Handbook for RRI for Adaptive Networks for Cognitive Architectures
(D3.7 – SGA3)

Figure 1: Exploring ethical issues of commercialisation and transfer of emerging applications

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Handbook for RRI for adaptive networks for cognitive architectures
Document Filename:	D3.7 (D26) SGA3 M34 SUBMITTED 230228.docx
Deliverable Number:	SGA3 D3.7 (D26)
Deliverable Type:	Report
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M34 / 31 Jan 2023
Actual Delivery Date:	SGA3 M35 / 28 Feb 2023
Author(s):	Achim ROSEMANN, DMU (P16), Michele FARISCO, UU (P83), Tonii LEACH, DMU (P16), Arleen SALLES, UU (P83)
Compiled by:	Tonii LEACH, Achim ROSEMANN, DMU (P16)
Contributor(s):	
WP QC Review:	Anita VAN OERS, Jeannette BOSCHMA, UM (P117)
WP Leader / Deputy Leader Sign Off:	Rainer GOEBEL, UM (P117)
T7.4 QC Review:	Martin TELEFONT, Annemieke MICHELS, EBRAINS (P1)
Description in GA:	The deliverable contains the processes required for responsible work on adaptive networks for cognitive architectures and a description of how these have been integrated into the work of the relevant HBP activities.
Abstract:	This document reports on work undertaken by Task T3.8 in relation to the social, ethical, and conceptual issues relevant to AI-related work undertaken in WP3 and the HBP more widely, within the context of an RRI framework.
Keywords:	Responsible research and innovation; social and ethical analysis of brain-inspired computing applications; regulatory analysis; exploitation and commercial use; industry relations; international technology transfer.
Target Users/Readers:	Researchers and scientists within the HBP; EBRAINS AISBL; EBRAINS service providers; the EBRAINS user community; the wider scientific community interested in areas of Artificial Intelligence (AI), AI ethics, Responsible Research and Innovation (RRI), neuroethics, and robotics.

Table of Contents

Table of Figures

1. Introduction

The development of artificial intelligence (AI) systems inspired by the biological structure of the human brain is expected to enable new types of AI and AI-powered robotics applications that combine an increasing number of cognitive functions, including higher cognitive skills such as planning, complex judgments and decision making. A focal point in the Human Brain Project's (HBP) Work Package 3 (WP3) has been the design of modular brain-inspired cognitive architectures: computational frameworks that consist of multiple modules that each emulate specific brain functions (Goebel 2022). Individual modules can serve as building blocks for many different architectures (HBP, 2022). The integration of these architectures in adaptive networks, and the possibility to add further modules over time, promises to enable the development of more complex and flexible AI systems that combine a variety of cognitive and behavioural tasks that can be used in a multiplicity of application domains. In each domain, architectures can be used for several purposes, including for industrial and military purposes (HBP, 2022; Aicardi et al., 2019). This wide range of possible uses and the potential of brain-inspired cognitive architectures to contribute to the development of artificial general intelligence systems (Ramos, 2022), raises important ethical, social and regulatory issues and responsibilities that will be identified and explored in the context of this Deliverable. These issues require the attention of everyone involved in the research, transfer, innovation, and (commercial) development of brain-inspired AI and robotics systems, including those involved in the regulation, governance, funding, management and administration of these processes.

This Deliverable reports key findings from the work conducted in Task 3.8 of WP3. The central aim of T3.8, as defined in the SGA3 HBP Grant Agreement (2020-2023), was to further develop the reflection about brain-inspired AI in order to define feasible conceptual and ethical tools for assessing the consistency of AI-related work in WP3 and the HBP with the RRI framework to ensure that societal benefits aimed for are achieved. To illustrate how this aim has been met, the Deliverable serves two purposes: Firstly, we summarise the findings from our analysis (i) of the processes required for responsible work on adaptive brain-inspired cognitive architectures. Secondly, we describe (ii) how these processes can be integrated into the work of emerging innovation activities in the HBP and EBRAINS.[1]

To achieve these purposes, we build our analysis around three key areas that require attention to support responsible innovation practices in the context of AI-related work in the HBP. These include:

- Issues related to the actionability of European and other relevant AI regulatory documents and guidelines
- The conceptual analysis of key terms for the ethical analysis of brain-inspired AI, bio-inspired modelling, and related applications
- The investigation of the social and ethical issues that arise in the context of the exploitation, commercialisation and transfer of AI-related inventions and technologies developed in WP3, in the context of the European Union (EU) and internationally.

Each of these three areas is critical to the facilitation of responsible innovation practices for AI-related work in in the HBP. For each of these three areas we identify topics that require attention, suggest processes through which identified concerns can be addressed, and provide options through which responsible work practices can be implemented in emerging innovation activities.

[1] The original formulation of the second of these purposes in the SGA3 Grant Agreement was to describe 'how these processes have been integrated into the work of relevant HBP activities'. However, because the exploitation and innovation potential of AI-related research outcomes from WP3, and in the HBP more widely, is only gradually emerging and will mostly take place in a post-HBP context, we had to minimally adjust this definition. Instead of describing how suggested processes have been integrated, we report now how they can be integrated, when the practical use and transfer of HBP research outcomes becomes a reality. Because the exploitation, transfer, and commercialisation of HBP research outcomes will be supported by EBRAINS, we provide options for implementing responsible innovation practices for researchers and management in both the HBP and EBRAINS.

Structure of the Deliverable

The overall structure of this Deliverable is as follows:

Part 1 starts with an overview of the methods that T3.8 has used and how these have shaped the analysis and findings described in the different sections. It also introduces key scientific and innovation studies concepts that have informed work in T3.8.

Parts 2–4 summarise the key findings of our research in the three above-defined areas, in which the development of feasible conceptual and ethical tools is required to support the implementation of responsible innovation practices in AI-related work in the HBP, EBRAINS and other scientific projects and infrastructures. Each of these parts discusses both, the potential processes needed for responsible work on adaptive brain-inspired cognitive architectures to be achieved (relating to Purpose 1 above), as well as possible options through which identified measures and processes can be integrated into the work of AI-related innovation activities in the HBP, EBRAINS, and other, similar projects and infrastructures (relating to Purpose 2).

Part 2 discusses issues around the actionability of AI regulatory guidelines and proposes strategies to improve the impact of AI regulation, and their applicability in the context of the ongoing research on the development of brain-inspired AI systems and applications.

Part 3 presents a summary of a systematic analysis of concepts that play a significant role in the normative discussion on AI and robotic applications in general, as well as in the context of brain-inspired cognitive architectures. It also discusses the potential contribution of conceptual analysis to AI ethics, including its role in reducing the gap between scientific research and public understandings in the HBP, EBRAINS and beyond.

Part 4 provides an overview of key ethical and social issues regarding the exploitation, transfer and commercialisation of AI-related inventions and technologies developed in WP3 and the HBP. Section 5.1 starts with a brief summary of scientific developments and emerging application areas of work conducted in WP3. Because the (commercial) exploitation and transfer of research outcomes from the HBP to other academic institutions and companies will be supported by EBRAINS, this section also discusses the anticipated role of EBRAINS to facilitate the transfer and innovation of AI-related inventions and technologies developed in WP3 and the HBP more widely. Section 5.2 examines risks and challenges that can arise in the context of the open transfer of software, algorithms, and related computational devices. Section 5.3 investigates challenges related to the closed, proprietary transfer of AI-related inventions and technologies developed in the HBP. Section 5.4 examines questions relating to the international transfer and commercialisation of software, code and algorithms that result from research in WP3 and the HBP.

2. Part 1: Methods and Key Concepts

2.1 Overview of the Methods

Given the scope and variety of themes covered by T3.8, a variety of methods have been employed to conceptualise, explore, and expand on the themes previously outlined.

2.1.1 Theoretical Approaches

Review of the Literature

All research undertaken within T3.8 has sought to build upon existing areas of knowledge. As such, engaging with the current literature has been vital in developing our understanding of the current state of the art in this field, and building towards expanding on this knowledge base. This approach has supplemented work undertaken utilising the other methods outlined here.

Conceptual Analysis

This comprises analytical reflection upon the semantics of key terms related to work undertaken in WP3. Exemplar key terms include: intelligence, artificial intelligence, consciousness, ethics, ethics of artificial intelligence, and bio-inspired artificial intelligence among others. A more detailed explanation on what this conceptual analysis entails and how it has been conducted, along with some illustrative examples, is provided in Section 4.

2.1.2 Empirical Approaches

Interviews

A number of interviews were undertaken with members of WP3 (including researchers, Task Leaders and WP management members) in order to gain a deeper understanding of the work carried out in the WP. Given the scope of these interviews and their role in shaping the development of T3.8 research activities across the period of SGA3, it is therefore unsurprising that these interviews have contributed to work undertaken in relation to all of the identified themes: applications, actionability, conceptual analysis, exploitation and international technology transfer.

Survey

Researchers across the HBP were asked to provide information about the AI systems they design, develop or work with, and on the ethical issues that they tend to associate with these systems via a comprehensive survey. The questions in the survey were derived from the 'Ethics guidelines for a Trustworthy AI' (AI HLEG, 2019) and the 'Assessment List for Trustworthy AI (for self-assessment)' (AI HLEG, 2020), and were designed to elicit in-depth responses where possible issues of ethics and trustworthiness are identified. This approach has contributed to work in relation to the themes of applications and actionability.

Workshop

A collaborative, joint workshop between WP3 T3.8 and EBRAINS was developed, with a focus on Technology Transfer, Commercialisation and Intellectual Property (IP) Rights. The workshop examined social, ethical, legal and political issues related to the transfer and commercialisation of technologies developed in WP3, that could lead to the development of new, bio-inspired Al and neurorobotics applications. This workshop contributed specifically to T3.8 work undertaken in relation to the themes of applications, exploitation, and international technology transfer. A more detailed explanation of the workshop methodology is provided at the start of Section 5.

2.2 Key Concepts

Key concepts have informed the work in T3.8 and have provided necessary background knowledge. We present them in this section. The section starts with an overview of the scientific concepts of 'brain-inspired artificial intelligence' and 'modular, brain-inspired cognitive architectures', which underlie and structure research in WP3. It then continues with a clarification of the concepts of 'AI Ethics' and 'Responsible Research and Innovation', which inform the analysis of the ethical, social, and philosophical analysis that T3.8 has conducted, and ends with an overview of the concepts of 'innovation', 'exploitation', 'knowledge transfer' and 'technology transfer', which play a central role in this Deliverable.

2.2.1 Brain-Inspired AI

In its broadest sense, talk about biological inspiration refers to the compatibility of AI with current knowledge in biology, particularly in neurobiology. Such a general description, however, while useful for the sake of introducing the concept, is not sufficiently constrained to be technically operationalisable. More specifically, an AI system is biologically inspired when its architecture and functioning include biological constraints that make specific parts of the system biologically realistic. Importantly, a biologically inspired AI system does not necessarily fully emulate or replicate the reference biological system, since different levels of biological realism are possible. Even if in theory biological inspiration can come from many different biological systems (Floreano and Mattiussi, 2008), the main trend today is to define biological realism of AI with specific reference to known biological principles of the brain, in particular mammalian and human brains. Of course, there is no such thing as *the* brain as brains vary substantially between both species and individuals of the same species. Furthermore, different organisational levels and regions of the same brain have different properties.

When compared to the human brain, current AI reveals a number of differences and limitations related to different domains and goals. These limitations are arguably of two main kinds: technical (e.g. limited ability to accomplish the task because of the limitation of the processing power) and conceptual (e.g. limited ability to accomplish the task because of the underlying paradigm). Technical limitations depend on the current technological stage of AI and are likely to be reduced and possibly overcome through further progress of knowledge and emerging technology. Conceptual limitations depend on the AI paradigms used, so overcoming them may require revised or new paradigms.

One possible strategy to improve the performance of current AI along the dimensions considered above is 'to introduce structural and operational principles of the brain into the design of computing algorithms and devices' (Poo, 2018). To illustrate this, present AI is still narrow, that is, it works for specific tasks in particular domains for which it is programmed and trained and fails if environmental conditions are different from the training context (Marcus and Davis, 2019). In this respect, its impressive success in specific applications is not yet translated into the capacity for solving broader and more general tasks. Other features of the biological brain that may be translated into AI in order to improve it are flexibility, robustness, and capacity for multi-tasking and minimal supervision, all accomplished with minimal energy supply.

2.2.2 Modular Brain-Inspired Cognitive Architectures

Brain-inspired cognitive architectures (BICA) are computational frameworks, which facilitate the building of AI systems inspired by the biological structure and processes of the human brain (Samsonovich, Gudwin and da Silva Simoes, 2020). The development of BICAs is expected to enable the creation of more complex (e.g. combining multiple cognitive tasks) and flexible (e.g. the ability to shift between different application domains) behaviours than other AI approaches, such as deep neural networks (HBP, 2022), and to play a key role in the development of artificial general intelligence (AGI) (Ramos, 2022)[2].

In the context of research in WP3, the term 'cognitive' refers to a variety of brain functions that range from lower-level abstraction, such as object recognition and saliency detection, the performance of sensorimotor tasks, towards higher cognitive skills such as planning and decision-making (Senden, 2022). The term 'modular' or 'modularity' describes the integration of different architectures that each focus on the performance of specific brain functions into adaptive networks, which combine multiple cognitive functions for the performance of more complex tasks.

