Brain-brain integration in 2035: metaphysical and ethical implications

Soraj Hongladarom
Department of Philosophy, Chulalongkorn University, Bangkok, Thailand

Received 7 October 2014
Revised 14 January 2015
Accepted 3 March 2015

Abstract

Purpose – The purpose of this study is to think ahead into the year 2035 and reflect on the ethical implications of brain-to-brain linking.

Design/methodology/approach – Philosophical argument.

Findings – It is quite likely that the direction of technological research today is heading toward a closer integration of mind and machine in 2035. What is interesting is that the integration also makes mind-mind or brain-brain integration possible too. There is nothing in principle that would prevent hooking up more than one brain to a machine, or connecting two or more brains together to harness their processing power to tackle a very complicated task. If that happens, the whole notion of what it is to be an individual and a self will have to be rethought. I have offered a way in which that can be done: Instead of viewing the self as being contained in a closed space traditionally defined by the skin, the self can expand outside of the skin and merge temporarily with other selves too. This also has profound implications on the notion of privacy, especially on how it is conceptualized and justified.

Research limitations/implications – This research is limited to theoretical argumentation only. It relies on the current empirical and scientific investigations that are going on at the moment and provide ethical reflections on them.

Practical implications – We need to anticipate technological innovations to be more proactive in deliberating and formulating policy and ethical guidelines; otherwise, ethicists will just muse after the fact, implying that there is nothing further to be done.

Social implications – Brain-to-brain linking has tremendous social implications, so is the ethical reflection on the issue.

Originality/value – Argument purporting to show the specific content in ethical guidelines on brain-to-brain interlinking based on the metaphysics of the self that is directly implicated by the technology has not been done before, according to the author's best knowledge.

Keywords Privacy, Computer ethics, Robotics, Bioethics, Human values

Paper type Research paper

Introduction

Talking about the future can be useful in letting us have an idea of what the world will look like and prepare for it. What I would like to do in this paper is to venture a

Research for this paper has been supported partially by a grant from the World Class University Project, Chulalongkorn University, grant no. WCU-H3-064-57. The first draft of the paper was presented at the ETHICOMP conference in Paris, in June 2014. Author would like to thank Charles Ess and Bernd Carsten Stahl for their comments and support. Special gratitude is also extended to the anonymous reviewers of this paper and to Profs Rao and Stocco for granting their permission to use the diagram of their research.

Journal of Information, Communication and Ethics in Society
Vol. 13 No. 3/4, 2015
pp. 205-217

1477-996X
DOI 10.1108/JICES-10-2014-0042

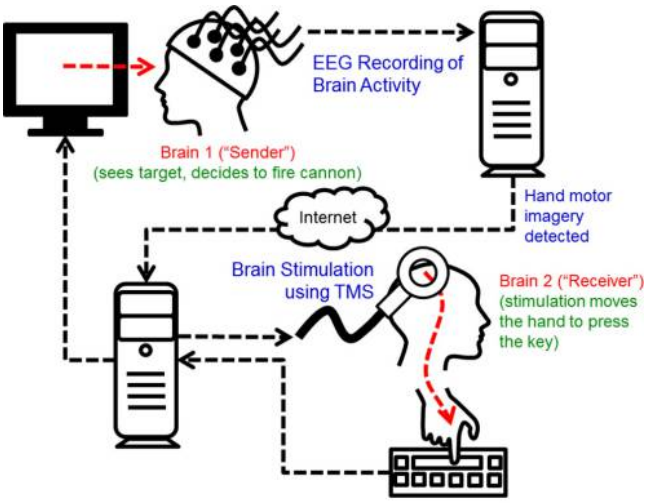
prediction: What will actually happen in the year 2035 in the field of brain-to-brain interface? What kind of technology and applications will be commonplace in the next 20 years? What are its metaphysical and ethical implications?[1] The specific issue the paper will be dealing with is brain-to-brain interface. One brain is hooked up to a machine and then the machine is hooked up with another brain, putting the two brains (or more) on the same network. It is thus not only a potentially very powerful technology but also one that we should pay serious attention to, as it has the potential of causing very serious ethical violations. A team at the University of Washington at Seattle in the USA, led by Rajesh Rao, has recently announced that they have successfully experimented with brain-to-brain interface where one colleague controls another's motion through the linking of their brains alone (Rao *et al.*, 2014; see also Direct Brain-to-Brain Communication in Humans, 2013, and news report in Armstrong and Ma, 2013). My prediction is that 20 years from now, the technology will become more powerful; instead of brain-to-brain interface, it will become brain-brain integration as the level of working together of the two or more brains becomes deeper.

