European Innovation Council

PATHFINDER CHALLENGE
Awareness Inside
CHALLENGE GUIDE
PART I

EIC Work Programme reference: HORIZON-EIC-2021-PATHFINDERCHALLENGES-01-03
Call deadline date: 27/10/2021
Programme Manager: Walter VAN DE VELDE (acting)
Challenge page: https://eic.ec.europa.eu/calls-proposals/eic-pathfinder-challenge-awareness-inside_en

CONTENT

The Appendices that are referred to in this document are common to the different Challenges and bundled in Part II of the Challenge Guide, published on the Challenge page on the EIC Website https://eic.ec.europa.eu/calls-proposals/eic-pathfinder-challenge-awareness-inside_en.

1. About this document

The Challenge Guide is the reference document accompanying a Pathfinder challenge along its whole life cycle, from call to achieving its objectives.

The Programme Manager in charge of this Pathfinder Challenge is the editor of the Challenge Guide. The Challenge Guide captures, at any moment, the state of play, achievements and remaining challenges, and documents the process by which the Programme Manager and Portfolio members jointly establish an evolving set of Portfolio objectives and a shared roadmap for achieving them. The most recent version can be found through the corresponding Challenge page on the EIC Website https://eic.ec.europa.eu/calls-proposals/eic-pathfinder-challenge-awareness-inside_en.

The Challenge Guide starts out as a background document to the initial Pathfinder Challenge call. It details the intention of the call by complementing notably the scope, objectives (see Section 5 - Challenge call text) or criteria (see Appendix 3: EIC 2021 Work Programme – Evaluation criteria) set out in the EIC Work Programme. In no case does it contradict or supplant the Work Programme text. After the call evaluation, the Challenge Guide further documents the initial Challenge Portfolio that resulted from the call.

As the actions in the Portfolio unfold, the Challenge Guide further documents the evolving Portfolio Objective(s) and the progress towards achieving them, notably through the Portfolio Activities that the Programme Manager puts in place.

The Challenge Guide serves as a reference for the common understanding, rules-of-play and obligations for the EIC beneficiaries that are involved in the Challenge Portfolio. Contractual Obligations are further reference material from the EIC Work Programme are collected in Part II of the Pathfinder Challenge Guide, published on the Challenge page on the EIC Website https://eic.ec.europa.eu/calls-proposals/eic-pathfinder-challenge-awareness-inside_en.

2. Overall objective of the Pathfinder Challenge

This section sets out the rationale of the Challenge, its scope and explains the overall objectives. This section should be read as further background and guidance to the Challenge specific part of the EIC Work Programme text (see extract in section 5).
Proposals to this Challenge are expected to explain how they relate to and intend to go beyond the state of the art, and how they interpret and contribute to the objectives of the Challenge.

Background – State of the Art and some possible directions[1]

Most scientific and philosophical accounts of awareness are based on a human subject perspective and at an individual level. They address the question of what it means for an individual human subject to be aware of, e.g., the environment, time or oneself and how one can assess awareness in this context. The origins of these subjective experiences has fascinated humankind for a long time but little scientific consensus on these has emerged. Nevertheless the problem is being approached from many angles, with new ideas and methodologies that start to provide partial insights. In a way, this call invites partnerships that bring some of these together to respond to the expected outcomes.

Probably the most challenging form of the Awareness issue, the hard problem of consciousness, was coined by David Chalmers and is widely debated since:[2]

> It is undeniable that some organisms are subjects of experience. But the question of how it is that these systems are subjects of experience is perplexing. Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion? It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does.

While the hard problem of consciousness is in a way the holy grail for awareness research, there are other ‘degrees’ of awareness that appear relevant and could provide stepping stones for the research.

The translation of consciousness findings is especially significant in the clinical domain (Kondziella et al. 2020[3]). The clinical field is one of the areas where consciousness research has progressed, because of the need to assess consciousness of patients in a clinical context.

