in this endeavour social neuroscience would benefit from partnering with neuroethics insofar as their respective areas and methods of explanation are complementary rather than in competition. We provide a richer account of neuroethics than the one given in social neuroscientists' common descriptions of that field, and suggest that, when understood in this richer (and in our view more adequate) fashion, neuroethics may open up productive avenues for research and play a key role in allowing us to determine social neuroscience's contribution to unveiling important epistemological as well as ontological notions. Accordingly, social neuroscience and neuroethics may form a constructive partnership.

Tennison, M., Giordano, J., Moreno, J. (2017) Security threats versus aggregated truths: Ethical issues in use of neuroscience and technology for national security. In: Illes J (eds.), Neuroethics: Anticipating the Future, Oxford University Press, Oxford: 531-553, doi: 10.1093/oso/9780198786832.003.0027

Abstract: This chapter explores current and possible near future uses of brain science for purposes of national security intelligence and defense, and discusses the neuroethical issues, questions and problems that these approaches foster, and posits ways that such ethical concerns can be identified and addressed. The chapter begins with a historical overview of military attempts to employ the tools and techniques of brain and cognitive science, and illustrates some of the ethical problems generated by these attempts. From this background, current and future uses of neurocognitive science in security and defense operations are presented. Extant ethical concerns are defined, with focus upon those ways that various

ethical systems and approaches might be utilized – and be limited – in guiding use and/or non-use of neuroscience and neurotechnology in military and security operations. With recognition of the global trends in neuroscientific and neurotechnological capabilities – and the power dynamics that such scientific capacity can yield - the importance of neuroethical preparedness is emphasized, and a paradigm for neuroethical risk assessment and mitigation is provided.

Farisco, M., Evers, K. (Eds.) (2016) Neurotechnology and direct brain communication. New insights and responsibilities concerning speechless but communicative subjects, Routledge. Taylor&Francis Group. London and New York, doi: 10.4324/9781315723983

Abstract: Neurotechnology and Direct Brain Communication focuses on recent neuroscientific investigations of infant brains and of patients with disorders of consciousness (DOC), both of which are at the forefront of contemporary neuroscience. The prospective use of neurotechnology to access mental states in these subjects, including neuroimaging, brain simulation, and brain computer interfaces, offers new opportunities for clinicians and researchers, but has also received specific attention from philosophical, scientific, ethical, and legal points of view. This book offers the first systematic assessment of these issues, investigating the tools neurotechnology offers to care for verbally non-communicative subjects and suggesting a multidisciplinary approach to the ethical and legal implications of ordinary and experimental practices. The book is divided into three parts: the first and second focus on the scientific and clinical implications of neurological tools for DOC patient and infant care. With reference to these developments, the third and final part presents the case for re-evaluating classical ethical and legal concepts, such as authority, informed consent, and privacy. Neurotechnology and Direct Brain Communication will appeal to researchers and postgraduate students in the fields of cognitive science, medical ethics, medical technology, and the philosophy of the mind. With implications for patient care, it will also be a useful resource for clinicians, medical centres, and health practitioners.

Salles, A. (2016) Brain Imaging and Privacy Concerns. In Farisco M, Evers K (eds.), Neurotechnology and Direct Brain Communication, Routledge, London & New York: 143-156.

https://www.taylorfrancis.com/chapters/edit/10.4324/9781315723983-11/brain-imaging-privacy-concerns-arleen-salles

Abstract: In this chapter, the author highlights some efforts to approach the issue of functional neuro-imaging and its possible threat to privacy in the neuroethics literature. Two main approaches or strategies are usually used in the discussion: the first strategy consists in a description and discussion of what neuro-imaging can and cannot do with a focus on the technical and methodological problems that bedevil the technology. The second strategy focuses on the metaphysical assumptions about the mind underlying concerns on the subject of neuro-imaging and mental privacy. Sometimes these two strategies are used jointly. There

is a third strategy, less common in the neuroethics literature, that brackets technical, methodological, and metaphysical issues to put the focus on the discussion of normative questions. The questions raised are: why would neuroimaging's impinging on privacy be problematic? What is valuable about mental privacy? Would it be morally undesirable to have less of it? My main aim is to outline the first two strategies clarifying their implications for the privacy debate, and then focus more on the third. I end by proposing to expand the normative discussion to incorporate some of the issues raised by a recent account of privacy as contextual integrity.