This involves the possibility to integrate additional architectures over time, promising the development of highly adaptable AI systems, that can be used in a multiplicity of application domains (HBP, 2022). The creation of modular cognitive architectures can start with the integration of relatively simple tasks, which can then be expanded over time, with additional modules being plugged in and out, including modules that perform higher-level brain functions, such as planning, complex judgements and decision making (Goebel, 2022). Extended, modular architectures can also be used in embodied settings, for example to facilitate complex navigation and manipulation tasks in robotic agents (HBP, 2022).

A modular approach to the development of more complex BICAs and networks allow different research groups to independently develop their own architectures, focusing on the realisation of

[2] Please note: AGI is not a research purpose in the HBP, but brain-inspired AI may offer tools that may enable AGI in the future.

specific functions and tasks, and then to integrate these in a larger framework, by using platform software such as Docker or AngoraPy, which enables communication and synchronisation between these heterogeneous modules (Goebel, 2022). A single module can in principle 'serve as a building block for many different architectures' (HBP, 2022).

In practice, the synchronisation of, and communication between, as well as learning and optimisation across different architectures (that can serve as 'modules' in larger, adaptive networks), are currently still hampered by various technical challenges (Senden, 2022). Research in WP3 (especially in T3.1) has started to address some of these challenges and aims to complete a reference architecture that enables modular integration (of some the BICAs developed in SGA2 and SGA3), and which can serve as a proof-of-concept to the wider research community before the end of the HBP in September 2023.

2.2.3 AI Ethics

AI research is growing rapidly, raising various ethical issues related to safety, risks, trust, transparency, and accountability, among others. This has resulted in the development of AI ethics (Coeckelbergh, 2020). In general terms, AI ethics addresses the social, regulatory, ethical, and philosophical dimensions raised by the design, development, and use of AI. It is an attempt to formulate and develop theoretical and practical approaches to minimise the potential adverse effects of AI research and applications across a diverse range of social and economic activities, and to enhance the advantages of AI for society. Different disciplines play a role (directly and indirectly) on the development of this field (computer science, cognitive science, philosophy, anthropology, moral philosophy) and, in this sense, we can say that AI ethics is characterised by its multidisciplinary character.

The focus and methodology of AI ethics can change as it is directed to different audiences (academics, designers, developers, policy makers). To illustrate, the AI ethics discussion within academic circles consists in practical and foundational reflection on a number of topics, ranging from potential breaches of privacy, biases, promotion of inequality and discrimination to arguably more unique issues such as increasing automated social surveillance, transformation of agency and AI as potentially a form of consciousness (Coeckelbergh, 2020; Floridi, 2013; Hildt, Laas and Sziron, 2020; Taddeo, 2009; Tolmeijer et al, 2020; Vakkuri and Abrahamsson, 2018). Beyond the academy, in the design, development, and deployment communities, the focus appears to be more on the creation and implementation of basic recommendations, as illustrated by the publication of multiple guidelines and documents intended to provide ethical support in response to the issues raised (EGE, 2018; Hagendorff, 2020; HLEG, 2019; IEEE, 2019; Jobin, Ienca and Vayena, 2019; Ryan and Stahl, 2020). In this Deliverable our goal is to go beyond the academic discussion to provide practice-oriented options for assessing the consistency of AI-related work by the HBP.

Accordingly, in the HBP, AI ethics research has been conceived and conducted within the RRI framework (see below). Beyond this specific framework, in several publications we have argued that AI ethics in general would benefit from an interface with other ethical disciplines, in particular, neuroethics (Salles, Evers and Farisco, 2020; Farisco, Evers and Salles, 2020; Farisco, Evers and Salles, 2022). We turn to this issue in the sections below.

2.2.4 Responsible Research and Innovation

Responsible Research and Innovation: RRI is broadly considered 'an approach that anticipates and assesses potential implications and societal expectations with regard to research and innovation, with the aim to foster the design of inclusive and sustainable research and innovation.'[3] There are a number of vectors across which RRI can be considered as operationalisable in relation to research relevant to this task: firstly, the AREA framework consisting of the principles of Anticipate, Reflect,

[3] https://www.humanbrainproject.eu/en/science-development/ethics-and-society/rri-glossary/

Engage, Act[4]; and secondly, the EU RRI Framework consisting of the dimensions of Public Engagement, Open Access, Gender, Ethics, and Science Education.[5]

Within the HBP, ethics-related work generally (and AI-ethics-related work more specifically) has been carried out within the framework of RRI. Although much of the ethics-related work across the HBP more broadly has been undertaken under the auspices of the EU RRI framework, the work of T3.8 in relation to the research of WP3 has focused on the reflection and anticipation dimensions of the AREA framework approach which emphasises early-stage upstream engagement with the longer-term implications of research and emerging innovation processes.

2.2.5 Innovation

The term refers to the transformation of an idea or invention into a service, good or product that serves specific purposes, for which others (e.g. users, customers, citizens, organisations, businesses, universities, etc.) have a demand and/or are willing to pay for (Velasco, 2022). In the context of the HBP and WP3 this means, for example, that inventions such as the development of new software tools, algorithms, or code, must be translated into transferable products or applications that can be distributed to identified user groups or markets, either on a commercial or open access basis. The practicalities of innovation often exceed the skillset and possibilities of scientists, that is why innovation takes often place through spin-offs, or in collaboration with external commercial partners (*ibid.*)

2.2.6 Exploitation

The concept refers to the direct or indirect transfer and utilisation of research results, tools or technologies developed in (usually publicly funded) research projects, by making these available to third parties, either for further research activities, or for innovation and other purposes. Exploitation can be commercial or aim at improving public and academic knowledge, with the aim to facilitate the active use of research outcomes, and to concretise their value and impact for society and the economy (Lorinczi, 2017). In the context of the Human Brain Project, 'exploitation can be understood as the commercial or non-commercial utilisation of a new or significantly improved hardware, software, service, process, model or database by neuroscientific, computing, or medical communities, academic institutions, research organisations, or industrial actors, and being such an utilisation the result of a process by which society and innovators become mutually responsive to each other with a view on the ethical acceptability of the process itself and its exploitable products' (Velasco, 2022).

2.2.7 Knowledge and Technology Transfer

The term 'knowledge transfer' encompasses a broad range of activities that focus on the sharing, distribution, and transition of knowledge from one entity to another. In academic research contexts, such as the HBP, knowledge transfer can entail the sharing of knowledge between scientific domains or disciplines, its distribution to other academic parties, such as universities or research networks, or to the commercial transfer of knowledge to firms and businesses (Lorinczi, 2017).

The concept 'technology transfer' can be seen as a subset of knowledge transfer, which involves the distribution of a technology or tool that can lead to the development of a technology, from the person, team or organisation that has developed and/or owns it, to another person an entity that aims to utilise the technology for the development of new products, services or more complex technological solutions or technological systems (that integrate a multiplicity of different technologies, that complement and interact with each other) (Lorinczi, 2017; Velasco, 2022).

[4] https://www.ukri.org/about-us/epsrc/our-policies-and-standards/framework-for-responsible-innovation/

[5] https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020_en

3. Part 2: Key Findings of Work Related to the Actionability of Regulatory Documents and Guidelines

As stated in the introduction to this document above, a key aim in T3.8 has been to propose feasible conceptual and ethical tools for assessing the consistency of AI-related work in WP3 and the HBP with the RRI framework. This has entailed identifying and analysing the processes used for responsible work on AI and suggesting strategies for improving them when necessary. The following paragraphs explore issues related to the actionability of European and other relevant AI regulatory documents and guidelines, which is a central aspect to achieve responsible innovation practices. They also propose a strategy and related processes that can improve the impact of AI regulation, and their applicability in emerging exploitation and innovation practices of research that targets the development of brain-inspired AI systems and applications.

One of the commonly used strategies to promote responsible AI has been the formulation of guidelines and recommendations.[6] They are intended to support and improve both decision-making from the design to the deployment of AI as well as the ethical governance and regulation of AI-related applications (Bentley, 2018; European Commission, 2019; EGE, 2018; Floridi et al, 2018; HLEG, 2019). Whether it is codes of ethics produced by professional bodies for their members and practitioners, or by other regulatory bodies, or governmental reports, statements, and declarations produced by ad hoc committees tasked with drafting policy documents, there are several such documents in both the public and private sectors.[7]

Despite diverse backgrounds, these documents tend to share a general objective, themes, and methodology. Regarding shared objective, while targeting different non-expert stakeholders such as policymakers, the general public, and professional associations (Jobin, Ienca and Vayena, 2019), guidelines and recommendations aim to provide ethical guidance across different activities, from AI design and development to potential and actual applications (Floridi and Cowls, 2019; Hagendorff, 2020). Regarding themes, these documents are usually particularly concerned with identifying and addressing the potential limitations of human beings, vis a vis AI, particularly in terms of human beings' limited knowledge in the area of AI and robotics, their limited decision-making capacity regarding a technology they don't fully understand, and their limited power to control the development of the technology and its impacts. Finally, regarding methodology, these documents tend to follow traditional practical ethics theorising and methodology, usually taking inspiration from professional ethics codes, such as medical ethics. They are generally characterised by a top-down approach starting from a few classical fundamental principles re-baptised in this context as 'human centred values' (e.g. human dignity, respect for autonomy, non-maleficence, beneficence, justice, and fairness) complemented by other principles more tailored to technology in general and AI in particular (accountability, effectiveness, trust, transparency, and explicability, amongst others) which are expected to jointly inform ethical decision making during the design, development, and application of AI (Floridi and Cowls, 2019; Hagendorff, 2020; Ryan and Stahl, 2020).

These documents are an important step forward: minimally, they show awareness of the ethically charged nature of the research and development of AI and of the need to anticipate and manage some of the social and ethical issues raised. The documents, however, face some challenges: some related to their impact and potential misuse, and others related to their flaws. Regarding impact, experience suggests that many designers and developers are either unaware of their existence or, if aware, not very familiar with their main recommendations. That this is the case, suggests the need to create adequate strategies not just to inform diverse publics (including those within the research and innovation community) of the existence of the documents but also to promote awareness of the

---

[6] According to AI Ethics Guidelines Global Inventory, these number 167 at present (https://algorithmwatch.org/en/ai-ethics-guidelines-global-inventory/).

[7] Some of the insights in this and the next paragraphs are based on published articles written as part of T3.8, especially: Farisco, Evers and Salles, 2022; and Farisco, Evers and Salles, 2020.

importance of anticipating, identifying, and addressing the issues raised by AI at different stages. There is also the risk that, in certain contexts, these documents might be (mis)used as a form of ethics washing, to legitimate commercial activities or ethics lobbying to prevent hard regulation (Floridi, 2019). These possibilities point to the need to promote a culture where every stakeholder is sensitised to societal and ethical considerations, and to further explore how to instil the view that attention to ethical and societal issues should be seen as part and parcel of the research and innovation processes themselves.

Regarding the documents themselves, there are concerns with respect to their applicability: a common objection found in the literature is that current AI guidelines risk being ineffective because of their level of abstraction and the difficulty in translating them into action-oriented recommendations (Hagendorff, 2020; Mittelstadt, 2019; Rességuier and Rodrigues, 2020). Accordingly, there have been recent attempts to focus on how actual organisations understand and address the ethical issues raised by AI (Stahl et al, 2021) and to develop frameworks for actionability of the guidelines (Morley et al, 2020; Stix, 2021). Efforts to address this issue include providing preliminary landscape assessments (to bridge the distinction between what should be done and what can actually be done), calling for a more productive interaction with diverse representative publics at different levels (so as to expand the scope of voices typically heard in the discussion of the issues), and for the creation of inclusive mechanisms for implementation.

An additional issue raised by many of those documents has to do with how AI ethics is conceptualised: there is a tendency to conceive of this discipline as a type of applied ethics that is methodologically principle-oriented, thus suggesting that ethical reflection is fully captured by application of general principles. Finally, AI ethics guidelines tend to be conceptually poor, showing insufficient engagement with foundational issues. In particular, there is scarce conceptual analysis of the terms used and a lack of a thorough clarification of their different meaning in different contexts and of their mutual relevance.

Possible strategy and related processes to address identified challenges and to integrate these in emerging exploitation and innovation practices

As a possible strategy for improving the impact of AI regulation, in several scientific publications we have proposed that AI ethics documents would benefit from interfacing with neuroethical reflection. We think that a mutual collaboration between the two fields would go a long way in addressing some of the challenges posed by guidelines and regulations. Within neuroethics, there have been specific attempts to use conceptual analysis as a tool for productive normative debate, and, among other things, the neuroethics team in the HBP has emphasised the crucial role played by a preliminary analysis of key concepts, as exemplified in particular by fundamental neuroethics.