There are a number of vexing metaphysical and ethical issues involved in the technology, perhaps the most important of which has to do with our understanding of the self. If it is possible that two brains can be integrated into one single unit, then are we talking about there still being two selves? Is it actually possible for two selves (inhabiting the two brains) could somehow merge into one? Brains can clearly be counted, but how do we count minds? Moreover, the ethical concerns are many, and I discuss only the most important ones in this paper[2]. Perhaps the most important ethical concern arising from this kind of technology is the possibility of mind control. It is one thing to control a colleague's arm whose brain is attached to ours, but it is totally another to be able to manipulate the content of her brain so as to serve our purposes. The possibility of mind control is strongly connected with the concern over privacy; only in this case, it is the privacy of the content of one's own memories and thoughts, what is there as content of one's own mind or brain, that is at issue. It is not the same issue as how to protect online data on one's social networking profile any more. Now, protection of personal data is intimately tied up with the sense of who we actually are. Manipulating the content of someone's brain has the potential of even altering who she really is. And that is certainly a very serious scenario.

Brain-to-brain interface: the basics

Brain-to-brain interface is a direct offshoot of brain-computer interaction. The basic idea is simple enough. The brain is in fact a giant network of neurons, each firing constantly at one another in all kinds of patterns. Linking up the brain to a computer thus is a matter of allowing the neurons to communicate with some components inside a computer, thus linking the brain and the computer together. Signals from the brain can then be interpreted by the computer. Many laboratories around the world are in fact developing applications for the interface, such as the ability to transmit images from within the brain onto a publicly viewable monitor, controlling external devices with thoughts alone (Figure 1).

In the experiment by Rao and his team, signal from Brain 1 ("Sender") is transmitted to a computer, which detects and interprets the signal as a command to move a finger. The signal is then transmitted via the Internet to a receiving computer, which then sends it to the receiving brain, which interprets the signal as a command to move a finger. The



Source: Rao *et al*. (2014). Permission given by Rajesh Rao and Andrea Stocco, University of Washington

Figure 1. Schematic representation of the University of Washington experiment

receiving brain then sends the signal to the finger, which moves accordingly. In the experiment, the Sender is playing a video game, which requires the player to push a key to fire at an enemy ship. The Sender watches the game on the monitor, but he does not push the button himself; instead, he transmits the "fire" or "not fire" signal to the Receiver who sits in another corner of the room with his back turned toward the monitor. The Receiver's index finger is on the fire key. Whenever his brain receives a signal from the Sender, his index finger pushes the button; whenever the enemy ship passes by. The team reports a satisfactory rate of hit using this method, which shows that direct brain-to-brain communication is at least a success at the beginning.

A question that emerges from studying this experiment is what happens to the self of the Receiver. In a Buddhist text, one of the criteria for determining which part of a body belongs to a self is whether the self can control that part or not. Normally, we can control our own bodily parts, such as our fingers, legs and so on. Therefore, if the Sender can control how a finger belonging to the Receiver moves, then there is a sense in which the finger belongs to the self of the Sender! This is intuitive from our basic understanding that if we can control a body part at will, then that part in a sense belongs to us; in other words, that movable part is part of our selves. Therefore, on the one hand, the finger clearly belongs to the Receiver because it is still attached to his hand and his body, but on the other hand, it quite clearly belongs to the Sender because he can move it at will. I will explore this intriguing metaphysical dilemma in the next section.