Evers, K. (2015) Can we be epigenetically proactive? In Metzinger T, Windt JM (eds), Open MIND. Frankfurt am Main: MIND Group. doi: [10.15502/9783958570238](10.15502/9783958570238)

Abstract: The human brain is an essentially evaluative organ endowed with reward systems engaged in learning and memory as well as in higher evaluative tendencies. Our innate species-specific, neuronally-based identity disposes us to develop universal evaluative tendencies, such as self-interest, control-orientation, dissociation, selective sympathy, empathy, and xenophobia. The combination of these tendencies may place us in a predicament. Our neuronal identity makes us social, but also individualistic and self-projective, with an emotional and intellectual engagement that is far more narrowly focused in space and time than the effects of our actions. However, synaptic epigenesis theories of cultural and social imprinting on our brain architecture suggest that there is a possibility of culturally influencing these predispositions. In an analysis of epigenesis by selective stabilisation of synapses, I discuss the relationships between genotype and brain phenotype: the paradox of non-linear evolution between genome and brain complexity; the selection of cultural circuits in the brain during development; and the genesis and epigenetic transmission of cultural imprints. I proceed to discuss the combinatorial explosion of brain representations, and the channelling of behaviour through "epigenetic rules" and top-down control of decision-making. In neurobiological terms, these "rules" are viewed as acquired patterns of connections (scaffoldings), hypothetically stored in frontal cortex long-term memory, which frame the genesis of novel representations and regulate decision-making in a top-down manner. Against that background I propose the possibility of being epigenetically proactive, and adapting our social structures, in both the short and the long term, to benefit, influence, and constructively interact with the ever-developing neuronal architecture of our brains.

Rose, N. and Abi-Rached, J. (2014) Historiciser les neurosciences. In Neurosciences et Société, ed. Brigitte Chamak and Baptiste Moutaud, 51-76. Armand Colin, Paris.

https://scholar.harvard.edu/jabirached/publications/historiciser-les-neurosciences

Abstract: Dans ce chapitre, nous tenterons de dessiner les contours du territoire conceptuel, technologique et sociopolitique de ces neurosciences. Nous identifierons d'abord les conditions qui ont permis la cristallisation de cette nouvelle vision de la recherche sur le

cerveau dans les années 1960 et sa métamorphose d'un simple programme (le NRP) en une discipline à part entière avec ses pratiques idiosyn-cratiques, ses discours, et une nouvelle communauté d'experts: les «neuroscientifiques». Nous commencerons par examiner les forces conceptuelles qui sous-tendent la naissance des neurosciences. Nous reviendrons ensuite sur les formes institutionnelles ou infrastructures qui ont permis l'expansion des neuro sciences et sa consolidation en une nouvelle discipline. Nous analyserons les forces économiques et l'impact de «l'esprit du biocapitalisme» (Rose, 2007), conditions clés pour la matérialisation d'une telle entreprise avec des implications aux niveaux personnels, épistémologiques et sociopolitiques. Enfin, nous analyserons le sens général de l'espoir, des attentes, du battage médiatique et de la prospective, associés aux neuro sciences.

Rose, N. (2014) What is diagnosis for? (Japanese Translation), In Gendai-Shiso (Review of Contemporary Thought). English version available.

http://nikolasrose.com/wp-content/uploads/2013/07/Rose-2013-What-is-diagnosis-for-IoP-revised-July-2013.pdf

Abstract: For some, especially British psychiatric clinicians, the debate over the new edition of DSM, the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association 2013), is over-excited. A diagnostic manual in psychiatry, they say, is just a provisional map of territory, a map of the kinds of troubles that psychiatrists work with in their practices. A manual like DSM is a kind of rough guide that will help the practitioner get oriented, maybe o read on the plane but best to put to one side when one is actually on the ground, that is to say, in the clinic. And indeed, in the UK, most practising clinicians don't use DSM classifications in their work, don't work through the DSM checklists in making their differential diagnoses, don't think of their diagnoses as absolute, let alone believe that they individuate a specific biological substrate in the brain as the cause of the disruption to the lives of the patients. So why all the fuss about DSM 5?

Salles, A. (2015) La neurociencia y la identidad: un debate abierto. In Esquembre CO et al (eds.), El mejoramiento humano, Editorial Comares, Granada, Spain: 57-66.

https://dialnet.unirioja.es/servlet/articulo?codigo=5390922

Abstract: Large brain projects worldwide, such as the American BRAIN initiative and the European Human Brain Project, are generating vigorous moral discussions on a number of topics. They range from how responsible research should be carried out and how to ethically use the findings, to critical questions about the impact of neuroscientific findings on human lives in general and subjective human experiences in particular. One important concern that has been voiced is that advances in brain research can potentially threaten human identity either by substantially altering it or by directly undermining it. In this paper, the author identifies and presents some of these identity related concerns.