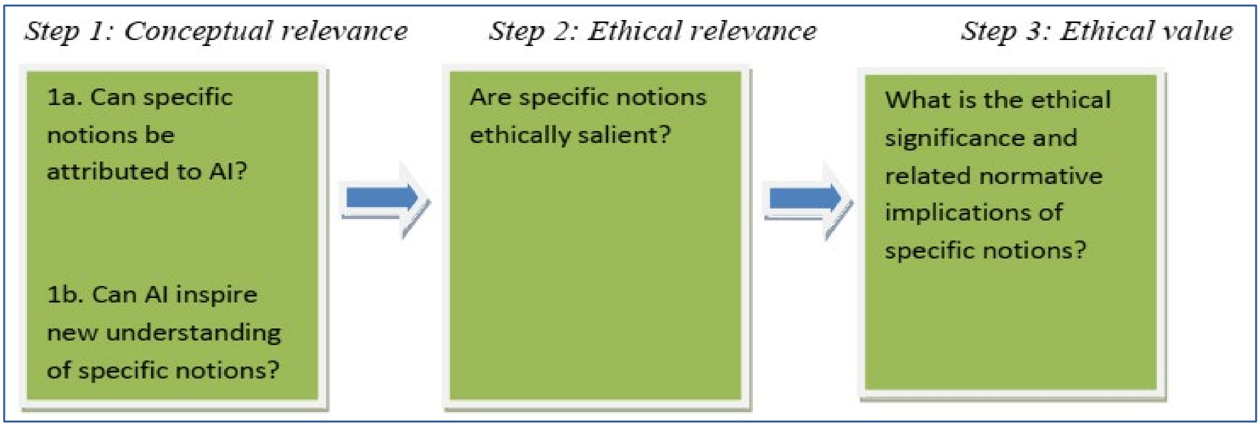

Figure 2: A conceptually informed ethical analysis applied to AI

Insofar as there are a number of notions underlying the normative discussion, and often assumed as relevant to AI in its ethical analysis and regulation, conceptual clarification of what they are and of their possible relevance to AI is necessary in order to fully address the impact of AI on society. The type of conceptual work that we suggest consists of three steps, as illustrated in Figure 2.[8] It aims at determining first, whether those notions are attributable to AI and whether AI can provide a new understanding of those notions; second, if attributable to AI, whether they are ethically relevant (i.e. salient); and, finally, if so, whether they are ethically valuable (i.e. assessing their ethical significance and normative implications).

The conceptual methodology proposed calls for a re-thinking of concepts ethically relevant to AI governance, further clarification of the meaning of AI and of the features usually attributed to it (such as intelligence, autonomy, agency, learning, trust, and transparency), and the clarification of the possible impact that AI language and related conceptual repertoire can have on traditional ethical notions such as autonomy, free-will, and responsibility (Farisco, Evers and Salles, 2020; Salles, Evers and Farisco, 2020). By providing clarity, this methodology could make AI guidelines less abstract and more effective and action inspiring.

Scientific research on adaptive networks for cognitive architectures, and AI research more broadly, should be complemented by ad hoc information about relevant regulation (both soft and hard). However, beyond awareness of relevant guidelines and recommendations, the operationalisation of such regulation can be facilitated by further implementation of RRI activities, including robust conceptual reflection as developed within the HBP. RRI services broadly understood can both offer guidance about regulatory documents, assistance in their application, and raise awareness about the ethical salience of research and emerging results, also through the interaction with representative stakeholders and the public at large. Importantly, RRI services should provide assistance not only in handling emerging ethical issues, but also in anticipating them.

4. Part 3: Key Findings from the Conceptual Analysis

Beyond methodology, neuroethics can contribute to AI ethics in terms of content (Farisco, Evers and Salles, 2022). Some specific ethical and societal issues raised by AI (e.g. the creation of potentially conscious AI, impact on autonomy and personal identity of AI-based brain implants, AI enabled or assisted monitoring of employment/academic performance), rely on notions such as intelligence, and on topics such as the ontological and ethical status of humans and machines that neuroethics has been addressing since its beginnings.

In this section, we present a summary of a systematic analysis of concepts that play a significant role in the normative discussion on AI and robotic applications in general as well as in brain-inspired cognitive architectures research: ‘intelligence’ and ‘brain-inspired AI’. We investigated these notions in collaboration with WP3 researchers. The reflections summarised below point to the potential contribution of conceptual analysis (as developed by HBP neuroethics team) to AI ethics, including its role in reducing the gap between scientific research and public understandings in the HBP, EBRAINS and beyond. The section ends with a discussion of a possible strategy and related processes to address identified challenges, and to complement AI research with a systematic conceptual analysis of key terms.

4.1 Intelligence

There is no agreement on what intelligence is and what it entails (for an overview see, for example, Legg and Hutter, 2007). This is probably the reason why some AI researchers, including Alan Turing, either do not explicitly define intelligence in their work, preferring to focus instead on how intelligent action is manifested (Prescott and Camilleri, 2018), or provide a working definition of

[8] For the sake of simplicity, this illustration of a conceptually informed ethical analysis does not include the background normative dimensions that affect and possibly bias conceptual analysis. In fact, the meaning we give to terms is often the result of implicit or explicit evaluations.

intelligence, sufficient for their purposes but possibly insufficient for an ethical analysis of AI and its social and economic impacts.

Within the life sciences, intelligence has been described as the ability of an organism to solve problems emerging in its natural and social environment (Roth, 2013). Gerhard Roth defines intelligence as behavioural flexibility and innovation capacity. Both abilities subserve the final aim of the intelligent behaviour of a living organism: self-preservation. A biological organism has an intrinsic ability to interact with its surroundings in order to preserve itself, and its intelligent behaviour can be instantiated in different specific features, e.g. associative learning, memory formation, behavioural flexibility, innovation rate and abilities requiring abstract thinking, concept formation, and insight.

It is true that behavioural flexibility and innovation capacity, which a biological account of intelligence recognises as critical, might also be expressed, at least partly, by AI. However, the needs and goals constitutive of these abilities are, at least to date, substantially different in biological organisms and in AI (Farisco, Evers, and Salles, 2020). In the first case, they are the result of an emotional interaction with the world, i.e. the ability to evaluate external stimuli differentiating their respective salience for fulfilling specific goals. While some AI applications are able to recognise/label human emotions, they do so in terms of information processes. This means that AI does not understand in the sense of empathising with emotions as humans who can experience them do. At least at present, AI arguably lacks abilities that neuroethical reflection recognises as ethically relevant and salient, namely what is generally called emotional and social intelligence (Gardner, 1985), notwithstanding some relevant conceptual and technical advancements in this direction (Kirtay et al, 2019), as well as a theory of mind. The conceptual and ethical reflection about intelligence provided by neuroethics can enrich the ethical discussion, for instance, on the use of AI to replace humans in some specific contexts (for example, care of older people and children) by fostering ethical reflection on how humans understand activities such as caring and what actions humans tend to value. Indeed, recognition of AI's lack of social and emotional intelligence might allow us to develop at least one criterion for assessing some AI uses, namely that regardless of the presence of actions commonly labelled as intelligent, the lack of some features (like emotional experience) typically taken to be morally relevant calls for caution when considering the role of AI in some specific human activities.

4.2 Biological Plausibility and Brain-Inspired AI

The concept of brain-inspired AI played a significant role in the work of WP3. As seen in Section 2.2.1, the concept of brain-inspired AI is multifaceted and necessitates, among other things, the identification of the specific brain level that is assumed as reference or target of the emulation/replication. For this reason, it is necessary to avoid general reference to brain-inspiration of AI without necessary specification in order to avoid, for instance, the interpretation of brain-inspired AI as sort of replication of the human brain. This will also help to avoid conceptual confusion when discussing the impact of brain-inspired AI, in the context of ethical reflection. Brain-inspired AI is an explicit topic of WP3 research, and its conceptual and ethical implications have been explored in more detail in an interdisciplinary paper coordinated by T3.8 that is being written in collaboration with several researchers from different WP3 Tasks (Farisco et al., in preparation).

The summary above reinforces a point that we argued for in several scientific publications and Deliverables since the beginning of the HBP: how terms are conceived and used is important, and it plays a role in the normative discussion (Evers and Salles, 2021). The importance of this clarification mainly consists in setting the stage for a more realistic and balanced ethical discussion, avoiding both unrealistic expectations and misplaced fears. Moreover, how terms are interpreted shapes policy making, funding and research priorities, and people's perception of and attitudes towards research and innovation: this importantly hinders the promotion of trust required for science to advance and innovations to develop responsibly. The issue of public perception is not minor: in the HBP we have been particularly concerned with informing and engaging with the public so that diverse communities have a voice. For such engagement to be productive, however, the concepts at stake must be clear. Obscure concepts increase the risk of hype, whether in the form of inspiring unrealistic expectations or unjustified concerns (Evers and Salles, 2021). Furthermore, because concepts often are not value neutral, unclear concepts may be particularly problematic insofar as they might hide underlying values that should be assessed.

Possible strategy and related processes to address identified challenges and to complement AI research with a systematic conceptual analysis of key terms

Concepts such as biological plausibility and intelligence, including the differentiation/comparison between biological and artificial intelligence, is generally not explicitly discussed in the work conducted in WP3 but rather implicitly assumed as a background reference. In this way, these terms play an important role, eventually impacting how the research results are interpreted. For this reason, we propose to complement research on adaptive networks for cognitive architecture, and AI research more broadly, with an explicit reflection about the meaning of key concepts involved. In particular, it is important to raise awareness about the anthropomorphic allure that terms used in AI have, and about its impact on the perception of AI by lay people. It is a well-known fact that AI functionalities and innovations are often anthropomorphised (i.e. described and conceived as characterised by human traits). Anthropomorphic attitudes are not only evident in the general public but also in AI research itself (i.e. in the very language of computer scientists, designers, and programmers). In the case of AI, considering the public's lack of understanding of what it is and what it can do, the anthropomorphism that often permeates the AI research community can have particularly pernicious epistemological and ethical consequences. To avoid these problems, and to align the language used by AI researchers with public perceptions, the creation of interactive spaces that facilitate their connection is desirable. We think that the kind of collaboration between researchers in AI and robotics on the one hand and researchers in the humanities and the social sciences on the other hand, as pursued in WP3, is an example of good practice towards the reduction of the gap between scientific research and public understanding. We encourage to apply this model also in the post HBP phase.

5. Part 4: Key Findings from the ethical analysis of the exploitation, transfer and commercialisation of AI-related research outcomes

This section examines ethical and social issues that can arise in the context of the exploitation, transfer and commercialisation of AI-related software, algorithms, and other computational tools and devices developed in WP3 and the HBP more widely. We explore the following question: What are the ethical and societal challenges of the transfer of AI-related inventions, tools, and technologies developed in the HBP to public and private sector organisations in the European Union and internationally? Which possible measures and processes could be used to address identified challenges and to integrate these in emerging transfer and innovation practices?

To answer these questions, we draw on literature research and findings from an international online workshop that we conducted in 2022 in collaboration with partners from EBRAINS and WP3, and a panel of inter-disciplinary experts, with specialisations in the ethical, legal, policy, and regulatory aspects of the international transfer of software, algorithms, and other digital technologies. The workshop involved the collective, inter-disciplinary evaluation of case studies of research outcomes and technologies developed in WP3 that shall be transferred to third parties via (or with the support of) EBRAINS and aimed to identify key challenges and requirements to support EBRAINS in the development of adequate technology transfer and IPR protocols. The scientists that developed these tools provided short presentations on possible uses of their work, targeted users and possible options for commercialisation and transfer to academic and private sector organisations. The workshop also involved the discussion of two hypothetical case scenarios that explored dilemmas around the illicit transfer and piracy of AI-related software tools. The introductory presentations of the case studies and hypothetical scenarios were followed by structured discussion with the invited panellists, as well as open discussion with the audience. The discussions explored participants' perceptions of (i) the key ethical and societal challenges of each of the presented examples and scenarios, (ii) the responsibilities that arise from these challenges for HBP researchers, their universities, and EBRAINS, as well as (iii) possible options to address and mitigate identified challenges. The analysis of the workshop findings has been supported by complementary literature research and joint reflection among T3.8 members, that has helped to contextualise identified issues, and to situate generated options within existing practices and policy approaches.

We start the discussion in Section 5.1 with an overview of emerging research outcomes and application areas of AI-related work conducted in WP3. Because the (commercial) exploitation and transfer of research outcomes from the HBP to other academic institutions and companies will be supported by EBRAINS, this section also discusses the anticipated role of EBRAINS in facilitating the transfer and innovation of AI-related inventions and technologies developed in WP3 and the HBP, as presented by EBRAINS's CIIO Steven VERMEULEN in the above-mentioned workshop. Sections 5.2, 5.3 and 5.4 do then present key findings from our ethical analysis. Section 5.2 examines risks and challenges that can arise in the context of the open transfer of software, algorithms, and related computational devices. Section 5.3. investigates challenges related to the closed, proprietary transfer of AI-related inventions and technologies developed in the HBP. Section 5.4. examines issues relating to the international transfer and commercialisation of software, code and algorithms that result from research in WP3 and the HBP.

Each of these three sections (and corresponding subsections, which discuss specific types of challenges), spell out possible measures and processes through which identified challenges can be addressed and options to ingrate these in the context of emerging transfer and innovation processes in the HBP, EBRAINS and other projects and infrastructures that work towards the development of new, brain-inspired AI applications.