Metaphysics of the self

The least that we can infer from the situation in the brain-to-brain experiment above is that the conception of the self–what we do take to be our own selves–is problematized. This is not a new phenomenon by any means. Apart from being the subject of perennial interest by philosophers, the self has become a focus of much attention by

neuroscientists (Kircher and David, 2003; Damasio, 2003, 2012; LeDoux, 2003). I have also tried to show elsewhere that the self is constituted through the information network that makes up its awareness moments, perceptions, thoughts, feelings, beliefs and so on (Hongladarom, 2011a, 2011b, 2013). In addition, each of these is different from the other, it is impossible to find an enduring entity that objectively ties up these strands together so as to represent a single self. Even the narrative account that many have cited as a rather promising account of self-identity does not appear to work well because a narrative can branch this way, and there is no single strand at which we can point and say that it, rather than the other branches, is what makes up the (real) self of this particular person[3]. It seems that one has to give up the idea that there is a single, enduring self that functions as the agent of the plot in a narrative. Rather, it seems to me that the self or the experiencer of the story also changes when the story unfolds. This is quite common in any kind of story – the main character changes as a result of the changing events in the story. However, in order for us to be able to pinpoint accurately what objectively makes up that particular person in the story and no one else, we have fixed a set of defining characteristics of the person; otherwise, we would not have any objective criteria at all. But it is unlikely that there can be such enduring characteristics of a person that cannot change, and if this is so, then any criterion we can use to fix an identity of a person has to come from outside of the person herself, such as our own account that we make of that person, something from within our minds that we use to fix that person[4].

It is quite well accepted by now that what we know as the self is in fact a construction out of a number of mental and physical episodes that causes a subjective viewpoint from one's perspective to occur, and this viewpoint is then understood to be "the one" who is experiencing the world. In other words, consciousness of the self arises when there is recognition of a unique vantage point from which the world is viewed. This vantage point, dubbed the "center of narrative gravity" by Dennett (2013), is then linked up with other such points during a period of time giving rise to a feeling that there is something continuous in time that functions as the locating point of all perspectives of the outside world. Furthermore, this vantage point also takes up a specific point in space as that place from which the viewing of the world from a first-person perspective occurs. All these give rise to the sense of a self, or of "oneself.". Findings in neuroscience indicate that the regions of the brain which are responsible for consciousness of the self are spread out, and there is no one specific region that is directly responsible for it, no "homunculus" sitting inside doing the work. Dennett's center of narrative gravity is actually a fiction, just like the prime meridian is a fiction. Nonetheless, as the prime meridian has proved very useful in anchoring the system of latitudes and longitudes, the center of narrative gravity is also very useful for each of us to construct our own personal narrative. Nevertheless, the center of narrative gravity can certainly shift as the narrative itself shifts. If the story of my life branches out in that way rather than this one, then my center of gravity will have moved. It will not be at the exact same location as my center is (metaphorically) located right now. Suppose, when I was young, I did not choose to study philosophy, but went to medicine instead. Then the center of my narrative gravity would be in a much different character than it is now. But, even if I had entered the medical school, my putative center of narrative gravity will still be there to perform a necessary logical function of tying up all the episodes of my life together into a single whole (similar to Kant's idea on the transcendental unity of apperception).

The current development in mind-machine integration and brain-to-brain interface, then, creates a very challenging conundrum on how we should understand the self. If the self is not to be equated with the homunculus and is neither constituted by any specific region of the brain nor by any closed network within the brain itself, and if the brain is essentially a collection of vast networks of neurons firing at one another, and if there are some connections between these neurons within the brain and their counterparts coming from outside in forms of technological artifacts, then those external counterparts should be considered as parts of the brain too, and, consequently, as parts of the sense of the self. In other words, the self could expand to include those added parts too (Clark and Chalmers, 1998). Imagine a tight integration that occurs when a chip is implanted in a brain to augment its capabilities; imagine further that we are the ones whose brains are being augmented in this way. Now we benefit from the added memory capacity or the more efficient calculation power. It is then easy to imagine further that these added capabilities are so seamlessly integrated with our brains that we do not notice any "bumps," so to speak. The added capabilities just become a natural part of who we are. From this point on, if we imagine further that another brain is hooked up to the same network, then our own sense of self extends toward that brain for the same reason, and the self that belongs to the newly hooked up brain extends to our own brain too. In other words, we have a real merging of two minds, where two minds literally become one.