Salles, A., Evers, K. (Eds.) (2014) La Vida Social del Cerebro, Editorial Fontamara, Coyoacán, México D.F.

Abstract: La vida social del cerebro, de la colección Derecho Salud y Bioética de la editorial Fontamara, contiene los trabajos realizados por un grupo de autores provenientes de disciplinas en neurociencias, medicina y filosofía. De esta forma, las coordinadoras Arleen Salles y Kathinka Evers —directora del Programa de Neuroética del Centro de Investigaciones Filosóficas en Buenos Aires y la codirectora de The Human Brain Project, respectivamente— invitan a la reflexión sobre preguntas que plantea el conocimiento del cerebro y su naturaleza social. Así, por ejemplo, se presenta el marco científico de las bases neuronales de la empatía y la posibilidad de regular la respuesta empática; los efectos de la pobreza sobre el desarrollo del cerebro; las cuestiones éticas planteadas por la lectura de la mente y la potenciación moral; los aportes de la neurociencia a la comprensión y el tratamiento de los trastornos de la conciencia; la muerte y el debate sobre la relevancia del conocimiento del cerebro y la deliberación moral. Animados por el deseo de fomentar un diálogo racional, los autores presentan visiones orientadas al avance en la comprensión de la naturaleza del cerebro y sus implicaciones éticas, sociales y legales.

Evers, K. (2013) Neuroethics, in Encyclopedia of Sciences and Religions, Springer Science + Business Media B.V., Dordrecht 2013: 1466-1471, doi: 10.1007/978-1-4020-8265-8_1529

Abstract: This chapter describes the discipline of neuroethics and some of its most recent developments; what makes neuroethics distinctive and how it is relevant to the scholarly area called "Science and Religion". It also describes ethical principles that guide this discipline some of its key-values of this discipline/sub-discipline and its views on concepts such as human being, life, reality, knowledge, truth, perception, time, consciousness, rationality/reason, mystery, self and meaning.

Rose, M., and Abi-Rached, J. (2013) Neuro: The New Brain Sciences and the Management of the Mind. Nikolas Rose and Joelle Abi-Rached. Princeton University Press, ISBN 9781400846337, doi: 10.1111/area.12253

Abstract: The brain sciences are influencing our understanding of human behavior as never before, from neuropsychiatry and neuroeconomics to neurotheology and neuroaesthetics. Many now believe that the brain is what makes us human, and it seems that neuroscientists are poised to become the new experts in the management of human conduct. Neuro describes the key developments--theoretical, technological, economic, and biopolitical--that have enabled the neurosciences to gain such traction outside the laboratory. It explores the

ways neurobiological conceptions of personhood are influencing everything from child rearing to criminal justice and are transforming the ways we "know ourselves" as human beings. In this emerging neuro-ontology, we are not "determined" by our neurobiology: on the contrary, it appears that we can and should seek to improve ourselves by understanding and acting on our brains. Neuro examines the implications of this emerging trend, weighing the promises against the perils, and evaluating some widely held concerns about a neurobiological "colonization" of the social and human sciences. Despite identifying many exaggerated claims and premature promises, Neuro argues that the openness provided by the new styles of thought taking shape in neuroscience, with its contemporary conceptions of the neuromolecular, plastic, and social brain, could make possible a new and productive engagement between the social and brain sciences.

HBP Reports

Aicardi, C., Datta, S., Mahfoud, T. and Rose, N. (2020) Formal and Informal Infrastructures of Collaboration in the EU's Human Brain Project. Research Report by Human Brain Project Foresight Lab at King's College London, doi: [10.1177/01622439221123835](10.1177/01622439221123835)

Abstract: This report builds on analytical insights gained through long-term engagement with the Human Brain Project, in addition to a recent round of interviews with scientists and engineers in the HBP. We seek to address the following questions: How do small communities of collaborative practices grow and/or merge into large scale, multicentric research and innovation communities? How are they supported, or hindered, by infrastructure? Our interviews show that while the HBP research infrastructure was designed to facilitate collaboration between scientists within and outside of the project, scientists have been collaborating using alternative means. While much of the the literature on infrastructure focuses on 'top-down', formal infrastructural design, we pay attention to the informal, invisible infrastructural assemblage involved in large-scale interdisciplinary collaborations. We suggest that the formal infrastructure built to facilitate and structure collaboration within large scale interdisciplinary research projects can sometimes render the informal infrastructure and collaborations invisible. Scientists and engineers within the HBP were often engaging in collaborations that were not visible to the project leadership, administration, and to the European Commission because they were not using the formal infrastructure built to support, and account for, these same collaborations.