5.1 Emerging Research Outcomes and Application Areas of AI-related work in WP3

As mentioned in Section 2.2.2, a central aim of WP3 is the development of modular, brain inspired cognitive architectures, and the integration of these architectures into larger, adaptive networks, which are able to combine multiple cognitive functions in AI systems that can perform increasingly complex and versatile tasks. At a material level, these emerging computational frameworks and tools involve the development of software, code, algorithms, learning rules and in some cases also new hardware elements (that are required to run newly developed software or code)[9].

These components serve as building blocks through which the computational instantiation of brain-inspired AI applications is realised, both in simulated and actual physical settings. Brain-inspired modular architectures are developed as software tools that, together with accompanying algorithms, learning rules, etc., can be transferred to other academic research groups, companies and other entities for further research, or purposes of innovation and the development of real-world applications.

A practical challenge is that because much of the work in WP3 is fundamental research that aims to develop reference architectures and proof-of-concept studies that has the potential to enable varied real-world applications in the years to come, the exact scope and types of future applications is, at present, difficult to predict. However, some teams in WP3 have targeted the development of real-world applications and have developed BICAs that can be commercially transferred as code and software tools to other parties, for innovation and the development of varied AI systems and neurorobotic applications.

Examples of application areas that emerge from work in WP3 include, amongst others:

1) Robotics research that focuses on the safe collaboration between humans and robots, conducted in WP3, T3.4. This includes work on improved visual and sensory robot perceptions, for example for proximity detection, refined engagement, and interaction of robots with human partners, such as in the context of the handover of objects, as well as research on context aware navigation

[9] For the most part, the computational frameworks developed in WP3 can be run on existing hardware systems that have been developed outside of the HBP, or by research groups in other HBP WPs, such as the neuromorphic computing systems SpiNNaker and BrainScaleS, that can be accessed via EBRAINS.

and related behavioural adjustments (e.g. in a crowded warehouse environment), and the integration of these skills/tasks in various types of industrial applications (Morel, 2022; Conradt and Morel, 2022, HBP Deliverable D3.5).[10]

2) Research on the building of spatial maps of complex, unknown environments - something that conventional AI networks can take a very long time for, and which can be applied in logistics applications and autonomously operating robotic devices. This research is conducted in WP3, T3.3., and previously in SGA2 in SP6 (Migliore, 2022).

3) Research on rapid machine olfaction and the development of an electronic nose that comes with a portable neuromorphic hardware device and corresponding software and algorithms, conducted in WP3, T3.2. This research enables robot olfaction (the detection of gases and odours) and that can be applied in robotic applications, for instance in the context of disaster management, environmental monitoring, and olfactory prosthetics amongst others (Schmuker, 2022).

4) Research on dextrous manipulation to equip robotic hands with human-like abilities such as the autonomous coordination of complex hand movements, for example in the context of in-hand manipulation of objects, as conducted in WP3, T3.1. The software and architecture that is used to train hand coordination in robots will become available via EBRAINS and can be developed further to focus on other aspects of motor coordination, for instance in the context of walking or movements of other body parts (Senden, 2022; HBP SGA3 Deliverable D3.5).[11] This has potential for varied industrial, medical, but also military applications.

5.1.1 Exploitation, Transfer and Commercialisation of WP3 Findings and the Role of EBRAINS

EBRAINS is in the first instance a service platform that provides users with access to its digital tools, software, and scalable computing facilities. These include the EBRAINS neuromorphic computing systems, neurorobotics platform, and medical informatics platform. Access to these services is currently restricted to academic researchers. EBRAINS both enables and mediates access to its technical infrastructure - it runs the services and allows researchers to use its software and computation facilities on a project-by-project basis and to store and access data, which are shared open access with the wider research community. Simultaneously, EBRAINS is also exploring options to commercialise access to its services in the short to mid-term. This could involve charges for some of its services, including for academic users, but also opening the platform to non-academic users (especially from industry) on a fee-basis. This could include cooperation with companies and other non-academic entities on a project basis, for example for pharmaceutical research, as well as the provision of paid consultancy services (Vermeulen, 2022).

As a successor organisation of the HBP, EBRAINS is also planning to support researchers with the transfer and commercialisation of their research outcomes (i.e. the knowledge, tools or technologies that researchers want to commercialise, and not share openly) to users from academia and industry within the EU and worldwide. This can include code, software, algorithms or other technologies and computing devices. To achieve this, EBRAINS would function as an intermediary between the HBP (and, in the future, also other) researchers and potential customers, to facilitate the commercial exploitation of HBP outcomes, innovation and the development of new products and applications[12]. Customers could be from academia, the private sector, as well as governmental and non-governmental research organisations (Vermeulen, 2022). However, not all researchers that utilise the EBRAINS platform, are likely to use this option. Some researchers may prefer to distribute and commercialise their inventions independently of EBRAINS, with the support of the knowledge or technology transfer offices (TTOs) of their universities, spin-off companies, or as part of existing co-operations with industry.

---

[10] https://sos-ch-dk-2.exo.io/public-website-production-2022/filer_public/0f/70/0f7011d7-7db9-4adf-a058-02c186cae39c/d35_d24_sga3_m21_accepted_220520.pdf

[11] Same weblink as in Footnote 11

[12] This is explored in more detail in Section 5.4.2.

5.1.2 EBRAINS as a Distribution Platform to Support the Transfer and Commercialisation of HBP Research Outcomes

At the time of writing, it was not clear whether EBRAINS would support the transfer and commercialisation of research outcomes on a minimal or more extensive basis. It was also not decided whether the operation of EBRAINS as a distribution platform would be organised on a commercial or non-commercial basis. Both options are currently being considered by the EBRAINS management team (Vermeulen, 2022). In a model with minimal involvement, EBRAINS would mainly serve as a 'pass through' that facilitates connections with possible clients, and possibly provides some information or guidance on the practical and Intellectual Property (IP) rights aspects of the transfer process. In this model, the setting up of contracts, development of license agreements, and other practical aspects of the transfer would then be handled entirely between the inventors/developers (i.e. HBP researchers and their universities, with the support of institutional TTOs) of the transferred goods and the clients (i.e. the companies, organisations and possibly individual users that buy or license the transferred technologies, tools or products). In a more extensive model, EBRAINS would get more actively involved in the transfer process, for example by vetting potential clients, supporting due diligence procedures, and co-developing the IPR and transfer agreements together with the transferring parties (which may include HBP researchers and the knowledge transfer offices of their home institutions). Both models can, in principle, be run on a non-commercial or commercial basis, with EBRAINS taking a commission or fee for the provided administrative and consultancy services (*ibid.*). Due to the higher workload of more extensive support, fees or remuneration for EBRAINS would most likely be higher in the latter model, but HBP and other researchers would also receive more support with licensing procedures and added safeguards (*ibid.*).

An option that is currently under consideration, and which would allow EBRAINS to provide more extensive support for the transfer of research outcomes and technologies to third parties - including industry and other non-academic entities - is through the formation of a spin-off commercial entity. A spin-off is a separate legal entity that functions as an accelerator to facilitate connections with companies, enable innovation, and support researchers with the IPR and contractual dimensions of the transfer and commercialisation processes (*ibid.*). This option would require the direct involvement of EBRAINS in the management of IPR and other technology transfer issues, and also require engagement with the ethical issues spelled out in Sections 5.2, 5.3 and 5.4.

5.2 Challenges and processes required to ensure the responsible open access transfer of AI-related software and computational tools

In line with HBP, EBRAINS and ERC open science requirements,[13] many of the software tools, code, algorithms, data, and other research outcomes of WP3 and other HBP Work Packages, will be distributed open access via EBRAINS, or openly shared via the researchers' home institutions or other channels. Open-source software can often be modified, refined, and developed further, to realise new functions or application areas, and can be shared independently by other users for any purpose, without requiring consent by the developer (e.g. HBP researchers) or the organisation that initially provided the software (e.g. EBRAINS) (Corbly, 2014). The targeted users of open-source transfer in the HBP and EBRAINS are mainly academic researchers, both in Europe and internationally. However,

[13] The principle of Open Science underpins Horizon Europe and constitutes a key requirement in SGA3. Open Science constitutes 'the sharing of knowledge, data and tools as early as possible in the Research and Innovation (R&I) process, in collaboration with all relevant knowledge actors, including academia, industry, public authorities, end users, citizens and society at large' (European Commission, 2021). Open Science practices include (but are not limited to) open access publications, data, software, and other tools, and the creation of open research infrastructures such as EBRAINS.

with EBRAINS considering to offer its services further to industry and other organisations, and the possibility for these entities to access and download software tools and data from EBRAINS for use in external computing platforms, the range of potential users extends far beyond the academic research community.

Recent years have seen the emergence of various government-regulatory concerns around the unrestricted dissemination of open source AI software, especially of those classified as general purpose AI (GPAI) that can be applied to several functions in different technical and application contexts (Future of Life Institute, 2022).[14] Debates around the finalisation of the EU AI Act (published in 2021 as a draft), for instance, raised concerns that the unrestricted availability of open-source GPAI software can lead to the development of dual use, misuse or other harmful or problematic uses, that pose 'risks to health, safety and fundamental rights' (Council of the European Union, 2022, p. 5). These concerns are driven by examples such as the use of an open-source algorithm for the development of an AI system to control navigation systems and sensors of U-2 spy planes of the US Air Force (Bernhardt, 2021), as well as reports of the strategic open-source procurement of AI-related software and tools for military and surveillance purposes by government-related entities in China, India and other countries (Arcesati and Meinhardt, 2021; Science Business Net, 2022). As the HBP Opinion on Responsible Dual Use states, 'the openness of scientific research makes it almost impossible to control its irresponsible uses, or its use by those with malign intent' (Aicardi et al, 2018, p.16).

Possible measures and processes to address identified challenges and to facilitate responsible open transfer practices

Discussions in the workshop clarified that for the providers of open-source software from the HBP, and for EBRAINS (as a distribution platform of open-source software developed in the HBP and other projects), uncertainties about future users, their integrity, motives and planned applications, pose an important challenge. Workshop participants stressed, that in order to mitigate risks of dual use, misuse, or other problematic applications, HBP researchers and EBRAINS should put clear and legally binding restrictions on the intended uses of open-source software, software tools, components and algorithms in the licenses through which they are distributed. This seems especially relevant, if open-source software will be available internationally, and for diverse users beyond the academic community.

Workshop participants suggested that EBRAINS could support this process by providing a list with possible risks, ethical challenges as well as criteria for good practices of AI-related open-source software transfer that researchers can use when developing license agreements in collaboration with the TTOs of their universities (or EBRAINS, if it decides to get involved in this process). They also recommended that the EBRAINS management commission the development of a self-assessment tool that HBP researchers can use to evaluate the potential of their software to lead to dual use, misuse or other problematic and harmful applications, and to identify the conditions under which such problems could emerge.

5.3 Challenges and processes required to ensure the responsible transfer of proprietary software and computational tools

In contrast to open-source transfer, where most of the future users and applications of the software remain unknown to the developer, in the context of the transfer of proprietary software or algorithms, the recipient or recipient organisation engage in a contract and/or license agreement with the software developer. This contract or agreement covers the software with copyrights (to prevent unauthorised copying), patents (that define the conditions for use by the patent owner) or both, as well as protection by contract law (Laurent, 2004). Two implications arise from these differences; a first point is that the researchers that developed the software, or EBRAINS (as the

[14] As clarified in Section 5, at least some of the BICAs that come out of WP3 fulfil this criterion.

institution through which the transfer is mediated) can conduct due diligence procedures (i.e. background checks to learn more about the recipient organisation) to mitigate risks related to the transaction or future use of the transfer software (Demand, 2017). Due diligence and other vetting procedures can provide information about an organisation's legitimacy, location, reputation, financial situation, customer demographics, possible subsidiaries, funders, or partner and parent organisations (Harvey et al., 2013; Lawrence, 2022). However, they do not involve an evaluation of the ethical dimensions and societal or environmental impacts or risks that can emerge from the actual use of the software, in the context of the development and testing of specific AI systems, products or applications (Busby, 2015; Raso et al., 2018; Krieger et al., 2020).

Considering the broad application potential, but also the possible effects (including potential adverse effects or unintended consequences) of brain-inspired AI and robotics applications on human society (Kasabov, 2019; Aicardi et al., 2020; Mehonic and Kenyon, 2022) a systematic assessment of the purposes for which recipient organisations intend to use the licensed software may be helpful for the software provider to make a fully informed decision, and to negotiate the conditions of its use in the license contract. This assessment should cover innovation plans, targeted client groups, business models, and the geographical context(s) in which possible products or applications shall be tested and become available (Demant, 2017; SYNOPSIS, 2022).

Possible measures and processes to address identified challenges and to enable the responsible transfer of proprietary software and computational tools

Workshop participants repeatedly emphasised the importance of due diligence procedures for HBP researchers that plan to transfer proprietary software and algorithms to third parties. They suggested that researchers, together with their institutions, and possibly with the help of EBRAINS, should conduct comprehensive background checks and vetting procedures on recipient organisations that should include, assessment of their reputation, aspects of an organisation's business, and capacity to use the transferred software tools in responsible ways.

To enable software developers to make better-informed decisions and negotiate the conditions of use for the transferred products in a license agreement, participants suggested that routine due diligence procedures (typically performed by universities' TTOs) should be accompanied by a questionnaire with clearly defined disclosure requirements, that explores key aspects of the research or innovation for which the licensed software would be used, such as information on research and innovation purposes, potential users, targeted applications, their potential for dual use, misuse or other problematic uses, the geographical contexts in which possible applications will become available, and other aspects of the innovation and commercialisation process.