But if this is the case, then when the Sender in Rao's experiment sends a signal to the Receiver intending to move the latter's index finger, and there is thus a sense in which the index finger belongs to the Sender. The finger, lying outside of the Sender's body and attached to the Receiver's, thus becomes part of the self of the Sender. It is as if during the experiment, the Receiver relinquishes this part of his own self to the Sender, lending his finger, so to speak, to the latter so that he can accomplish his task of shooting the enemy ships. We can certainly imagine further that if the same set of hookup is wired so that the Receiver can also move the Sender's finger, then the selves of the two persons will be merged. For A to be able to move B's finger, B's finger becomes a part of A's self, and if B can do the same to A's finger, then A's finger becomes a part of B's self too. The selves of A and B will thus be intertwined.

Still this does not automatically imply that the self can exist outside of the body. There is a strong tradition that maintains that the self is closely tied to the body (Merleau-Ponty, 1962), so according to this tradition, the idea of the self outside of the body borders on the murky (reminiscent of the immaterial soul hovering above the body) or the ridiculous. But a key point here is how we determine which bodily part belongs to someone's body. The Buddhist text mentioned earlier gives us an idea that whatever bodily part I can move at will becomes part of my body, and that is intuitive enough. I believe that the direction of technological development that I am talking about here points to a conclusion along this line. The circuitry that is added to the brain is a physical addition to the already existing material manifestation of the conscious self. Hence, if I can move your finger, then your finger at that moment becomes part of my body, and vice versa.

Furthermore, when two brains are connected together in this way, what could happen is that the first-person perspective of the two persons will merge together. Suppose, A's brain and B's brain are wired together, A will then have access to the first-person perspective of B, and vice versa. In this case, then, a complex philosophical problem ensues. What would it mean for there to be two selves, in this case, A's and B's?

Or would the merging result in an emergence of a new self, a synthesis of A and B? We tend to think of being a self as having a monologue within the mind. Imagine yourself to have a monologue inside your mind and the person who speaks the monologue is your own self. This very familiar picture makes it difficult to imagine a situation where the two brains are wired together and the two selves are merged. Would it mean that there are two voices speaking to each other, or that each is engaged in their own monologue? But the person hearing the voices will also be the one who says the voices, or at least one of them, so does it mean here that the person can enter any of the voices at will? And when the two voices are merged together completely so that there remains only one voice, whose voice is it going to be? A's, or B's? Or the newly synthesized self?

The merging does not always have to produce nightmarish scenarios. We can also imagine cases where the two minds collaborate closely and form a close-knit team to accomplish a complex task. In these cases, the integration of the two brains would occur not only at the level of direct mind-to-mind conversation, but also at the level wherein two minds merge completely as one which is entirely focused and concentrated on the task at hand. In this case, the two minds do not "talk" with each other, but become one mind which is intensely focused on the task.

The possibility of merging two brains, as described above, is admittedly remote, even 20 years from now. But still there are certain indications that the development in technology could be going into this direction. In any case, we can imagine this type of scenario as a thought experiment. What the thought experiment seems to show is that the conception of the individual self is not only constructed, but it also can extend outside of the physical body, and merge and engage in interplays with other brains/selves in interesting ways. When the two minds are so concentrated on some common task, the consciousness of there being a "self" that is opposed to and distinct from the "object" being thought of disappears and the two minds become one. The self can be extended outward and merge with other selves so that a new, expanded self-results. The first-person singular pronoun then gives way to the first-person plural pronoun, only that the "plural" in the "we" actually becomes a singular[5].