Bitsch, L., Bang Bådum, N., Campion, N., and Jørgensen, M. L. Insight into Europeans' informed and considered views on Artificial Intelligence – Results from EuropeSay on AI (2020) Human Brain Project, Public Engagement, the Danish Board of Technology Foundation.

https://tekno.dk/app/uploads/2021/05/D12.3.1-Report-on-Citizen-Engagement.pdf

Abstract: This report presents the results of the first European-wide citizen engagement on artificial intelligence, EuropeSay on AI, which ran from September 2019 to January 2020. The consultation focused on getting insights into what European citizens think about artificial intelligence, its potential benefits and risks, as well as some of the applications that it can be put to. The consultation found that participants were generally positive to AI, and applications in medical and health research were generally seen as acceptable, whereas support for use by public authorities was more contingent. However, there were also widespread concerns, e.g. about its potential application for political purposes and for making accurate inferences. In continuation of this, there was also concerns about the lack of control over personal data. In addition, it was found that black-boxing of AI was not acceptable, and that AI systems should be explainable and transparent. Further, the consultation found that the vast majority of participants believed that AI should be subject to regulation, and means to consumer empowerment should be implemented.

Bitsch, L., Bang Bådum, N., Palsberg, A. Lessons from stakeholder engagement and dialogue (2020) Human Brain Project, Public Engagement, the Danish Board of Technology Foundation.

Abstract: This report collects up on the Ethics and Society activities on engagement of stakeholders and the European public in the Human Brain project (HBP). The report covers the period from start of the HBP in October 2013 to the end of the third funding period in March 2020. The activities described in this report were carried out by the Danish Board of Technology Foundation. Overall, we have worked with four main themes: data protection and privacy, potential political, security, intelligence or military (PSIM) uses or research, artificial intelligence (AI), and inclusive community building. In the final phase of the HBP grant period priorities could include: 1) Validating the HBP information material on data collection, use, policies, procedures and consent with citizens and other external audiences for e.g. in relation to ease of understanding and transparency on what how and why data is used in the HBP, 2) Investigating and developing step-by-step approaches to support researchers in anticipating societal and ethical issues, 3) Engaging with professional stakeholders, publics, researchers and other brain projects in international debate on neuroethical issues, including issues related to dual use research of concern and AI to stimulate debate, trust, mutual understanding and collaboration, 4) Supporting education on societal and ethical issues of research in the HBP. In SGA3, we contribute to activities in works package nine (WP9) to develop a training programme in responsible research and innovation for EBRAINS, 5) Supporting the exploration and development of an HBP strategy for commercial exploitation of its work, and assist with the citizen and stakeholder's perspectives and input on the strategy.

Aicardi, C., Datta, S., Mahfoud, T. and Rose, N. (2019) Machine Learning and Big Data for Neuro-Diagnostics: Opportunities and Challenges for Clinical Translation. Scoping Report, Human Brain Project Foresight Lab, King's College London.

Abstract: In this report, we examine some developments in neurodiagnostics that make use of machine learning and other algorithms, with a particular focus on the potentials and challenges for clinical translation. These issues have to be understood in the context of more general questions concerning the social and ethical implications of brain research for psychiatric and neurological clinical practice, and, more widely, in relation to identifying biomarkers that will allow more accurate and individualised diagnosis and treatment. We investigate the use of algorithms and machine learning to search for patterns in data that can individuate the neurobiological correlates of a disorder in ways that could be used to aid diagnosis, to target treatments and hence to improve prognosis. This research is based on long-term engagement with the Human Brain Project as part of its 'Ethics and Society' sub-project since 2013, as well as ethnographic fieldwork, and stakeholder interviews in the UK and Europe in 2018-2019. We draw attention to the challenges faced in relating probabilistic predictions derived from such algorithms to individualised clinical interventions, and we

highlight the importance of trust in the relationships that enable clinical translation of technologies – trust between researchers, clinicians, patients, and regulators.

Bitsch, L., Ramchandra Kotnis, S., Palsberg, A., Bang Badum, N., Jørgensen, M. L. and Klüver, L. (2019) Report from the workshop – AI 360 Copenhagen. The Danish Board of Technology Foundation, Copenhagen, Denmark.

Abstract: The report reproduces the proceedings from the AI 360 COPENHAGEN workshop. This report is put together with the core recommendations from the invited experts as its main focus. Hence the structure of the report starting with a schematic presentation of the experts' recommendations for how to steer the development of AI technologies towards desirable applications and in societally beneficial directions. This is followed by a brief passage on the background of the workshop as well as an overview of the "360 method" and "360 tool" used to guide the discussions of the workshop. Finally, an overview of the challenges and uncertainties raised by AI technology in the various areas of society, as identified and debated by the workshop participants, is provided.