Participants proposed that EBRAINS, and other large-scale international digital research infrastructures, could support these processes by providing a checklist with clear criteria and checkpoints for the transfer of AI-related software, algorithms, and related computational devices and tools. Another idea that emerged was that EBRAINS itself could set up a TTO, that would help researchers to assess and navigate the legal and ethical challenges of the transfer of inventions and technologies that are developed using EBRAINS digital services and tools. Still another idea was that EBRAINS could offer fee-based consultancy services to HBP researchers or other EBRAINS users (provided either in-house or through collaboration with external consultants).

5.4 Challenges and processes required to ensure the responsible international transfer of AI-related inventions and technologies developed in WP3

This section examines issues related to the international transfer and commercialisation of software, code, and algorithms that result from research in WP3 and the HBP. Specifically, this section examines the ethical, societal, and institutional governance challenges that arise from the commercial transfer of AI-related inventions and technologies from the HBP and EBRAINS to third party users from academia, industry, and other entities that are situated in different countries in the European Union, and globally. This includes an investigation of issues related to the transfer of AI-related inventions and technologies to states with authoritarian and semi-authoritarian governments, countries involved in military conflicts, as well as jurisdictions with poorly developed, or ineffectively implemented regulatory infrastructures and rules. At the end of each section, we will offer a number of options to address and mitigate some of the identified issues that emerged in the discussions of the workshop. While these do not provide comprehensive solutions, they provide an entry point for ongoing deliberation on how EBRAINS and other infrastructures can navigate the complex challenges related to the cross-border transfer of AI-related innovations. As in Sections 5.2 and 5.3, we draw on findings from our 2022 workshop and complementary literature research.

5.4.1 Challenges Arising from Differences in the Regulation of IP Around the World

At an international level, the commercial transfer of technology and knowledge is regulated under the TRIPS (Trade-Related Aspects of Intellectual Property) agreement, which has established minimum standards for the regulation of IP among the 164 member states of the World Trade Organisation (WTO). TRIPS aims to safeguard the economic rights of creators and inventors and covers various areas of IP, including patents, copyrights, trademarks, and trade secrets (WTO, 2022). All of these can be used to protect the IP of AI-related inventions (Thales, 2022). TRIPS also provides details for the enforcement of IP rights, including rules for injunctions, damages, and other penalties. In doing so, it creates a basis to forestall and prosecute violations of licensing or copyright provisions, for example, in the context of the unauthorised use, theft or copying of AI-related software, tools or technologies, or other uses that are not permitted by the owner (WTO, 2022).

Yet, as a multi-lateral agreement TRIPS only provides a set of baseline standards. Numerous WTO member states have issued further bilateral agreements, to provide additional IP protections (often referred to as TRIPS plus provisions) (Morin and Theriault 2018). Moreover, members are left free to identify the appropriate methods to implement TRIPS rules, depending on national economic circumstances, existing legal systems and other factors (WTO, 2022). For these reasons, there is significant variation in how countries have implemented TRIPS (or TRIPS plus) arrangements, with numerous jurisdictions enforcing these rules only partly or in adaptable ways (Syam, 2022). Large emerging economies such as India and Brazil, for example, have pushed for a flexible implementation of TRIPS, and their efforts have been followed by many other developing countries (Serrano, Oswald, and Burri, 2019). China, on the other hand, has largely embraced WTO rules, but has nevertheless faced ongoing accusations of IP violations, including of illegal and extra-legal access and transfer of foreign AI technologies (Hannas and Chang, 2019; Muehlfield and Wang, 2022).

This heterogeneous situation, which is underpinned by global differences in wealth, social and business cultures, scientific capacities, and political systems, poses various challenges to the commercial, international transfer of AI-related software, tools, algorithms, devices, etc. from HBP researchers and EBRAINS. These include, for example, lack of effective enforcement, insufficient awareness of IPR and IP culture, unauthorised copying and other modes of copyright infringement, inadequate or poorly funded IP administration infrastructures, and other factors (Olubiyi et al., 2022).

Possible measures and processes to address identified challenges and to integrate these in emerging transfer and innovation practices

Participants of the 2022 workshop recommended that EBRAINS supports HBP researchers (and other researchers that will use the EBRAINS platform in the future) with adequate information resources (and if feasible related consultancy services) to understand and navigate the fragmented and complex international IP landscape, which includes multi-lateral, bilateral as well as regional sets of rules for the protection of IP rights. While the identification of the challenges that can arise in the context of the transfer of IP assets to different countries and jurisdictions outside of the EU (e.g. risks arising from variation in legal structures, lack of enforcement, etc.) falls typically under the responsibility of universities' TTOs, the consulted experts suggested that EBRAINS could provide complementary support that specifically addresses IPR challenges and risks of the international transfer of AI-related inventions, tools and technologies (*cf*. WIPO, 2019), or also for other types of technologies that research collaborations with EBRAINS enable.

More specifically, workshop participants proposed three options through which support could be provided: (i) through the development of an online open access information portal that provides researchers and the TTOs of their institutions with relevant information and other resources (e.g. process flow charts, checklists, etc.) to facilitate the IP management process;[15] (ii) through specialist consultancy services for international IP and tech transfer, either in-house or by external professionals, and (iii) through the creation of an EBRAINS TTO that provides legal and business counselling, and actively supports the management of IP assets that emerge from collaborations with EBRAINS or the use of EBRAINS tools and services. The development of an ERBAINS TTO, participants pointed out, could enable tailored IPR strategy support for HBP (and other researchers), including the assessment of infringement risks and other challenges, the development and negotiation of international license agreements and patents, advice on IP insurance and coverage, and other aspects).

5.4.2 Illegal or Unauthorised Distribution of AI-related Technologies and Inventions

A more specific challenge concerns the illicit distribution of AI-related software, code, algorithms, and tools, etc., which enables the unauthorised use of these inventions by unidentified third parties, such as, for example, foreign government units, criminal and terrorist organisations, commercial actors, or other entities (Rasser et al., 2019). Software is intangible and easy to transfer across borders, which makes is simple to circumvent export or other trade controls (Carrozza et al., 2022). Illicit transfer can involve the legal acquisition of IPR protected inventions, tools or technologies by 'legitimate' actors that serve as shadow subsidiaries, which then transfer the acquired product to undisclosed end users (Darrowby, 2019). Workshop participants stated that this is a possible scenario for international infrastructures such as EBRAINS but can also happen to individual researchers and/or their institutions. Entities that operate as subsidiaries can be academic, commercial, or non-governmental actors (Hannas and Chang, 2021). They can, but do not necessarily have to, be located in the country of the intended end users. In principle, shadow organisations that license or acquire inventions from the HBP and EBRAINS can be based in the EU and transfer their acquisitions to recipients internationally, or in EU countries (Reisman, 2006; Hannas and Chang, 2021). Illicit transfer can also include cyber theft and software piracy, where an AI-related software product sold under a single-use license is illegally distributed, copied and used by multiple users for unknown purposes (Rasser et al. 2019; Hannas and Chang, 2021).

Possible measures and processes to address identified challenges and to integrate these in emerging transfer and innovation practices

In workshop discussions on how HBP researchers could mitigate risks of illegal or unauthorised transfer of AI-related inventions and technologies, a variety of options emerged. In addition to the importance of due diligence and other background checks and vetting procedures of recipient organisations (as mentioned in Section 5.3), participants stressed the need for legally binding disclosure obligations that would require potential recipient organisations to reveal associate organisations, collaborators, or other entities or users that may obtain access to the transferred product in the future. This suggestion complements previous suggestions for mandatory disclosure in Section 5.3 about targeted applications, business models, and other aspects of the innovation processes for which transferred inventions and technology shall be used. Participants stressed that, although disclosure of this information may not be enforceable, it may provide a basis to take legal action if illegal distribution happens.

Workshop participants also proposed the development and use of a checklist to recognise indicators of intent to illegal technology transfer. Such a checklist could help HBP or other researchers to recognise possible signs for plans of unauthorised transfer. Indicators that were suggested included reluctance to provide information about the end uses of the transferred invention, or about whether the transferred technology shall be used for domestic use or export, or for use by entities or partners

[15] An existing resource published by the World Intellectual Property Organization, that the recommended EBRAINS information portal could refer to is here: https://www.wipo.int/reference/en

in other global regions.[16] Participants also recommended considering technical solutions to prevent and/or track illegal transfer, such as digital fingerprinting that allows the tracking of unauthorised distribution, with the aim to identify possible users and uses, and to use this information as a basis for legal action.

5.4.3 Transfer to Authoritarian Regimes or Societies with Authoritarian Tendencies

According to The Economist's 2021 Democracy Index which assesses the state of democracy in 167 countries, 59 of these countries have authoritarian governments, 34 countries are hybrid regimes (that combine democratic features with autocratic ones, such as pressure on political opposition and the media, non-independent judiciaries, and election manipulation, etc.), 53 countries are classified as flawed democracies (in which free and fair elections are honoured, but media, political opposition and critics are sometimes suppressed), and only 21 countries as full democracies (The Economist, 2021). The European Union itself has experienced a partial decline of democracy, with some countries displaying authoritarian tendencies (Sesko, 2021). Many authoritarian and hybrid regimes have imposed oppressive actions against their citizens, deploying increased surveillance, manipulating information and elections, and violating human rights. Some of these countries have heavily invested in technology and innovation and support the development of their domestic innovation base with the acquisition and transfer of inventions and innovations from overseas, including ICT, AI and other digital technologies. China, for example, has an extensive technology transfer programme, through which it tries to access foreign AI technologies, that comprises both legal, extra-legal ('grey area'), and illegal forms of transfer (Hannas and Chang, 2019).

Brain-inspired AI systems and components developed in the HBP (and in the future through EBRAINS) can, in principle, inform the development of a variety of technologies with dual use and other problematic purposes that have the potential to further the interests of authoritarian regimes, as well as hybrid regimes and democracies that display authoritarian or other suppressive tendencies. These include the creation of advanced military and defence technologies, AI-powered surveillance, manipulation of public opinions, and other applications that help to solidify the power of these regimes and their control over citizens (Unver, 2018; Aicardi, et al., 2018; Feldstein, 2019; Polyakova and Meserole, 2019).

These factors place specific responsibilities on HBP researchers (and, moving forward, EBRAINS) that aim to transfer AI-related tools, components and inventions to public and private sector organisations in authoritarian countries, as well as other societies that express oppressive tendencies towards their citizens, or specific individuals or groups of their populations, including in the European Union.

Possible measures and processes to address identified challenges and to integrate these in emerging international transfer and innovation practices

Participants of the 2022 workshop proposed that EBRAINS could support the development of a comprehensive assessment procedure that HBP and other researchers (or the TTOs of their institutions) can use, to evaluate whether the planned transfer of software or other technologies to authoritarian countries or societies with authoritarian tendencies, including in the EU, could lead to the misuse or dual use of the transferred inventions, or to other problematic applications, that could result in the violation of human rights, crimes against humanity, or processes of racial or ethnic discrimination, segregation and apartheid. They also suggested that EBRAINS could introduce an internal procedure to check whether the planned international transfer of software or other technologies violates existing international sanctions regimes.

[16] An example of such a checklist developed by the US Department of Commerce's Bureau of Export Administration is here: https://www.bis.doc.gov/index.php/all-articles/23-compliance-a-training/51-red-flag-indicators (Please note, we only include this list as an illustration; the use of such a list by EBRAINS would require significant adjustments.)

5.4.4 Countries Involved in War or Supporting Military Action

According to the World Population Review 29 countries across the world are currently involved in war and military conflicts. These include occupation of foreign territories, civil wars, ethnic violence, and drug wars - as well as terrorist and paramilitary insurgencies (World Population Review, 2023). Other countries or global regions face the threat of military intervention in the future. While the global Arms Trade Treaty and other regulations try to restrict the international weapons trade, the worldwide transfer of military technologies is booming. The global trade of weapons is dominated by around 25 countries, with the USA and Russia (prior to the Ukraine war) accounting for more than half of global arms exports, followed by France and China (Stebbins, 2022).

AI and other digital technologies play an increasing role in warfare, not only in the development of autonomous weapon systems, but also in cyber and information warfare (US Naval Institute, 2022). Military competition and concerns about a global AI arms race have reportedly accelerated the procurement, testing and adoption of AI technologies for defence purposes in many countries, and led to new strategic partnerships (Scharre, 2021). China's Digital Silk Road initiative, for example, has created a new sphere of technological influence, with partnerships reaching from Latin America and the Caribbean to Africa, Central Asia, the Middle East, and Eastern Europe, enabling new possibilities for digital technology cooperation and military engagement (Heidbrink and Becker, 2022). Russia has also heavily invested in AI in the last decade and given particular attention to AI-driven military technologies (Polyakova, 2018). Whilst, with the war in Ukraine, access to foreign AI has become more difficult, Russian-Chinese collaboration in defence technology have opened new opportunities (Rossiter, 2022).