Ethical implications

Brain-to-brain technology naturally gives rise to a number of serious ethical issues. One is about who exactly will be responsible for an action performed by a multiple brain unit. As the unit is composed of two (or more) brains connected together and performing as one unit, should the two (or more) persons whose brains participate in the team each be held responsible? So, should the responsibility be shared? At any rate, it is clear that our standard notion of agency and responsibility need to be revised in light of the possibility of brain-to-brain teams. Could one of the members of the team say that he or she is not responsible for the action of the team because his individuality was suspended during its operation? That, however, would be similar to a member of a football team saying that he is not responsible for the action of the whole team, such as winning a match, because his individuality was suspended during the game. Actually, it was precisely because team members suspend their differences and individualities and concentrate every bit of their mental and physical effort into playing that is responsible for the team's winning performance. Common sense seems to tell us that it is all the members who participate in the team who should be responsible, but then it is not each member taken separately, because each has specific roles to play in the team. Although there are

certainly bound to be glitches in the communication – there might be failure of communication among the team members (which happens all too often), or there might be loss of fidelity in the communication between the connected brains, these are errors that can be corrected, and if the technology is to move forward, a way needs to be found to correct these glitches. Nonetheless, it is common sense that in a team performance, it is the whole team that is responsible (if the glitches are not too serious to undermine the team's performance).

The issue of responsibility becomes more serious and difficult when considered in the context of brain-to-brain interface because the linking is much more tightly integrated than in a typical team performance. In the former, it is possible that a new form of consciousness could emerge that results from merging the two brains together. The emerging consciousness is there only when the brains are connected and disappears as soon as they log out, but it is this emergence consciousness that we think is responsible for whatever task it accomplishes. Practical considerations, however, say that somebody or something has to be responsible. Suppose the mind-merged team has committed something wrong, it is clear then that someone has to be held accountable. But as we have seen, it is not immediately clear where to find the responsible party. The most likely one should be the team itself during its moment of full operation, i.e. with all of its members losing their individuality and completely merging into the collective whole.

If we imagine that the mind-merged unit has its own independent decision, that is free from the thoughts of each of the two persons inside, then it is not an easy matter to decide who is in fact responsible. When the two minds or brains are tightly integrated in this way, there is a chance that something entirely new will emerge, a new consciousness or a new identity perhaps[6]. The emergent consciousness does not have to be identical to any of the persons who are connected to the network. A way to see how this is possible is to consider that the emerging consciousness is a result of two brains joined together, whereas the consciousness of the persons who join comes from only one brain, their own. We can imagine cases of split-brain patients who experience a split in their personalities. This seems to show that personalities are a function of the brain or its parts, thus presumably, if we join two brains together we could have a new, emergent personality and consciousness, which does not have to be the same as any of its components. This single, emergent unit of personality then can be held to be responsible for their action, whereas the brains or the minds of each of the persons composing the emergent unit know almost nothing about what is going on. In this case, common sense indicates that it is the emergent consciousness that should be responsible. However, this does not imply that the persons whose brains are joined in this way can evade their responsibility altogether. In most cases, if the technology is available, then the merging will likely be done only for a short period of time to accomplish a complex task, and once the task is finished then there is no longer reason for the two brains to remain conjoined. In this case, the emergent consciousness does emerge for only a short time. Hence, if during the merging the emergent consciousness performs an action, then after the brains are separated, it does not have to mean that no one is responsible for the action. Even though during the merging each of the owners of the two brains is not directly responsible, they still are materially responsible because those are still their brains that are involved. In any case, this issue will occupy philosophers and scholars in related fields for some time to come.