[1] This call topic originated from a workshop organised by the EU to identify Challenge topics for the EIC WP2021, January 30th, 2020. This section relies on inputs from some of the participants, explaining the heterogeneous style of referencing. Special thanks go to Aureli Soria-Frish and Roumen Borissov for help with background research and textual contributions.
[2] Chalmers, D. (1995). Facing up to the problem of consciousness. *J. Consciousness Stud.* 2, 200–219.
[3] https://onlinelibrary.wiley.com/doi/epdf/10.1111/ene.14151

In medical practice, Steven Laureys linked consciousness to different clinical study fields (Laureys 2005[4]). He represented consciousness in a two-dimensional plot accounting for the content of consciousness or awareness, and the level of consciousness or wakefulness. This two-dimensional plot serves the representation of several consciousness states of clinical relevance (Figure 1).

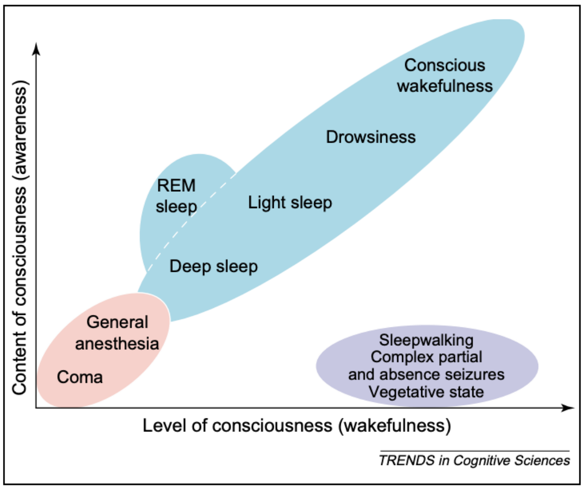

Fig. 1. Laureys (2005) two-dimensional plot representing consciousness states

In patients suffering from Disorders of Consciousness (DOC), which constitutes Laureys starting point, the level of awareness determines its diagnosis as suffering from Unresponsive Wakefulness Syndrome (UWS) or Minimal Conscious State (MCS). There are still misdiagnosis rates of about 40% in patients with MCS (Wang et al. 2020[5]), which make advisable the introduction of objective markers (Kondziella et al. 2020) to characterize consciousness levels. Low level and content of consciousness characterize unconsciousness under anesthesia. Here the main clinical issue is what is the correct dose of anesthetic drug that has to be provided in order to reach the right unconsciousness state avoiding both under- and overdose (Warnaby et al. 2017[6]). Clinical consequences in this case range from traumatic chirurgic experiences in the former to temporary states of delirium in the latter. In this context it is worth pointing out the usage of psychedelic drugs as anesthetics, which challenge the clinical application of consciousness altering substances. Indeed, the employment of psychedelics for the treatment of mental diseases opens a complete new dimension in the study of consciousness as recognized in a relative recent opinion article.[7] Other consciousness clinical fields are related to sleep. Its relationship to the study of consciousness has been addressed for instance in Boly et al. (2013). Lucid dreaming presents the largest awareness degree among sleep stages, which therefore makes it a very special consciousness state. Its study presents an interesting window for the understanding of consciousness with characteristic electrophysiological correlates (Voss et al 2009[8]).

Meditation practices have been associated to consciousness already for a while.

[4] https://pubmed.ncbi.nlm.nih.gov/16271507/
[5] https://bmcneurol.biomedcentral.com/articles/10.1186/s12883-020-01924-9
[6] https://pubs.asahq.org/anesthesiology/article/127/4/645/19832/Investigation-of-Slow-wave-Activity-Saturation
[7] https://www.nature.com/articles/d41586-019-02207-1
[8] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2737577/

Furthermore meditation has been claimed to result in "pure consciousness" in a recent paper (Metzinger 2020[9]). Besides its philosophical dimension, the clinical relevance of meditation has been related to cognitive enhancement and well- being not only in patient populations, but in healthy ones as well, e.g. (Champion et al. 2018[10]).

Computational approaches

Starting from this neuroscientific and clinical interest in the topic of consciousness, its study has reached the development of new technologies, through so-called Machine Consciousness (Koch & Tononi 2017[11]). Further beyond Artificial General Intelligence[12], which aims to confer engineering systems with the capability of solving problems without specifically being programmed for any particular purpose, Machine Consciousness targets the realization of technology able to "subjective feeling", i.e. become sentient and therefore speak, reason, self-monitor and introspect (Koch 2019).