Aicardi, C., Datta, S., Mahfoud, T. and Rose, N. (2018) Infrastructure and Community Building. Scoping Report, Human Brain Project Foresight Lab, King's College London.

Abstract: This report aims to inform the Science and Infrastructure Board (SIB) and the Project Coordination Office (PCO) of the Human Brain Project about some of the issues that may be encountered in the transformation of the Human Brain Project into a sustainable infrastructure project on the 2020 Roadmap of the European Strategy Forum on Research Infrastructures (ESFRI), which was set-up in 2002 to shape policy on research infrastructures across the European Union. There has been much previous social scientific research on the nature and importance of infrastructure, which has also revealed very significant differences in how this term is understood and how infrastructures are implemented. Similarly, while it seems undeniable that research is best done among a community of scholars, and indeed this is the ambition of declarations about the importance of 'open science' by a range of scientific funding bodies, there are significant differences in forms of research communities, and in practices for building such communities. This report aims to summarise key findings of this research, to set out European policy on Research Infrastructures, to consider a number of existing European Research Infrastructures, and to examine some specific issues that might affect the capacity of the Human Brain Project to build a successful research community and a durable Research Infrastructure within the ESFRI framework. We conclude with a number of recommendations to the SIB and the PCO of the HBP which we hope will support its ambitions in community building and infrastructure development.

Aicardi, C., Datta, S., Mahfoud, T. and Rose, N. (2018) Dual Use in Neuroscience and Neurotechnology. Scoping Report, Human Brain Project Foresight Lab, King's College London.

Abstract: The Ethics and Society sub-project of the Human Brain Project (HBP) and the independent HBP Ethics Advisory Board intend to publish an 'opinion' on the mitigation of 'dual use' of research in the HBP in particular, and more generally in current research and development in neuroscience and neurotechnology. The present report is a scoping document to provide evidence and background for that opinion. It aims to provide evidence the contemporary and envisaged deployments of neurotechnologies in political, security, intelligence and military contexts; the current legal and regulatory framework; the work within the HBP with 'dual use – or what we prefer to term PSIM 'political, security, intelligence and military (PSIM) potential; and finally, some discussion of existing recommendations concerning mitigation of these potentials.

Mahfoud, T., Aicardi, C., Datta, S., Ulnicane, I. and Rose, N. (2018) Neuromorphic Computing: Machine Learning, Open Science, Military and Industry Partnership. Workshop Report, Human Brain Project Foresight Lab, King's College London.

Abstract: The Human Brain Project (HBP) Foresight Lab at King's College London and Ethics Support at De Montfort University co-organised a researcher awareness workshop titled "Dual Use and Neuromorphic Computing" with the SpiNNaker team at the University of Manchester on 18 June 2018. The aims of this workshop were to 1) discuss dual use issues with scientists and engineers who are researching and developing neuromorphic computing hardware and software, and 2) feedback the research of the HBP Ethics and Society subproject into the everyday work of scientists and engineers. This report is organised in the following sections: Ethics and Society in the Human Brain Project; Ethical and social implications of neuromorphic computing; Dual Use; Multilateral Treaties and the EU's Dual Use Policies; Citizens' and Policy-makers' Perspectives on HBP and Dual Use; General Discussion; Scenario Discussion Sessions; Concluding Discussion; Ideas for Dual Use Action Plan and Next Steps.

Mahfoud, T., Aicardi, C., Datta, S. and Rose, N. (2018) Neuromorphic Computing: Artificial Intelligence, Open Science, and Industry Collaboration. Workshop Report, Human Brain Project Foresight Lab, King's College London.

Abstract: The Human Brain Project (HBP) Foresight Lab at King's College London co-organised a workshop titled "Dual Use and Neuromorphic Computing" with the Neuromorphic Hardware team at the University of Heidelberg on 9 November 2017. The aims of this workshop were to 1) discuss dual use issues with scientists and engineers who are researching and developing neuromorphic computing hardware and software, and 2) feedback the research of the HBP Ethics and Society sub-project into the everyday work of scientists and engineers. This report is organised in the following sections: Ethics and Society

Evers, K., Farisco, M., Giordano, J., Salles, A. (2017) Dual Use in Neuroscientific and Neurotechnological Research. A Report on Background, Developments and Recommendations for Ethical Address, Assessment and Guidance of Human Brain Project Activities. HBP Neuroethics and Philosophy (SP-12) CRB-Uppsala University Report.