A key challenge to HBP researchers and EBRAINS is that many AI systems are multi-purpose, enabling technologies that can be used for a wide range of applications, including for military and other dual use purposes. Current research in WP3 on the modular integration of different brain-inspired cognitive architectures, which are able to combine divergent motor, perceptual and cognitive functions (Senden, 2022), exemplifies this point. If the longer-term vision of WP3 can be realised (that different architectures can be used as modules by other research teams around the world and be combined and recombined in larger cognitive architectures for an extensive variety of AI and neurorobotics applications (HBP, 2022)), interest in the military potential of these computation tools is likely to emerge.

This puts great responsibility in the hands of HBP researchers and EBRAINS to ensure that research results are transferred to trustworthy recipients and will not be used to the detriment of human rights, fundamental freedoms, or the dignity of individuals and communities (*cf.*, UNESCO, 2017). In the European Union, the use of technologies and other potential dual use items (including software, digital technologies and computing devices) is regulated through the updated 2021 Dual Use Exports Controls legislation (European Union, 2021a). This legislation aims to control the export, trade and provision of technical assistance to countries outside the EU for the development of dual use applications, such as the development and use of weapons of mass destruction or cyber-surveillance (European Union, 2021b). Transfer of technological components, tools or inventions that classify as dual use items, to non-EU countries, requires export authorisation through competent authorities, such as the State Agency for Export Control of Dual Use Goods (SBDU) in France, or the Federal Office of Economics and Export Control (BAFA) in Germany (European Union 2021c). However, with the exception of particular sensitive issues, dual use items may be traded freely between EU countries (European Union, 2021b).

Commentators of the EU's 2021 Dual Use legislation have argued that AI technology can pose specific problems for export controls, because key aspects of these technologies are based on software, that can be easily updated and rewritten. This means, that an initial export may have received export authorisation, but a subsequent update modifies the software to a product that would not have passed controls (Carrozza et al., 2022). AI software is also often offered as a service, and not as a standalone product that is exported to a single recipient under clearly defined conditions. AI services can potentially be used by a wide range of users, in different countries, which may cause difficulties in controlling who uses the AI system, and for which purposes (*ibid.*). Other have suggested that, because export licensing still resides with national authorities, the EU's dual use export regime is vulnerable to incoherent implementation across member states, which could give rise to divergence of the ways in which dual-use sanctions are applied in practice, with some countries applying a less stringent approach (Meissner and Urbanski, 2022). Still others have commented, that with the transitioning of some EU countries into 'soft authoritarianism', the option to freely move dual use items between EU member states could enable new forms of oppression and human rights violation in these countries (Lynch, 2022).

Possible measures and processes to address identified challenges and to integrate these in emerging international transfer and innovation practices

Workshop participants suggested that EBRAINS directorial and management board should carefully consider the limitations of existing EU export controls for the international transfer of AI-related software and technologies (and possibly other inventions and technologies that the use of EBRAINS services and tools enables), and to commission the development of an internal procedure that would allow to assess the potential for dual use and misuse of technologies that shall be transferred on a case-by-case basis. As Carrozza et al. (2022) have stressed above, in the case of the transfer of AI software, such assessments should not only examine the (immediate) dual use or misuse potential of the software as transferred, but also consider the possibility that it might be modified or updated for possible dual use purposes by third parties in the future.

While checks of EU sanction regimes and the filing of export control applications fall under the responsibility of researchers' home institutions, participants proposed that EBRAINS could encourage or possibly require additional checks, such as the use of so-called 'entity lists' that file parties of concern in export transactions, such as organisations, firms, individuals, governments and other entities that have been involved in criminal activities, human rights violation, terrorist funding, financial support for problematic military operations, or in technology cooperation that further the interests and oppressive actions of authoritarian governments, or other problematic activities. Examples of such lists are the Australian China Defence University Tracker,[17] that provides an overview of public research organisations in China involved in military, defence and intelligence research, or the US Departments of Commerce's Bureau of Industry and Security's Lists of Parties of Concern,[18] which publishes the names of foreign organisations, persons and entities that are subject to specific license requirements for the export and transfer of specified items, including software and other technology products.

5.4.5 International Transfer in a Heterogeneous AI Regulatory Landscape

The regulatory situation for the development and deployment of AI systems is an emerging issue and varies significantly between jurisdictions globally (Jobin, Ienca, and Vayena, 2019). Whilst supranational bodies such as OECD, the World Economic Forum and UNESCO have provided various international guidelines, the rules that guide AI research and applications at national and regional levels vary widely. While many countries have in recent years adopted AI-specific guidelines, action plans and policy papers, in others the regulation of AI and related applications is still minimal (Smuha, 2021). In some jurisdictions, although regulatory frameworks are nominally in place, effective implementation is hindered due to a lack of resources, inadequate institutional structures, and other factors (Science Business Net, 2022). Laws that protect personal data and privacy are also often absent, or inefficiently enforced, in many societies (UNCTAD, 2023).

This heterogeneous regulatory situation creates ethical and societal risks that researchers in the HBP and EBRAINS who seek to commercialise AI-related inventions should be aware of. One risk is that stakeholders and companies from countries with low or insufficiently developed regulatory infrastructures purchase or access AI-related technologies or services via the HBP and/or EBRAINS and use these in the context of premature or inadequately tested (commercial) AI applications, which can lead to accidents or other safety issues (*c.f.*, Cheatham et al., 2019). Another challenge is what is sometimes referred to as 'ethics dumping', where researchers or companies from more

[17] https://unitracker.aspi.org.au
[18] https://www.bis.doc.gov/index.php/policy-guidance/lists-of-parties-of-concern

stringently regulated countries (for example in the EU) move their research and R&D operations to leniently regulated jurisdictions. The aim of this is to avoid regulatory restrictions and related expenses 'at home', or also to develop, test, and commercialise applications that are prohibited or discouraged in other world regions (Schroeder et al, 2019). Dual use regulation also varies widely across the world, providing new opportunities to develop AI-based weapon systems in less stringently regulated jurisdictions, which can then be exported and used in other countries (Ienca and Vayena, 2018). In all these scenarios, local citizens, communities, and infrastructure are subjected to possible risks, including risks to physical wellbeing, psychological integrity, or also misuse of personal data, and other unintended consequences of premature, unregulated or irresponsibly regulated AI applications.

Possible measures and processes to address identified challenges and to integrate these in emerging international transfer and innovation practices

Participants of the workshop suggested that HBP and institutions that plan to transfer inventions or technologies developed with the support of EBRAINS, should actively consider the potential consequences of the distribution of their research to societies with low or insufficiently developed regulatory infrastructures. They proposed that EBRAINS should support this process through development of an internal ethical review process, conducted for example by an EBRAINS institutional review board, that would evaluate the possible risks that arise from the transfer of a technology or invention developed with the help of EBRAINS on an individual basis. This approach would resemble institutional ethics committee approval in medical research. A designated group of multi-disciplinary experts makes an independent assessment of the safety and broader societal implications of the transfer of AI-related technologies to specific jurisdictions, advising to abstain from the transfer if identified risks are too high (Tiell, 2019; Veliz, 2019).

5.4.6 Forced Technology Transfer (FTT)

Forced transfer of technology is also a potential cause for concern in relation to the work being undertaken in WP3. The term refers to practices in which 'governments force foreign businesses to share their tech in exchange for market access' (Frankenfield et al., 2022). For HBP researchers, this may primarily be of relevance if they collaborate with a company that aims to enter the market in a country in which FTT policies are in place. Forced technology transfer differs from mutually agreed transfer conditions as required under TRIPS in the context of transfer to the least developed countries in the world, where IP holders from developed countries are advised to share IP to support the development of a viable technological base in these countries (Andrenelli, Gourdon, and Moïse, 2019). FTT can take place through formal policies, but can also involve informal or illicit practices, including the hidden transfer of knowledge or technology from research institutes, for example, government departments or military research institutions. Some countries, such as China, also initiate or fund international collaborations with foreign research institutes or support the development of start-ups through 'angel fund' investors, both with the aim to access strategically important technology (Hannas and Chang, 2019).

In an academic context, the sharing of both technology and knowledge occurs often in a more informal way in research collaborations at the individual or team level, without licensing or formal transfer protocols (Link et al., 2007). Whilst informal sharing may be a practical workaround to perceived institutional bureaucracy, at an international level it is not without risks. In China, for example, there exist policies that can be used to ask research institutes and companies to hand over data or technology to government bodies, including military and intelligence research units. This means that collaborations with Chinese universities can be leveraged for domestic military, surveillance or other purposes in China (Saraswat, 2020). China also uses front companies to conceal the hand of the government obtaining access to research results and technology (NCISC, 2018). Other countries are likely to pursue similar policies. The severity of the risks associated with such leaks depends on the context, types of users, and purposes for which the transferred technology could be used. Loss of control can lead to possible misuse or dual use of concern applications, or more generally, unanticipated purposes that move beyond those considered or agreed by the IP owner.

Possible measures and processes to address identified challenges and to integrate these in emerging international transfer and innovation practices

Participants stated that EBRAINS, as well as HBP and other researchers that plan to internationally transfer technologies developed with the support of EBRAINS, should be actively aware of the different (overt and hidden) forms through which FTT happens in some countries, and to assess the risk of FTT on a case-by-case basis. An option that was suggested, was to examine the involvement of potential recipient organisations in national defence research or their involvement in cyber espionage through the use of international lists of parties of concerns (as already mentioned in Section 5.4.4), for example, the China Defence University Tracker,[19] which has been published by the Australian International Cyber Policy Centre.

5.4.7 Transfer of Datasets

Previous research in T3.8 (Stahl and Leach, 2022) has highlighted some further potential issues around ITT in relation to the possibility of bias in datasets, which various AI systems are being trained on. Respondents to the HBP-wide survey conducted as part of T3.8 were asked if the AI system they worked with was being trained on data that may have issues of bias, discrimination, or that may suffer from a lack of diversity, with a number of respondents indicating that these issues may be present in their dataset. The identified bias stems from the fact that the datasets in use are Eurocentric in nature - a feature of a system being used and deployed exclusively in an EU context, which may become problematic were the system to be deployed in other, non-EU areas (Ntoutsi et al., 2020).

If an AI system is transferred to another region, country or society, and it is trained on data that are not representative of the area to which they are being applied, there are a number of performance risks that might occur. Firstly, there is a risk that the system may not work as anticipated in the new context – for example, the digital soap dispenser that was trained on a dataset of white hands being unable to fulfil its purpose when presented with hands with darker skin tones, the pulse oximeter that overestimates O2 levels in people of colour, or the auto caption software inaccurately predicting non-received pronunciation accents (Waight et al., 2022). Secondly, there is a risk that the system performs in a biased way and could produce discriminatory outcomes for the new population within or upon which the system is deployed, for example western datasets being used to train financial decision systems in used in African countries which discriminates against certain ethnic groups, restricting their access to financial products (Access Partnership, 2018). Whilst this has not been identified as a specific concern related to WP3 research at this time, we have decided to include it in this section, as it may constitute a future issue related to research in this area depending on the future possibilities for transfer and commercialisation of WP3 research.

Possible measures and processes to address identified challenges and to integrate these in emerging international transfer and innovation practices

One way to address this challenge is that, when researchers consider the transfer or commercialisation of technologies to a different global context, they re-examine the data that were used to train the model, to determine if there are any specific issues related to bias in datasets that might be problematic in the deployment of the system outside of the EU context (*cf*. Ntoutsi et al., 2020).

6. Conclusions

This Deliverable reports work undertaken by T3.8 of HBP WP3 with the aim to advance the reflection on brain-inspired AI and to define feasible conceptual and ethical tools for assessing the consistency of AI-related work in WP3 and the HBP within the RRI framework to ensure that societal benefits aimed for are achieved. To achieve this aim, this Deliverable has summarised findings from our analysis of (i) the processes required for responsible work on adaptive brain-inspired cognitive architectures. It has also described (ii) how these processes can be integrated into the work of emerging innovation activities in the HBP and EBRAINS.

[19] https://unitracker.aspi.org.au

We have structured our analysis around three key areas that are critical to support responsible innovation practices in the context of AI-related work in the HBP. These areas address (i) issues related to the actionability of European and other relevant AI regulatory documents and guidelines (which was explored in Part 2), (ii) the conceptual analysis of key terms for the ethical analysis of brain-inspired AI, bio-inspired modelling, and related applications (explored in Part 3), and (iii) the investigation of the social and ethical issues that arise in the context of the exploitation, commercialisation and transfer of AI-related inventions and technologies developed in WP3, in the context of the European Union (EU) and internationally (explored in Part 4).

Part 2 has suggested that a possible strategy for improving the impact of AI regulation, is to interface AI ethics with neuroethical reflection, and that interaction between these two fields would go a long way in addressing some of the challenges posed by existing AI guidelines and regulations. It has proposed a conceptual methodology that facilitates clarification of the possible impact that AI language and related conceptual repertoire can have on traditional ethical notions such as autonomy, free-will, and responsibility, which play a central in AI regulation, but which are often insufficiently defined and operationalised. By providing clarity, this methodology could make AI guidelines less abstract and more effective and action inspiring. This part has also suggested that brain-inspired AI research in the HBP and EBRAINS should be complemented not only by ad hoc information about relevant regulation (both soft and hard), but also by further implementation of RRI activities, including anticipation of critical ethical issues that emerge in the context of exploitation, innovation processes and emerging applications of brain-inspired AI systems.