The work in (Dehaene et al 2017[13]) discusses the different computation principles that are requested for conscious behavior. Interesting enough most of current Computational Intelligence systems fulfil the requirements of so-called unconscious processing found in brains, which are categorized as C0. Functions like view-invariant recognition or decision making are based on unconscious neural processes that are already implemented in most modern AI algorithms. Far beyond these, the implementation of genuine Machine Consciousness requires computer systems that implement global availability of relevant information (which is denoted as C1) and self-monitoring of oneself (denoted as C2). The following list illustrates the different functions associated and leading to consciousness, which have been demonstrated in the cognitive neuroscience literature, and their accountability to one or other category (see complete table in (Dehaene et al 2017)):

- C0 Unconsciouss processing
  - Invariant visual recognition
  - Cognitive control
  - Reinforcement learning
- C1: Global availability of information
  - Stabilization of short-lived information for off-line processing
  - Flexible routing of information
  - Sequential performance of several tasks
- C2: Self-monitoring
  - Self-confidence
  - Evaluation of one's knowledge
  - Error detection

Not only systems showing these behaviors should be implemented for achieving consciousness, but also the interaction between C1 and C2, which is sometimes diverging

[9] https://philosophymindscience.org/index.php/phimisci/article/view/46
[10] https://pubmed.ncbi.nlm.nih.gov/30596696/
[11] https://spectrum.ieee.org/computing/hardware/can-we-quantify-machine-consciousness
[12] http://www.scholarpedia.org/article/Artificial_General_Intelligence
[13] https://science.sciencemag.org/content/358/6362/486

sometimes converging. (Dehaene et al 2017) state that such systems can be implemented and gives some examples of already existing computation architectures that fulfil these principles. In some cases they seem to offer clear performance advantages and represent an advance towards Artificial General Intelligence (Fernando et al 2017[14]).

Embodiment

Within general machine consciousness the research field of Conscious Robots is gaining on relevance (Chella et al 2019[15]). Although the topic is primarily focused in implementing robots that present visual experiences, bodily sensations, mental images, or emotions among other features, the ultimate goal seems addressing self-awareness in an embodied set-up. Indeed some works claim that selfhood is based on the existence of a body (e.g., Seth and Tsikris 2018[16]).

Besides the ethical dimension, the controversy between the implementable and non-implementable nature of Machine Consciousness seem to be a dispute among apparently competing consciousness theories. The role of the physical substrate and its emergent associated properties (Koch & Tononi 2017), the computational principles in which consciousness is based (Dehaene et al 2017), and the requirement of embodiment (Seth and Tsikris 2018) are some of the points being disputed.

Theoretical Frameworks

Different Theories of Consciousness propose a variety of frameworks for the study of consciousness. The Templeton World Charity Foundation's Accelerating Research on Consciousness initiative (ARC)[17] proposes a research program based on open science and adversarial collaboration to advance in the understanding of consciousness. These are some of the proposed frameworks:

Global Neuronal Workspace Theory

The Global Neuronal Workspace Theory (GNW) (Baars 1988, Dehaene et al. 1998) refers to the coexistence of distributed local processing areas operating in parallel with a global workspace network. These processors compete for access to the global workspace. Consciousness is reached when content enters the workspace and is made globally accessible to the rest of the brain (termed "ignition"), due to a whole brain activation. Once in the GW this content is the only one to be conscious about.

Dynamic Core Hypothesis

The Dynamic Core Hypothesis (DCH) (Edelman and Tononi 2000a) focuses on the existence of a core functional unit responsible for representing consciousness in the brain, whereby segregated and differentiated processes are integrated to give rise to conscious experience. Conscious experience can be measured though the estimation of neural complexity that

[14] https://arxiv.org/abs/1701.08734
[15] https://www.frontiersin.org/articles/10.3389/frobt.2019.00017/full
[16] https://psyarxiv.com/6snfm/
[17] https://www.templetonworldcharity.org/our-priorities/accelerating-research-consciousness

relates to structural connectivity, functional integration and functional differentiation of neuronal activity in thalamocortical networks and is based on mutual information between complementary subsystems of a neuronal network (Tononi & Edelman 1998).