Abstract: Neuroscience employs a variety of methods and technologies to evaluate and influence neurologic substrates and processes of cognition, emotion, and behaviour. In general, brain science can be either basic or applied research. Basic research focuses upon obtaining knowledge and furthering understanding of structures and functions of the nervous system on a variety of levels by employing methods of the physical and natural sciences. Applied research seeks to develop translational approaches that can be directly utilized to understand and modify the physiology, psychology, and/or pathology of target organisms, including humans. The techniques of both basic and applied neuroscience can be further categorized as those used to assess, and those used to affect the structures and functions of the nervous system, although these categories and actions are not mutually exclusive. For example, the use of certain ligands, toxins, and probes that are used to elucidate functions of various sites of the central and peripheral nervous system can also affect neural activity. Neuroscience is broadly considered to be a natural and/or life science and there is implicit and explicit intent, if not expectation to develop and employ tools and outcomes of research in clinical medicine. Given the goals of medicine to elicit right and "good" treatment in patients' best interests, neuroscientific research is conducted in accordance with an undergirding maxim of non-harm (non-maleficence). However, absence of harm cannot always be assured for the use of research findings and/or products. This latter point has become somewhat contentious and is the focus of this report as regards the potential and actual uses of neuroscientific research that are distinct from intended applications, and/or specifically intended to incur demonstrably threatening consequences to individual and public health and/or environmental integrity. Such applications of scientific and technological research are referred to as "dual use".

Aicardi, C., Reinsborough, M., and Rose, N. (2016) Foresight report on future computing and robotics: A Report from the HBP Foresight Lab.
https://kclpure.kcl.ac.uk/portal/files/86508137/KCLForesightLab_2016_Future_computing_robotics.pdf

Abstract: As computing is increasingly integrated into all aspects of societies, economies and everyday life, and as robotic machines play an ever-larger role in industrial production and elsewhere, carrying out many tasks that were previously only possible by skilled human labour, there is a growing public debate about the ways these developments are influencing

different aspects of our societies, our economies and our lives. Lately, such developments have led to a growing amount of speculation about so called 'intelligent machines' with some suggesting that we will soon reach the 'singularity', the point at which machine intelligence overtakes human intelligence. The view that we take in this report is that such speculations are, at the very least, premature, and divert our attention from more pressing social and ethical issues that are arising in connection to the spread of not-so-intelligent machines. It is these less dramatic, and less speculative, implications that we focus on. The HBP Subprojects can potentially contribute to many areas of ICT and robotics, leading to a variety of future applications and cutting across many domains. Instead of a piecemeal inventory of possible products, we adopt a holistic approach, looking at hardware and software, machines and humans, as parts of larger systems. Our goal is to identify social and ethical challenges related to the potential contributions of the Project to future ICT and robotics. We first single out two cross-cutting topics: intelligent machines and human-machine integration. Then we look beyond the direct contributions of the HBP research and explore potential social and economic challenges that more general developments in ICT and robotics may bring, focusing in particular on affective computing and the impact on jobs and the economy.

Bang Badum, N. and Jørgensen, M. L. (2016) European Citizens' View on Neuroscience and Dual Use Synthesis Report of Citizen Workshops. The Danish Board of Technology Foundation, Copenhagen, Denmark.

http://hbp.tekno.dk/wp-content/uploads/2018/03/Synthesis-Report-of-Citizen-Workshops.pdf

Abstract: As part of the HBP citizen consultation on dual use, citizen workshops were carried out in 8 European countries, and a total of 241 European residents took part in this face-to-face consultation. The research showed that the citizens were generally concerned about the HBP research and the potential uses that it could be put to. The participating citizens were most frequently concerned about how these technological advances could lead to or be used for dehumanization of society, reduction of self-determination and free will, manipulation and political and social control and, lastly, privacy and surveillance. The overall conclusion of the face-to-face workshops was that the citizens, despite their concerns, were in favour of continuing neuroscience research even if it could have dual use, as long as it contributes to developing society, science and technology in a beneficial way. The citizens' support for continued neuroscience research was contingent on the development of international legislation and ethical guidelines for the research and use of neuroscience, and they suggested setting up a monitoring and enforcement body. To the citizens, policy-makers should play a central role in defining what neuroscience research and use is acceptable.

Dudai, Y., Evers, K., Second report on simulation, brain, body and environment (2016), the Human Brain Project.