Merging of minds in this way is a very powerful way to enhance human capabilities. Two brains or more are always better than one brain, and certainly, seem to be better than many brains (or persons) traditionally working together as a team, as the brains here are tightly merged together to become essentially one unit. Another, related, ethical concern arises from the fact that the individual person (or individual brain) is submerged to the collective brains working together. The interest of the individual is then set aside to make room for the interest of the collective. However, the individual brain itself consists of several regions, each specializing in different tasks, and these regions each consists of billions of neurons forming countless patterns and connections. There is no single "seat of consciousness" that commands all the other regions of the brain as the homunculus. All these imply that even an individual person herself can be regarded as a collective in this way. Moreover, it is also well known that humans cannot live alone without any help or connection to other people. They have to give away something; for example, they have to follow the rules of the community to become a part of the latter and thus receive their support. Thus, when brains do merge together, this does not have always to be a nightmarish scenario, such as total loss of individuality, uniqueness or individual memories. When the task that requires merging is complete, then team members return to their own individual selves.

Another ethical concern has to do with the possibility of one brain controlling the other during the merge. This is what the University of Washington experiment is about from the beginning, and the possibility could well become the most serious ethical violation as it involves direct control of one human being by another. In the experiment, the Sender takes control of the Receiver's finger for a brief moment. But in 20 years' time, A's brain could take much fuller control of B's brain during their merging session so that A himself can move any of B's limbs without B being able to do the same in return. This is clearly a cause for concern, as it shows that the relation between A and B is obviously unequal. Apart from moving B's limb, a more insidious action would be for one to implant false memories on the other or to steal memories from one brain to the other. These acts could well be a part of reality in the future as the technology of linking brains to machines is more advanced. Furthermore, political authorities might find the technology to their advantage. They might install a virus in the machine so that whenever a brain is hooked up with the machine, information inside the brain will be secretly uploaded to the government server. This would be far more powerful than what the National Security Agency (NSA) has been doing now in 2014 with their current surveillance technologies. The government might try to manipulate the content of their citizens' brains so that the latter become more docile and would not ask too many difficult questions. This would be the 1984 scenario intensified to the much higher order of magnitude. What the government can do might be again to install a virus in the machine, and when a brain of a citizen is hooked up, the virus will gain access to the content of the brain and manipulate in such a way that the owner of the brain ends up viewing the government is a much more positive light. Furthermore, when the virus gains access to the content, the citizen will then hold no secret against the government any longer, any items in the citizen's mental inventory that goes against the government's liking can then be ground for prosecution.

Having access to the information inside someone else's brain means that feats such as mind reading or retrieval of someone else's memory, can become a reality (Erler, 2011; Liao and Sandberg, 2008). This will have very profound implications on the notion of

privacy – how it is conceptualized, justified and defended (Räikkä, 2010). The notion of privacy is traditionally tied up with that of the autonomous, singular individual person. But as I have written elsewhere, another conception of the individual person (i.e. one that does not rely on the metaphysical conception of the autonomous individual typical of the liberal tradition in the West, such as the Buddhist one that I allude to in my papers; Hongladarom, 2007, 2008) gives rise to another conception of how privacy is justified, one which is tied up with the more pragmatic notion of what it is to be a free citizen in a democratic society. In a nutshell, the idea is that a free and democratic society requires that its citizens be respected and accorded a list of rights which ensure that there is a balance between the need for the authorities to maintain peace and order, and the ability of the citizens to monitor the political authorities and to influence the direction of public policies. In order for the citizens to be able to achieve, these they need to be protected from unjustified intrusion into their own private spheres; in other words, their right to privacy has to be respected. This is still valid even if, metaphysically speaking, the core individual self does not objectively exist. Privacy is justified more through its role in maintaining the democratic order rather than through the putative metaphysical status of individuals. And I have argued elsewhere that the former conception seems to be stronger and more responsive to the changing times (Hongladarom, 2007).