Note that thanks to advances in brain imaging and probing, neuronal correlates of consciousness can now be studied, allowing also the identification of parts of the brain that are responsible for consciousness. The cerebral cortex has long been claimed essential for explaining consciousness. Parvisi and Damasio[18] and most recently Solms[19], position the basic form of consciousness as a core biological process of life regulation. This identifies RAS (Reticular Activating System) as an important locus for consciousness. This area connects the brain stem to the cerebral cortex through various neural paths. The approach opens a path to decouple consciousness (completely for Solms, partially for others) from its cognitive manifestation by linking it to homeostatic regulation, the processes that living systems use to fight entropy.

Integrated Information Theory

Integrated Information Theory (IIT) (Tononi et al. 2016) states that the level of consciousness depends on the amount of information the brain can integrate. This capacity depends on the system's causal and intrinsic dynamical interactions. Information integration allows conscious brains to present a larger amount of information than its components (Saffron 2020), which constitutes a fundamental property of complex systems. IIT suggests a novel quantity, called Phi, which measures the richness of the whole brain dynamics by calculating the irreducibility of the system (Anil et al. 2011). Higher degrees of Phi reflect higher levels of integration and at the same time segregation, consequently higher consciousness levels. Being Phi computationally intractable, several mathematical proxies exist, such as the Perturbational Complexity Index (PCI), which calculates the complexity of the system after perturbing it (Casali et al. 2013).

Theories based on Consciousness as Prediction

The Entropic Brain Hypothesis (EBH) (Carhart-Harris et al 2014[20]) postulates that the richness in content of any conscious state can be described by the magnitude of entropy of the associated brain activity. According to this theory, entropy refers to its information theoretic sense, and is influenced by the notion that greater entropy equals greater uncertainty and information content.

The EBH theory is related to the Predictive Coding (PC) theory, which is based on the Free Energy principle (Bucci & Grasso 2017, Clark 2013, Jakob 2013, Friston 2009, Roumen 2016). Specifically, both theories are based on information theory, and are closely linked to Shannon entropy. In particular, EBH measures the uncertainty of spontaneous neuronal fluctuations over time, while PC measures the uncertainty of beliefs or high-level priors encoded by such neuronal fluctuations (Carhart-Harris et al., 2018). Together these two

[18] Parvizi J, Damasio A. Consciousness and the brainstem. Cognition. 2001 Apr;79(1-2):135-60. doi: 10.1016/s0010-0277(00)00127-x. PMID: 11164026.
[19] Solms, M. (2021) The Hidden Spring: A Journey to the Source of Consciousness. W.W. Norton Company.
[20] https://www.frontiersin.org/articles/10.3389/fnhum.2014.00020/full

theories form the recently proposed ”relaxed beliefs under psychedelics” (REBUS) model that states that psychedelics relax high-level priors liberating bottom-up information flow that can help modify any pathological priors (Carhart-Harris and Friston, 2019).

The algorithmic information theory of consciousness or K-theory of consciousness (KT) (Ruffini 2017) focuses on the generation of structured experience in the brain arising from the ability of agents to model and track the external “world”. Structured experience — a form of consciousness that builds on primal consciousness by accounting for complex, hierarchical representations of data — arises from the comparison of reality with its internally existing model. The mathematical foundation for this theory is algorithmic information theory (AIT), the result of combining Shannon information theory with the theory of computation of Turing. The theory connects readily with Predictive Coding (Clark 2013, Friston 2009, Jakob 2013) and IIT, but seeks to provide a framework for understanding graded, structured human experience.

Fundamental limitations and alternatives

Several voices claim for the impossibility of implementing both Artificial General Intelligence (see Fjelland 2020[21]) and Machine Consciousness (Metzinger 2003[22]). This may lay on the type of available computer principles and machines (Koch & Tononi 2017). However, it is also argued that classical physics is not enough to explain the hard problem of consciousness and that quantum physics is at the heart of it.[23] Others claim that fundamental revisions of physics are necessary to explain consciousness (e.g., Hoffman[24] proposing a concept of agency within a revised space-time construct), making it the hard problem ‘less hard than it seems’.

Proposals along these critical or alternative lines of work are not excluded, provided they do manage to formulate significant contributions to the objectives of the Challenge as described in the call text.