Part 3 has shown that a systematic conceptual clarification, understood as the description and analysis of key terms that are ethically relevant is crucial for a balanced ethical analysis of AI. As illustrated, a key contribution of conceptual analysis to AI ethics includes its role in reducing the gap between scientific research and public understandings in the HBP, EBRAINS and beyond. To avoid this challenge, and to create a shared understanding of the languages used by AI researchers and different groups of the public, the creation of interactive spaces that facilitate mutual learning, and the identification of possible connection is needed. The collaboration between researchers in AI and robotics on the one hand and researchers in the humanities and the social sciences is a starting point to achieve but should be expanded to include a more direct dialogue with potential technology users and other stakeholders and groups involved in, or affected by, the development of brain-inspired AI and robotics applications.

Part 4 has documented that awareness of, and active engagement with, the ethical and social aspects of the exploitation, transfer and commercialisation of AI-related research outcomes is a crucial aspect to implement responsible innovation practices in the HBP, EBRAINS and other projects and infrastructures that target the development of brain-inspired AI systems and applications. Our analysis has shown that both the open and the proprietary transfer of AI-related software, code, algorithms and other computational tools that come out of WP3 and the HBP, raises critical ethical issues that need to be evaluated and addressed on a case-by-case basis. This part has described possible measures and processes to address identified challenges, and to facilitate responsible transfer and innovation practices, as research in the HBP reaches its final stages and is (in some cases) continued through the services and tools provided by EBRAINS. The suggested options have been defined to support EBRAINS and other digital research infrastructures in the development of robust and appropriate technology transfer protocols and policies, and to support HBP and other researchers in identifying and addressing the ethical risks and managerial and legal challenges that can arise as their research reaches the exploitation stage, and is commercialised, distributed and used by public and private sector entities in both the EU and internationally.

Appendix: References

Access Partnership (2018). Artificial Intelligence for Africa: An Opportunity for Growth, Development and Democratisation, Access Partnership, University of Pretoria, URL: https://accesspartnership.com/wp-content/uploads/2022/11/WP-AI-for-Africa-Nov-2018.pdf

AI HLEG. (2019). Ethics Guidelines for Trustworthy AI. European Commission - Directorate-General for Communication. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

AI HLEG. (2020). Assessment List for Trustworthy AI (ALTAI). European Commission. https://futurium.ec.europa.eu/en/european-ai-alliance/pages/altai-assessment-list-trustworthy-artificial-intelligence

Aicardi, C., Akintoye, S., Fothergill, B., Guerrero, M., Klinker, G., Knight, W., Klüver, L., Morel, Y., Morin, F.O., Stahl, B.C. and Ulnicane, I. (2020). Ethical and social aspects of neurorobotics. Science and Engineering Ethics, 26(5), pp. 2533-2546. http://dx.doi.org/10.1007/s11948-020-00248-8 (P2580).

Aicardi, C., Bitsch, L., Bang Bådum, N., Datta, S., Evers, K., Farisco, M., Fothergill, T., Giordano, J., Harris, E., Jørgensen, M. L., Klüver, L., Mahfoud, T., Rainey, S., Riisgaard, K., Rose, N., Salles, A., Stahl, B., Ulnicane, I. (2018) Dual Use Opinion. https://sos-ch-dk-2.exo.io/public-website-production/filer_public/f8/f0/f8f09276-d370-4758-ad03-679fa1c57e95/hbp-ethics-society-2018-opinion-on-dual-use.pdf

Aicardi, C., Akintoye, S., Fothergill, B., Guerrero, M., Klinker, G., Knight, W., ... & Ulnicane, I. (2020). Ethical and social aspects of neurorobotics. Science and Engineering Ethics, 26(5), 2533-2546.

Aicardi, C., Bitsch, L., Burton, S.D., Evers, K., Farisco, M., Mahfoud, T., Rose, N., Rosemann, A., Salles, A., Stahl, B.C. and Ulnicane, I. (2021.) Opinion on Trust and Transparency in Artificial Intelligence. https://zenodo.org/record/4588648#.Y8GLGS2catA (P3056)

Andrenelli, A., Gourdon, J., and Moïsé, E. (2019). International technology transfer policies. https://doi.org/10.1787/7103eabf-en

Arcesati, R. and Meinhardt, C. (2021). China's Open-Source Tech Development. Mercator Institute for China Studies. URL: https://merics.org/sites/default/files/2021-05/MERICS%20Primer%20Open%20Source%202021_0.pdf

Baark, E. (2019). China's Indigenous Innovation Policies. East Asian Policy, 11(02), pp. 5-12.

Basu, S., and Taylor, R. (2010). Use of Consultants in Technology Transfer. AUTM Technology Transfer Practice Manual,, 4(Part 1).

Bentley, P.J., Brundage, M., Häggström, O., and Metzinger, T. (2018). Should we fear artificial intelligence? In In-depth Analysis (Brussels: European Union - STOA).

Bernhardt, M. (2021) Open-source software: freedom from ethics. E&T Engineering and Technology. URL: https://eandt.theiet.org/content/articles/2021/04/open-source-software-freedom-from-ethics/

Busby, L. (2015). The Limits of Due Diligence: Edited by Lorraine Busby. The Serials Librarian, 69(1), 13-19.

Carrozza, I., Marsh, N., and Gregory, M. Reichberg (2022) Dual-Use AI Technology in China, the US and the EU: Strategic Implications for the Balance of Power, PRIO Paper. Oslo: PRIO. URL: https://www.prio.org/publications/13150

Cheatham, B., Javanmardian, K., and Samandari, H. (2019). Confronting the risks of artificial intelligence. McKinsey Quarterly, 2, 38.

Coeckelbergh, M. (2020). AI Ethics. Cambridge, MA; MIT Press.

Conradt, J., and Morel, Y. (2022) Brain-based technology to support safe, human-robot collaboration. 28.10.2022. https://www.youtube.com/watch?v=jvEriT9UMLY

Corbly, J. E. (2014). The Free Software Alternative: Freeware, Open Source Software, and Libraries. Information Technology and Libraries. 33(3): p. 65. https://doi.org/10.6017%2Fital.v33i3.5105

Council of the European Union (2022). Recommendation for a COUNCIL DECISION authorising the opening of negotiations on behalf of the European Union for a Council of Europe convention on artificial intelligence, human rights, democracy and the rule of law. URL: https://data.consilium.europa.eu/doc/document/ST-11818-2022-INIT/en/pdf

Darrowby, J. (2019). Intellectual property, artificial intelligence, and ethical dilemmas: China and the new frontiers of academic integrity. Made in China Journal, 4(1), 24-28.

Dehaene, S., Lau, H., and Kouider, S. (2017). What is consciousness, and could machines have it? Science 358, pp. 486-492.

Demant, C. (2017). Software Due Diligence: Softwareentwicklung als Asset bewertet. Springer-Verlag.

Descartes (2022) Export Control System (EU): A Quick Overview & Resources. URL: https://www.descartes.com/resources/knowledge-center/export-control-system-eu-quick-overview-resources

European Commission. (2018). Statement on Artificial Intelligence, Robotics and ´Autonomous´ Systems. Brussels; Publications Office.https://data.europa.eu/doi/10.2777/531856

European Commission. (2019). Building Trust in Human-Centric Artificial Intelligence. Brussels; European Parliament.

European Commission (2020) Emerging Technologies - Developments in the Context of Dual-Use Export Controls: Factsheets. URL: https://trade.ec.europa.eu/doclib/docs/2021/september/tradoc_159791.pdf

European Commission. (2021). Horizon Europe, open science: early knowledge and data sharing, and open collaboration. Brussels; Publications Office. https://data.europa.eu/doi/10.2777/18252

European Union (2021a). Dual-use export controls. URL: https://eur-lex.europa.eu/EN/legal-content/summary/dual-use-export-controls.html

European Union (2021b). Exporting dual-use items. URL: https://policy.trade.ec.europa.eu/help-exporters-and-importers/exporting-dual-use-items_en

European Union (2021c). National authorities empowered in the member states to grant export licences, to prohibit the transit of non¬community dual-use items and to grant authorisations fortheprovisionofbrokeringservices. URL: https://trade.ec.europa.eu/doclib/docs/2019/september/tradoc_158374.pdf

Evers, K., and Salles, A. (2021). Epistemic challenges of digital twins & virtual brains: Perspectives from fundamental neuroethics. SCIO Journal of Philosophy, 21, pp. 27-53.http://dx.doi.org/10.46583/scio_2021.21.846 (P3011)

Farisco, M., and Evers, K. (2017). The ethical relevance of the unconscious. Philos Ethics Humanit Med, 12, p. 11. http://dx.doi.org/10.1186/s13010-017-0053-9 (P1244)

Farisco, M., Baldassarre, G., Cartoni, E., Leach, T., Petrovici, M. A., Goebel, R., Rosemann, A., Salles, A., Stahl, B., van Albada, S.J. (In progress) A Model for the Ethical Analysis of Brain-Inspired AI.

Farisco, M., Evers, K. & Salles, A. (2020). Towards Establishing Criteria for the Ethical Analysis of Artificial Intelligence. Sci Eng Ethics 26, pp. 2413–2425. https://doi.org/10.1007/s11948-020-00238-w (P2577)

Farisco, M., Evers, K. & Salles, A. (2022). On the Contribution of Neuroethics to the Ethics and Regulation of Artificial Intelligence. Neuroethics. 15, 4. https://doi.org/10.1007/s12152-022-09484-0 (P3225)

Feldstein, S. (2019). The road to digital unfreedom: How artificial intelligence is reshaping repression. Journal of Democracy, 30(1), 40-52.

Floreano, D., and Mattiussi, C. (2008). Bio-Inspired Artificial Intelligence. Cambridge, MA; MIT Press.

Floridi, L. (2013). The ethics of information. Oxford; OUP.

Floridi, L. (2019). Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. Philosophy & Technology. 32. pp. 185-193.

Floridi, L., and Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review. 1(1).

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F. and Schafer, B. (2018). AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds & Machines.28 (4). pp 689–707.

Flynn, C. (2020). Recommendations on Export Controls for Artificial Intelligence. Centre for Security and Emerging Technology.

Frankefield, J., Rasure, E. and Skylar, C. (2022) Forced Technology Transfer. Investopedia. URL: https://www.investopedia.com/forced-technology-transfer-ftt-4687680#:~:text=Forced%20technology%20transfer%20(FTT)%20is,practice%20is%20common%20in%20China.

Future of Life Institute (2022) General Purpose AI and the AI Act. URL: https://artificialintelligenceact.eu/wp-content/uploads/2022/05/General-Purpose-AI-and-the-AI-Act.pdfGardner, H. (1985). Frames of mind: the theory of multiple intelligences. London: Heinemann.

Goebel, R. (2022). Introduction. HBP WP3 Workshop Amsterdam, 20.09.2022.

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. Minds & Machines 30. pp 99-120.

Hannas, W. C. and Chang, H. M. (2021). China's "New Generation" AI-Brain Project. PRISM. 9(3), pp.18-33.

Hannas, W.C., and Chang, H.M., (2019). China's Access to Foreign AI Technology: An Assessment. Center for Security and Emerging Technology. https://doi.org/10.51593/20190053

Harvey, M., Tihanyi, L., Novicevic, M. M., & Dabic, M. (2003). Technology transfers to Central and Eastern Europe: developing an adequate due diligence format. Journal of East-West Business, 8(2), 5-38.

HBP (2022). Brain-inspired artificial intelligence. Human Brain Project website. URL: https://www.humanbrainproject.eu/en/science-development/focus-areas/brain-inspired-cognitive-architectures/

Heidbrink, C. and Becker, C., (2022). Framing the Digital Silk Road's (De) Securitisation. Journal of Current Chinese Affairs.

Hildt, E., Laas, K., and Sziron, M. (2020). Editorial: Shaping Ethical Futures in Brain-Based and Artificial Intelligence Research. Sci Eng Ethics. 26,5. pp. 2371-2379.

IEEE (2019). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.

Ienca, M., and Vayena, E. (2018). Dual use in the 21st century: emerging risks and global governance. Swiss Medical Weekly, 148(4748), p.w14688.

Jobin, A., Ienca, M., and Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence. 1. pp. 389-399;

Kasabov, N. K. (2019). Time-space, spiking neural networks and brain-inspired artificial intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg.

Kirtay, M., Vannucci, L., Albanese, U., Laschi, C., Oztop, E., and Falotico, E. (2019). Emotion as an emergent phenomenon of the neurocomputational energy regulation mechanism of a cognitive agent in a decision-making task. Adaptive Behavior. 29(1). pp 55-71.

Koch, C. (2019). The feeling of life itself: why consciousness Is widespread but can't be computed. Cambridge, MA; MIT Press.

Krieger, H., Peters, A., & Kreuzer, L. (Eds.). (2020). Due diligence in the international legal order. Oxford University Press.

Lake, B.M., Ullman, T.D., Tenenbaum, J.B., and Gershman, S.J. (2017). Building machines that learn and think like people. Behav Brain Sci. 40.

Laurent, A. M. S. (2004). Understanding open source and free software licensing: guide to navigating licensing issues in existing & new software. O'Reilly Media: Sebastopol, USA.