Abstract: Brains are inherently influenced by context. We propose a taxonomy for such context, based on the source of information, from either within the body (intracorporeal) or outside the body (extracorporeal). These types of context must be taken into account in modelling and ultimately in simulating the brain. Many of the difficulties in qualitatively, let alone quantitatively, estimating the effect of context, or lack of context, on realistic brain function, stem from our lack of knowledge concerning the meaning and resolution of information in both intracorporeal and extracorporeal context that is obligatory for proper operation, and in the context of brain modeling and simulation, particularly from the daunting complexity required to simulate an intracorporeal but extraencephalic context that is sufficiently identical to the operating human body. Furthermore, we do not yet know what is the role of context in permitting or even generating types of consciousness. All in all, such gaps of knowledge notwithstanding, we posit that a large-scale brain simulation project, even if successful in the far future, will never be satisfactory completed in the absence of incorporation of bodily, social and environmental contexts.

Aicardi, C., Reinsborough, M., and Rose, N. (2015) Foresight report on future medicine: A Report from the HBP Foresight Lab.

https://kclpure.kcl.ac.uk/portal/files/86508529/KCLForesightLab_2015_Future_Medicine.pdf

Abstract: The Human Brain Project (HBP) is one of the Future and Emerging Technology Flagship initiatives, which are funded by the European Commission and promote ambitious and science-driven research initiatives. The HBP is a 10-year project in medicine, neuroscience and computing that brings together scientists and institutions from 20 nations across Europe. As part of this, the builders of the HBP Medical Informatics Platform (MIP) are working to identify the biological basis of brain diseases, with the long-term goal of informing the development of new treatments for neurodegenerative and psychiatric disorders. Developers of the HBP MIP are federating medical data from European hospitals and deploying statistical clustering strategies to identify the brain signatures of these disorders. This report by the HBP Foresight Lab at King's College London (Work Package 12.1) outlines the major societal and ethical challenges faced by MIP developers as they work to federate data and translate key health findings into clinical practice. It focuses on two key issues: Data federation and privacy; and Disease signatures and personalised medicine. Data protection and data privacy present some major challenges for the federated organisation of the HBP Medical Informatics Platform, which our Foresight research suggests will be key to the future of the Project. We identify three such challenges that are interwoven in the multi-layered architecture of the MIP: legality, trustworthiness, and privacy. Some of these challenges may be addressed by measures for technology management; others may be addressed via community-building activities around the MIP. These activities could involve

clinicians, the pharmaceutical industry and other professional stakeholders, as well as patient groups and the general public.

Aicardi, C., Reinsborough, M., and Rose, N. (2015) Foresight report on future neuroscience: A Report from the HBP Foresight Lab,

https://kclpure.kcl.ac.uk/portal/files/86508305/KCLForesightLab_2015_Future_Neuroscience.pdf

Abstract: After a technical review run by the European Commission, some relevant initiatives organised by the HBP Consortium, and a workshop organised at the Fondation Brocher (Hermance, Switzerland), it was decided to focus on the study of the possibilities, issues and practicalities in collaborative neuroscience, paying heed to the between diverse brain modelling communities and approaches. Specifically, the two themes of this report are: a) building an infrastructure for Future Neuroscience, b) building a community for Future Neuroscience. We studied these issues in the frame of a short timescale, because we believe that they may have implications for strategic decisions that have to be made concerning the management of that aspect of the HBP's work.

Dudai, Y., Evers, K. (2015) First report on how far brain simulation can explain mechanisms of the mind, the Human Brain Project.

Abstract: Simulation is a powerful method in science and engineering. In neuroscience, problem-oriented computer simulations of specific systems and functions of the brain are extensively used to test predictions, validate conclusions and models, and to guide hypothesis-driven experiments and new models at various levels of analysis. Rapid advancements in neuroscience and in computing draw increasing attention to large-scale brain simulations. Against this background, we raise the question: 'how far can brain simulation contribute to the explanation the brain and the mind?' We delineate three types of issues that relate to the potential explanatory power of large-scale brain simulations. We note that, whereas some types of issues are expected to be resolved with the advance of neuroscience and computing technology, others pose more profound and long-lasting conceptual obstacles that should be taken into account in managing the expectations from the approach.

Guidelines

Grasenick, K. (2020). Report on the Implementation of the HBP Gender Action Plan (SGA2, M1-M24). Zenodo. doi: 10.5281/zenodo.5535676

Abstract: The report compares the Gender Action Plan (GAP) for SGA2 with the actual achievements during M1-M24. This achievements could be realised by an increased collaboration across the HBP: especially of the GAC, the PCO, the Education Programme and further SP managers openly discussing structures and processes, sharing internal guidelines and offering advice, the support of the Directorate, the Science and Infrastructure Board and the Ethic Rapporteurs. This positive development led to an increase for certain planned activities, especially contributions to guiding documents as well as to conferences and workshops.