Some possible directions

Let it be clear from the above that there are numerous possible frameworks and approaches for the study of consciousness. The present Pathfinder Challenge is open to any of these, or others that build bridges among them or propose other, alternative routes. Different fields in awareness research are starting to deliver partial results and there is a high potential for cross fertilisation among these. The call explicitly seeks to diversify from the human frame of reference, towards approaches that can be applicable to other kinds of systems and entities and other kinds of awareness than those that the human is naturally grasping.

[21] Fjelland, R. Why general artificial intelligence will not be realized. *Humanit Soc Sci Commun* **7,** 10 (2020). https://doi.org/10.1057/s41599-020-0494-4

[22] Th. Metzinger, Being No One, MIT Press, 2003.
[23] Penrose, Hammeroff, Rovelli, and many others..
[24] http://cogsci.uci.edu/~ddhoff/ConsciousRealism2.pdf

The starting point of the Challenge can be captured by the question 'what is it that a user would expect from a service or device that has 'awareness inside'. Here 'user' refers to any person interacting with a 'machine' of any kind (physically embodied or not, anthropomorphic or not, biological or not). What, in other words, is the extra that consciousness brings to a system? Following from that, proposals are expected to develop a technological approach for achieving this extra, in a measurable and verifiable way, and to argue the benefit this would bring and – ultimately – the innovation potential it would unlcock (impact).

This Pathfinder Challenge looks for novelty in any of its main objectives, related to scientific and theoretical underpinnings, to demonstration and application, and to the technical approach being taken. Below are some possible directions, by no means prescriptive or exhaustive, in which such novelty might be found.

Consciousness and Broad Intentions

Recent work (rooted in Ethics of AI) has demonstrated that people expect different things from other people than from machines doing similar things in the same context.[25] First, people judge humans by their intentions and machines by their out-comes. The second principle is that people assign extreme intentions to humans and narrow intentions to machines. People tend to accept mistakes from machines when these can be explained by such narrow intentions, but not when physical harm is caused.

Broad intentions are beyond the state of the art in Artificial Intelligence, yet they seem needed for trustworthiness. Consciousness can be seen as a form of 'broad intention': it modulates behavior and problem solving to comply with a system of boundaries and constraints that determine what are viable and acceptable solutions to the problem, or acceptable, fair or just behaviors (excluding, for instance, physical harm, in the example used above).

Consciousness and self-regulation

The plausible link between consciousness and homeostatic processes (Parvizi, Damasio, Seth, Solms, above) has to our knowledge, not been widely explored for artificial consciousness. Homeostasis captures a lot of processes that we are not aware of (breathing, heart), but also our behavioral reactions to sensations and feelings which evolution has wired into patterns for survival and well-being. The link to the Cortex, for those species having it, allows making these cognitively explicit, be it always approximate since this reification has (at least to some accounts) no causal role.

Beyond demonstrators: theoretical framework

[25] Hidalgo, C.A. (2021) How Humans Judge Machines. MIT Press. Also at https://www.judgingmachines.com

The call text provides some examples of devices and services that could benefit from certain features of consciousness: ‘safer robots or self-driving cars, for better resilience of critical infrastructure, in artefacts that compensate for consciousness disorders, in decision support (e.g., for surgery, economics or epidemiology), or for chatbot-based conversation, language learning or translation’. However, projects are encouraged to go beyond such a specific and single demonstration by selecting, proposing or adapting an underlying theoretical framework, and by developing a credible onset of a ‘toolbox’ (a series of concepts, methods, techniques and technologies) for realising certain aspects or forms of awareness and consciousness with a clear (and preferably demonstrated) potential to be useful for a wider range of applications, and preferably applicable to very different kinds of systems, in multiple, open or evolving contexts, covering diverse awareness features and for different purposes (robustness, development, etc.). Obviously, the technological approach is expected to reflect key features of the theoretical framework.

It is a priori not excluded that the general problem of consciousness does not respond to any unifying theoretical treatment. As Daniel Dennett argues, consciousness may well be a bag of tricks of our mind, nothing more. If such is the starting point of a project one would nevertheless expect a plausible set of such tricks to be reflected in the toolkit, so as to account for a reasonable range of features of consciousness with broad applicability.