Lawrence, G. M. (2022). Due Diligence in Business Transactions. Law Journal Press.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature. 521. pp 436-444.

Legg, S., and Hutter, M. (2007). A Collection of Definitions of Intelligence. In: (Eds) Goertzel, B., and Wang, P. Advances in Artificial General Intelligence: Concept, Architectures and Algorithms. Amsterdam; IOS Press.

Link, A.N., Siegel, D.S. and Bozeman, B., (2007). An empirical analysis of the propensity of academics to engage in informal university technology transfer. Industrial and corporate change. 16(4). pp.641-655.

Lorinczi, L. (2017). Dissemination and Exploitation of Research and Innovation Project Results. Horizon 2020 Coordinators' Day. URL: 2015/RTD/J5/OP/PP-02682-2015-CSSERR (europa.eu)

Lynch, J. (2022). Iron net: Digital repression in the Middle East and North Africa, Policy Brief, European Council on Foreign Relations. URL: https://ecfr.eu/publication/iron-net-digital-repression-in-the-middle-east-and-north-africa/

Marcus, G., and Davis, E. (2019). Rebooting AI: building artificial intelligence we can trust. New York; Pantheon Books.

Martínez-Zarzoso, I., and Chelala, S. (2021). Trade agreements and international technology transfer. Rev World Econ. 157. pp 631–665. https://doi.org/10.1007/s10290-021-00420-7

Mehonic, A., and Kenyon, A. J. (2022). Brain-inspired computing needs a master plan. Nature, 604(7905), 255-260.

Meissner, K. L., and Urbanski, K. (2022). Feeble rules: one dual-use sanctions regime, multiple ways of implementation and application?. European Security, 31(2), 222-241.

Metzinger, T. (2021). An Argument for a Global Moratorium on Synthetic Phenomenology. Journal of Artificial Intelligence and Consciousness. 8. pp 1-24.

Migliore, M. (2022). Single-shot learning using Hippocampal circuitry. Presentation at workshop: Technology transfer, commercialisation and IPR: Practical, ethical and governance challenges in the transition to a Post-HBP future, 7.11.2022

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. Nat Mach Intell. 1. pp. 501-507.

Morel, Y. (2022). Progress on Cobotics and Logistics Demonstrators. Introduction. HBP WP3 Workshop Amsterdam, 20.09.2022.

Morin, J. F., and Theriault, D. (2018). How Trade Deals Extend the Frontiers of International Patent Law. Centre for International Governance Innovation, Canada. CIGI Papers No. 199. URL: https://www.cigionline.org/publications/how-trade-deals-extend-frontiers-international-patent-law/

Morley, J., Floridi, L., Kinsey, L., and Elhalal, A. (2020). From What to how: An Initial Review of publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics. 26. pp. 2141-2168.

Muehlfeld, K., and Wang, M. (2022). Intellectual Property Rights in China—A Literature Review on the Public's Perspective. Frontiers in Socioliology, 7. https://doi.org/10.3389/fsoc.2022.793165

NCISC (2018). Foreign Economic Espionage in Cyberspace. National Counterintelligence and Security Centre. https://www.dni.gov/files/NCSC/documents/news/20180724-economic-espionage-pub.pdf

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., and Kompatsiaris, I. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 10(3), p. e1356.

Olubiyi, I. A., Emerole, U. A., and Adetula, A. F. (2022). Contemporary Challenges to Intellectual Property Rights in Developing Countries: Looking Beyond the Laws. IIC-International Review of Intellectual Property and Competition Law, 53(1), pp. 5-30.

Pennartz, C., Farisco, M., and Evers, K. (2019). Indicators and Criteria of Consciousness in Animals and Intelligent Machines: An Inside-Out Approach. Frontiers in Systems Neuroscience. 13(25). (P2013)

Polyakova, A. (2018). Weapons of the weak: Russia and the AI-driven asymmetric warfare. Brookings. URL: https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/

Polyakova, A., & Meserole, C. (2019). Exporting digital authoritarianism: The Russian and Chinese models. Policy Brief, Democracy and Disorder Series (Washington, DC: Brookings, 2019), 1-22.

Poo, M.-m. (2018). Towards brain-inspired artificial intelligence. National Science Review. 5(6). p 785.

Prescott, T., and Camilleri, D. (2018). The Synthetic Psychology of the Self. In: Eds. Aldinhas Ferreira, M.I., Silva Sequeira, J., and Ventura, R. Cognitive Architecture. Basel; Springer International Publishing. pp. 85-104.

Ramos, F. (2022). Brain-Inspired Cognitive Architectures for Artificial General Intelligence. Special Issue Call for Papers. Cognitive Systems Research. https://www.journals.elsevier.com/cognitive-systems-research/forthcoming-special-issues/brain-inspired-cognitive-architectures-for-artificial-general-intelligence

Raso, F. A., Hilligoss, H., Krishnamurthy, V., Bavitz, C., & Kim, L. (2018). Artificial intelligence & human rights: Opportunities & risks. Berkman Klein Center Research Publication, (2018-6).

Rasser, M., Lamberth, M., Riikonen, A., Guo, C., Horowitz, M. and Scharre, P. (2019). Counter illicit technology transfer. In: M. Rasser et al. (Eds.) The American AI Century – Report. Center for a New American Security. URL: https://www.jstor.org/stable/pdf/resrep20445.10.pdf?refreqid=excelsior%3A88f1826511965c68485f8203467c922b&ab_segments=&origin=&acceptTC=1

Reber, A.S. (2019). The First Minds: Caterpillars, 'Karyotes, and Consciousness. New York; OUP.

Reisman, A. (2006). A taxonomic view of illegal transfer of technologies: A case study. Journal of Engineering and Technology Management, 23(4), 292-312.

Rességuier, A., and Rodrigues, R. (2020). AI should not remain toothless! A call to bring back the teeth of ethics. Big Data & Society 7(2). pp 1-5.

Rossiter, A. (2022). Russian-Chinese collaboration in defense technology: prospects and limits. Trends Research and Advisory. URL: https://trendsresearch.org/insight/russian-chinese-collaboration-in-defense-technology-prospects-and-limits/

Roth, G. (2013). The Long Evolution of Brains and Minds. Dordrecht; Springer Science.

Ryan, M., and Stahl, B. C. (2020). Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. Journal of Information, Communication and Ethics in Society.

Salles, A., Evers, K., and Farisco, M. (2020). Anthropomorphism in AI, AJOB Neuroscience, 11:2, pp. 88-95, https://doi.org/10.1080/21507740.2020.1740350 (P2506)

Samsonovich, A.V., Gudwin, R.R., and da Silva Simões, A. eds. (2020). Brain-Inspired Cognitive Architectures for Artificial Intelligence BICA* AI 2020: Proceedings of the 11th Annual Meeting of the BICA Society (Vol. 1310). Springer Nature.

Saraswat, A. (2020). Understanding the National Intelligence Law of China: Why India banned Tik Tok?. Diplomatist. URL" https://diplomatist.com/2020/09/05/understanding-the-national-intelligence-law-of-china-why-india-banned-tik-tok/

Scharre, P. (2021). Debunking the AI Arms Race Theory. The National Security Review. URL:https://tnsr.org/2021/06/debunking-the-ai-arms-race-theory/#:~:text=There%20is%20no%20AI%20arms%20race.%20However%2C%20military,leading%20to%20the%20deployment%20of%20unsafe%20AI%20systems.

Schmuhl, M. A., & Na, C. (2019). Globalization and software piracy within and across 103 countries. Crime, Law and Social Change, 72(3), 249-267.

Schmuker, M. (2022). Accelerating gas sensing and improving odour-guided robotic navigation using principles from Neuroscience. Presentation at workshop: Technology transfer, commercialisation and IPR: Practical, ethical and governance challenges in the transition to a Post-HBP future, 7.11.2022

Schroeder, D., Chatfield, K., Singh, M., Chennells, R. and Herissone-Kelly, P. (2019). Equitable research partnerships: a global code of conduct to counter ethics dumping. Springer Nature.

Science Business Net (2022). Is global AI harmonisation actually achievable?. URL: https://sciencebusiness.net/data-rules/news/global-ai-harmonisation-actually-achievable

Searle, J.R. (2007). Biological Naturalism. In: Eds. Velmans, M., and Schneider, S. The Blackwell Companion to Consciousness. Malden MA, Oxford, Victoria; Blackwell Publishing Ltd. pp. 325-334.

Senden. M. (2022) Modular Brain-Inspired Cognitive Architectures. HBP WP3 Workshop Amsterdam, 20.09.2022.

Serrano, O., and Burri, M., (2019). TRIPS implementation in developing countries: likely scenarios to 2025. In: Eds. Elsig, M., Hahn, M., and Spilker, G. The Shifting Landscape of Global Trade Governance. Cambridge; Cambridge University Press. pp. 275-294.

Sesko , Z. (2021). Eastern European countries adopting authoritarian measures in face of Covid. URL:https://www.theguardian.com/world/2021/dec/29/eastern-european-countries-adopt-authoritarian-measures-covid

Smuha, N.A. (2021). From a 'race to AI'to a 'race to AI regulation': regulatory competition for artificial intelligence. Law, Innovation and Technology. 13(1). pp. 57-84.

Stahl, B.C., and 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 and Ethics, pp.1-23. http://dx.doi.org/10.1007/s43681-022-00201-4 (P3701)

Stahl, B.C., Andreou, A., Brey, P., Hatzakis, T., Kirichenko, A., Macnish, K., Laulhé Shaelou, S., Patel, A., Ryan, M., and Wright, D. (2021). Artificial intelligence for human flourishing – Beyond principles for machine learning. Journal of Business Research 124. pp. 374-388.

Stebbins, S. (2022) Special Report: The 25 countries selling the most weapons. 24/7 Wall St. URL: https://247wallst.com/special-report/2022/07/25/the-25-largest-arms-exporting-countries-and-who-they-sell-to/#:~:text=The%20United%20States%20and%20Russia,of%20all%20global%20arms%20exports

Stix, C. (2021). Actionable Principle for Artificial Intelligence Policy: Three Pathways. Science and Engineering Ethics. 27(1). Pp. 1-17.

Syam, N. (2022). A review of WTO disputes on TRIPS: Implications for use of flexibilities for public health.

SYNOPSYS (2022). White Paper: Software Due Diligence – Checklist and Guidelines. URL: https://www.synopsys.com/content/dam/synopsys/sig-assets/whitepapers/wp-software-duediligence-buyersguide.pdf

Taddeo, M. R. (2009). Defining Trust and E-trust: Old Theories and New Problems. International Journal of Technology and Human Interaction.5(2). pp. 23-35.

Thales (2022) Software Intellectual Property: What It Is & How to Protect It. URL: https://cpl.thalesgroup.com/software-monetization/protecting-software-intellectual-property

The Economist (2021). Democracy Index 2021. URL: https://www.eiu.com/n/campaigns/democracy-index-2021/

Tiell, S. (2019). Create an ethics committee to keep your AI Initiative in check. Harvard Business Review, 15.

Tolmeijer, S., Kneer, M., Sarasua, C., Christen, M., and Bernstein, A. (2020). Implementations in Machine Ethics: A Survey. ACM Computing Surveys. 53(6). pp. 1-38.

UNCTAD (2023). Data Protection and Privacy Legislation Worldwide. United Nations Conference on Trade and Development. URL: https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

UNESCO (2017). Recommendations on Science and Scientific Researchers. https://en.unesco.org/themes/ethics-science-and-technology/recommendation_science

Unver, A. (2018). Artificial intelligence, authoritarianism and the future of political systems. EDAM Research Reports.

US Naval Institute (2022). Information Warfare: AI, Cyber and Maneuver. URL: https://www.usni.org/magazines/proceedings/2022/february/information-warfare-ai-cyber-and-maneuver

Vakkuri, V. and Abrahamsson, P. (2018). The Key Concepts of Artificial Intelligence. 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). Stuttgart. pp. 1-6.

Velasco, G. (2022). Innovation Management and Exploitation of Research Results. HBP CDC Training Module, 10.11.2022

Véliz, C. (2019). Three things digital ethics can learn from medical ethics. Nature Electronics, 2(8), 316-318.

Vermeulen, S. (2022). EBRAINS Commercial Options. Presentation at workshop: Technology transfer, commercialisation and IPR: Practical, ethical and governance challenges in the transition to a Post-HBP future, 7.11.2022

Waight, N., Kayumova, S., Tripp, J., and Achilova, F. (2022). Towards Equitable, Social Justice Criticality: Re-Constructing the "Black" Box and Making it Transparent for the Future of Science and Technology in Science Education. Science & Education. pp.1-23.

WIPO (2019) DRAFT ISSUES PAPER ON INTELLECTUAL PROPERTY POLICY AND ARTIFICIAL INTELLIGENCE, World Intellectual Property Organization Secretariat. URL: https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_2_ge_20/wipo_ip_ai_2_ge_20_1.pdf

World Population Review (2023) Countries currently at war 2023. URL: https://worldpopulationreview.com/country-rankings/countries-currently-at-war

WTO (2022) Overview: The TRIPS agreement. URL: https://www.wto.org/english/tratop_e/trips_e/intel2_e.htm