Grasenick, K. (2019). Gender Action Plan for SGA2. Zendo. doi: 10.5281/zenodo.5535662

Abstract: The Gender Action Plan identifies key measures for SGA2 to enhance equal opportunities and a better gender balance for the HBP. It is based on a proposal by convelop, the preliminary considerations of EAF Berlin during SGA1, feedback from the HBP Directorate and suggestions by the HBP Gender Advisory Committee.

Grasenick, K. (2020). Career Guideline: "If you don't know where you want to go you might end up at any place.". Zenodo. doi: 10.5281/zenodo.5236363

Abstract: This document offers information on career planning: From identifying your own skills and developing a vision accordingly to building networks and keeping a life-domain-balance, this document guides you through different steps of career development. Further aspects treated are: communicating priorities, tips on writing, mentoring, salary conversation, and many more.

Grasenick, K. (2019). Talent Guideline: "I don't care who they are, I just want the best person.". Zenodo. doi: 10.5281/zenodo.5236298

Abstract: This guideline is intended to support reviewers in considering unconscious bias and diversity management when evaluating CVs, recruiting new staff and interviewing candidates. Step-by-step suggestions and tools are provided for each level in the process to improve diversity and equality in recruiting, that is: job advertisements, interpretation of CVs, job interviews, guidance for applicants.

Grasenick, K. (2019). Publication Guideline: "You know how to write an excellent paper - do you know how to plan your publication strategy?". Zenodo. doi: 10.5281/zenodo.5236352

Abstract: A Guideline on how to design a publication strategy for researchers, their sponsors, and mentors

Grasenick, K., Trattnig, J., Romero, P. F. (2021). Toolkit on equality in governance structures, procedures and implementation. Zenodo. doi: 10.5281/zenodo.5575773

Abstract: This Toolkit provides support for integrating EDI (Equality, Diversity, Inclusion) in governance principles of complex, large-scale, international, publicly funded research infrastructures (CoLIPRIs) and projects, like EBRAINS or the HBP. It is based on experiences and processes developed in the HBP. It covers important elements of governance such as organisational power relations, stakeholder management, questions of leadership and career development, principles for event planning and development of procedures. Additionally, the European framework for equality as well as terminology related to EDI are elaborated.

Dissertation

Farisco, M. (2019) Brain, Consciousness and Disorders of Consciousness at the Intersection of Neuroscience and Philosophy. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1597 (63 pp). Uppsala: Acta Universitatis Upsaliensis ISBN 978-91-513-0749-7. http://uu.diva-portal.org/smash/get/diva2:1347252/FULLTEXT01.pdf

Abstract: The present dissertation starts from the general claim that neuroscience is not neutral, with regard to theoretical questions like the nature of consciousness, but it needs to be complemented with dedicated conceptual analysis. Specifically, the argument for this thesis is that the combination of empirical and conceptual work is a necessary step for assessing the significant questions raised by the most recent study of the brain. Results emerging from neuroscience are conceptually very relevant in themselves but, notwithstanding its theoretical sophistication, neuroscience is not sufficient to provide a complete interpretation or an appropriate understanding of their impact. Consequently, the present thesis starts from the need for an interdisciplinary and hybrid field of research, i.e. fundamental neuroethics. Within this framework, the thesis takes consciousness and related disorders (i.e. Vegetative State/Unresponsive Wakefulness Syndrome, Minimally Conscious State and Coma) and the addicted brain as illustrative cases of the potential fruitful collaboration between empirical and conceptual investigations. The general goal of the thesis is to contribute to the overall development of bridging the gap between empirical and conceptual understandings of consciousness. The first paper sets the theoretical framework, providing an empirically-based description of the brain with significant philosophical implications for an understanding of consciousness. The last three papers of the thesis try to apply the theoretical framework to illustrative cases. Papers II and III analyse the possible application of science and technology for an easier detection and clinical care of patients with disorders of consciousness, with particular attention to communication mediated by neurotechnology and the simulation of the conscious brain, respectively; paper IV provides a potentially new ethical analysis of addiction within the elaborated general conceptual framework. The conclusion of the thesis is that the impact of neuroscientific results needs that a dedicated conceptual approach reveals and investigates their conceptual meaning. This conceptual analysis is not exclusive but integrative and complementary to the empirical science. The case of consciousness, analysed from both an ethical and conceptual point of view, is highly illustrative in this respect. In the end, a conceptual/linguistic work of clarification is urgently needed.