Doughnut Intelligence: generalising self-regulation

The concept of ‘Doughnut Intelligence’ is an onset for a potential approach to an extended form of consciousness, applicable to a wide range of systems.[26] It is directly inspired by the idea of ‘Doughnut economics’[27] The Doughnut of social and planetary boundaries visually captures the viable space for sustaining civilisation on our planet, squeezed between undershooting (the boundary of a sufficient social foundation) and overshooting (the ecological ceiling) the use of resources. If successful, the planet has a socio-ecological homeostatic state that is viable, both socially and ecologically. It is widely used in relation to the UNs Sustainable Development Goals (SDGs).

Doughnut intelligence applies this idea to any form of awareness of such viability, i.e., from the individual level, to the planetary level, as in Raworth. It seems plausible that several levels in between (household, city, energy or transport infrastructures, eco-systems, etc.) can be conceptualised in similar ways, requiring or enabling different forms of awareness (e.g., awareness of environmental footprint of some specific action). Obviously, however, as the space-time scales of these levels extend, the human condition is less and less adapted for the corresponding awareness (think of the difficulty to stimulate behavioral change for ecological objectives). It is, nevertheless, a possible path for exploring new forms of

[26] Webinar on ‘Responsible AI’ https://youtu.be/PbzoXm0qsEc

[27] Raworth, K (2018) Doughnut Economics, Seven Ways to Think Like a 21st-Century Economist. Random House UK Ltd.

awareness and consciousness that can be used to more effectively keep socio-ecological systems into their just and viable working range.

Artificial Intelligence

The call is also an opportunity to renew perspectives in Artificial Intelligence. While the field is currently booming with applications, the technological basis on which these are built is actually quite narrow. In a recent critical paper[28] the author reviews the history of AI from its ups-and-downs through AI-springs and AI-winters and anticipates a new AI winter. While she suggests that 'common sense' is an essential missing ingredient for general AI, the present call invites for an alternative answer around a new approach to consciousness (although a deep connection between both, consciousness and common sense, is plausible).

Ethics

The ethical dimension is another critical aspect of the call. All projects are expected to take it into account and further develop it for their specific approach. This should build on the most recent insights, expert advice and proposals for a regulation on AI that addresses the risks of AI.[29] This may need further specific considerations when targeting forms of awareness and consciousness. The most recent AI regulation poses a definition of AI that may be too limited (being rooted in the currently dominating data-driven paradigm), and the analysis of risk categories may need to be revisited.

In light of the discussion above, one can anticipate possibly meaningful extensions of the ethical debate. A first one would be to take into account the notion of 'just and viable' operating space as an ethical requirement. This could be linked, for instance, to the European 'do-no-significant-harm principle' as a fundamental ethics building block. Another direction could take into account the ethics of extended awareness from the way it may influence one's own perception as a coherent unity ('self') in the world. This perception of coherence is obviously challenged by awareness at larger scales, beyond immediate perception, and by a non-localised awareness of impacts of one's own actions, ultimately in emotional (as driver in a generalised homeostatic system) terms.

Projects funded under this call will be expected to collaborate around the ethical issues, possibly identifying further interpretation of existing regulations, or for identifying gaps and needs for further regulation regarding conscious machines.

Impact: relevant problems in a relevant way

The impacts sought by this Challenge stem from enabling features of consciousness in a range of systems and applications, leading to a clear (possibly quantifiable) advantage in

[28] Mitchell, M. (2021) Why AI is Harder Than We Think. arXiv:2104.12871v2 **[cs.AI]**
[29] Proposal for a Regulation laying down harmonised rules on artificial intelligence | Shaping Europe's digital future (europa.eu)

terms of performance, robustness, user experience, or any other combination of relevant indicators. As described above we expect this to go beyond a single application. Impacts can be expected to be larger when progress is rooted into a solid theoretical and conceptual framework for consciousness, with implications beyond human consciousness, and that can be translated into a range of technological ideas for achieving consciousness (or aspects there-of) into a broad (even open-ended) range of systems and contexts of use. Such wide applicability will ultimately unlock innovation potential in specific applications. In line with the Pathfinder orientations, concrete exploitations are not expected to occur in the course of the funded projects. Nevertheless, outreach activities to potential users and interested parties, and to society more generally, will be important to probe the societal viability of the proposed approach.

3. Proactive portfolio management

This section describes the EIC proactive management as applied to this Pathfinder Challenge. It starts by building the portfolios; i.e. by allocating actions into portfolios (a). Proactive management will allow to define and to update portfolio's objective and roadmap (b). Portfolio members will benefit from portfolio activities and from the access to the EIC Market Place (c).

a. Allocation of the actions into the portfolio

This section provides the Challenge specific elements of the way in which the evaluation results in a coherent Challenge Portfolio. It should be read in conjunction with the overall evaluation process as described in the EIC Work Programme text (Appendix 3).This section provides guidance to proposers on how to align their proposal with the architecture of the Challenge Portfolio as envisaged by the Programme Manager.

At the second evaluation step, the evaluation committee, chaired by the Programme Manager, builds a consistent Challenge portfolio. In order to do so, the evaluation committee will allocate proposals into categories.

For this Challenge, the evaluation committee will categorize proposals according to

1. the type of entity to be made aware. These types will be created at the time of evaluation. They are likely to include 'robotic system', 'human', 'device', 'service', 'distributed system' or others that the committee finds useful. Proposals can fall under different types, either by what they aim to concretely demonstrate, or by how the approach is arguably applicable to a range of entities.
2. The awareness features (roughly stated: what is a system conscious about in a way that it influences its framework for action). These types will be created at the time of evaluation. They are likely to refer to directly or indirectly perceivable features (internal or external to the conscious entity), at different scales of space and time, or of social extent.

The evaluation committee, guided by the Programme Manager, will use the following considerations to constitute the Challenge portfolio, i.e., for selecting excellent proposed actions from the different categories:

1. The portfolio aims to have a high diversity of type of entities to be made aware.
2. The portfolio aims to have a high diversity of awareness features to be addressed.

We anticipate that this will lead to diversity of theoretical foundations with inherently broad scope of applicability. In case of indecision, the committee will further aim to maximise the interdisciplinarity dimension of the portfolio.

The selected projects, once funded, will be included in the EIC Challenge Portfolio.

The contractual obligations subsequent to the participation of a project into a Challenge portfolio are described in Appendix 1.

b. Portfolio objectives and roadmap

Portfolio Objectives are overarching objectives for the collection of projects in the portfolio, to be achieved by joint activities among the projects' participants. They are set by the Programme Manager, in close discussion with the participants. They will be updated following discussion with beneficiaries, and revised on a regular basis, for instance based on projects' achievements, new technology trends, external inputs (other projects, new calls…), and discussions with stakeholders/communities.

For this Challenge, the Portfolio Objectives and Roadmap will be co-designed and agreed between the Programme Manager and the participants of the projects, once the Portfolio is established. This will build on the objectives set out in the relevant call text (see the WP extract in Section 5), this Guide (in particular its Section 2: Overall objective of the Pathfinder Challenge – e.g., joint work on ethics), and other relevant actions, other interested members of the EIC Community and other third parties.

c. Portfolio activities and EIC Market Place

Portfolio activities will be proposed and designed by the EIC Programme Manager and in consultation with the beneficiaries of funded projects. They aim at developing the cooperation within the EIC Portfolio in order to:

- achieve the Portfolio Objectives or the objectives of the actions,
- enhance research,
- prepare transition to innovation,
- stimulate business opportunities,
- and strengthen the EIC Community.

Such activities may cover - notably but not only – organization of, and participation to conferences, workshops or any EIC Portfolio or networks meetings, experience and data sharing, as well as participation in any relevant EIC Business Acceleration Services events.

The responsible Programme Manager will manage the portfolio through actions, he or she sees most fitting to advise the participants in the orientations of their work, or exploring potential synergies with others, including with businesses and start-ups, for mutual benefit.

To enhance cross-fertilization activities, and to stimulate potential innovation, the EIC Programme Manager may request any beneficiary to make available - through the EIC Market Place - information on preliminary findings and results generated by the action, with the aim to probe their potential for further innovation. The Programme Manager aims to accelerate the most promising results using EIC tools such as, fast tracking a project to Accelerator (see Appendix 1, Active management section) and attributing additional funds (up to 50K that can be used more than once, see Appendix 4) where rather unexpected outcomes open new opportunities.

At this point, the following Portfolio Activities are being scheduled:

Date	**Description**	**Main outcomes**	**Reference/report**
tbd	Portfolio kick-off meeting	Challenge Objectives and Roadmap	

Table 1: Portfolio Activities for Challenge Awareness Inside

4. Challenge Portfolio

(this will be completed after the Call evaluation)

5. Challenge call text

(extract from https://eic.ec.europa.eu/system/files/2021-03/EIC%20Work%20Programme%202021.pdf)

Awareness and consciousness have been high on the Artificial Intelligence (AI) research agenda for decades. Progress has been difficult because it has been hard to agree on exactly what it means to be aware. Most researches would agree though that we do not have any truly aware artificial system yet, that awareness is much more than a sensorial sophistication, and that it is much more than any Artificial Intelligence as we know it. But, what is it then that a user would expect from a service or device that has 'awareness inside'?

Most scientific and philosophical accounts of awareness are based on a human subject perspective and at an individual level. They address the question of what it means for an individual human subject to be aware of, e.g., the environment, time or oneself and how one can assess awareness in this context. The problem is relevant, certainly, since many clinical and cognitive conditions can be linked to awareness issues. The concept is also relevant to emerging technologies as it has been argued, for instance, that humans will not accept robots (or chatbots, or decision support systems) as trustable partners if they cannot ascribe some form of awareness and true understanding to them.

The individual human-centric concept of consciousness hinders the application of awareness as a measurable feature of any sufficiently complex system. The study of awareness in other species and artefacts, or even more elusive concepts such as social awareness require a new perspective applicable to many systems. It can then also serve to attack the inter-subjective state and experience of awareness (i.e., what is it like to interact with an aware robot that, most probably, does not have the same kind of awareness than the human?), or to include non-conscious objects into the sphere of awareness (e.g., to become aware of the time without looking at the watch).

For technologies, awareness principles would allow a step-up in engineering complex systems, making them more resilient, self-developing and human-centric. Awareness is a prerequisite for a real and contextualised understanding of a problem or situation and to adapt ones actions (and their consequences) to the specific circumstances. Ultimately, awareness serves the coherent and purposeful behaviour, learning, adaptation and self-development of intelligent systems over longer periods of time.

Specific conditions for this challenge

Proposals are expected to address each of the following three expected outcomes

- New concepts of awareness that are applicable to systems other than human, including technological ones, with implications of how it can be recognised or measured. It will require to elucidate the relationship between, among others, complexity and awareness, information structure and representation, the environment and its perception, distributed versus centralized awareness, and time awareness. This will lead to better approaches for defining aspects of awareness over different temporal, spatial, biological, technological and social scales.

- Demonstrate and validate the role and added-value of such an awareness in an aware technology, class of artefacts or services for which the awareness features lead to a truly different quality in terms of, e.g., performance, flexibility, reliability or user-experience. The specific expected outcome is a proof of principle of technologies far beyond the current state of the art or a laboratory-validated prototype enabling evaluation of the proposed technology's awareness features, relying where relevant on neuroscientific and psychological methods, and possibly in a range of application areas. As examples, projects could investigate the implications of 'awareness inside' for safer robots or self-driving cars, for better resilience of critical infrastructure, in artefacts that compensate for consciousness disorders, in decision support (e.g., for surgery, economics or epidemiology), or for chatbot-based conversation, language learning or translation.

- Define an integrative approach for awareness engineering, its technological toolbox, the needs and implications and its limits, including ethical and regulatory requirements. On this aspect specifically, the projects that will be funded under this challenge are expected to collaborate and contribute to the wider ethical, societal and regulatory debate since, ultimately, new awareness concepts may lead to a redefinition of how we look at the relation between humans, other species and smart technologies. The gender dimension in research content should be taken into account, where relevant, to maximise user experience.[30]

This Challenge is only open to proposals for collaborative projects with at least 3 partners following the standard eligibility conditions. Proposals are required to comply with the Trustworthy Artificial Intelligence principles (see Annex 2).

[30] https://ec.europa.eu/info/news/gendered-innovations-2-2020-nov-24_en.