Furthermore, the need for robust privacy protection schemes is even stronger as a result of brains linking technologies than ever before. A technology has to be developed that aims at protecting the privacy of individuals when they consent to having their brains linked to the network. This technology has to be developed in tandem with that of brain-to-brain interface. For example, there could be a technology that acts as a firewall preventing unauthorized access to someone's brain content, even when the brain itself is connected to a network. There should be a regulation requiring that whenever and wherever the brain networking technology is developed, privacy enhancing technology must also be developed alongside the other one. At least the user should be able to determine what level and what kind of information inside their brain can be shared and to which circles of friends and acquaintances. As for the merging of individual brains to accomplish a task, the brains should be uncoupled as soon as the task is finished. There is actually no point in keeping the brains connected together when there is nothing substantial to gain any longer. When a team finishes playing a game, each member then separates and goes their own way. The same should also apply for the brain-merged team. Otherwise, this could lead to a superorganism having more than one brain. Perhaps there is no task so complex as to require brains to be merged together at all times.

Most importantly perhaps, political authorities should be kept out of the technology as much as possible. Their role should only be to ensure that the privacy regulations are followed, but they should not infringe upon the citizens' right to privacy themselves. As much as the NSA in the USA is resented for its intrusive practice into people's lives, when the brain networking technology becomes more mature, the likes of the NSA will be much more powerful. They could not only gather information much more powerfully than before, but they can alter the memories, beliefs and desires of citizens through manipulation of their brain contents. By doing so, they can even change the very identity of the citizens themselves. The only way to curb the power of an organization, such as the NSA, is through political means. This will represent one of the greatest challenges to democracy in the years to come.

One might wonder, before we end this section, whether the ethical implications of brain-to-brain communication would be any different from the implications of other-related technologies. What we have discussed so far seems to apply equally well to the older kinds of technologies. After all, issues such as privacy and individual control have been around for a long time. However, I think there are some unique ethical problems that arise out of brain-to-brain communication. It is certainly the case that privacy, individual control and responsibility are going to loom large on any discussion of ethical implications of the technology, but as we have seen, the possibility that brains can be linked together in a network also pose a set of ethical concerns that are unique to the technology. These concerns here are directly related to the metaphysics of the self as discussed earlier. It would be highly unethical to totally absorb the selfhood and individuality of another person through linking up her brain in a network. I do not think anybody would object to that. Absorbing and dissolving the unique individuality of a person would seem to be tantamount to dissolving her own person altogether. This ethical issue is uniquely related to the technology that we are talking about in this paper. Therefore, what we can conclude regarding the problem whether there is a unique set of ethical concerns is that there is indeed such a set. However, that does not mean that the other sets of concerns, such as those related to privacy, would be any less important.

Conclusion

It is quite likely that the direction of technological research today is heading toward a closer integration of mind and machine in 2035. What is interesting is that the integration also makes mind-mind or brain-brain integration possible. There is nothing in principle that would prevent hooking up more than one brain to a machine, or connecting two or more brains together to harness their processing power to tackle a very complicated task. If that happens, the whole notion of what it is to be an individual and a self will have to be rethought. I have offered a way in which that can be done: Instead of viewing the self as being contained in a closed space traditionally defined by the skin, the self can expand outside of the skin and merge temporarily with other selves too. This also has profound implications on the notion of privacy, especially on how it is conceptualized and justified.

Notes

1. See Clausen (2013) for a useful review of the ethical implications of deep brain stimulation and brain computer interfacing. As for the metaphysical analysis, see Witt *et al.* (2013), for an analysis of what it means to remain an identical person before and after deep brain stimulation.
2. The ethical concerns discussed in the paper are not based on any particular religion, though in the paper, I sometimes mention Buddhism as an inspiration. The reference to Buddhism, however, does not imply that I am doing a religious critique of brain-to-brain interface technology. For a religious critique of the same type of technology, see Campbell *et al.* (2008).
3. For an account of the narrative view on personal identity, see Schechtman (2007). Françoise Baylis argues that the narrative account makes it highly unlikely that deep brain stimulation could alter personal identity (Baylis, 2013).
4. Another result from empirical studies that corroborates the point I am making here is the "Extinction of Thought" hypothesis advanced by Kübler and Birbaumer (2008). They note

that patients who lack the ability to move their limbs afterwards lose the ability to control brain-to-computer interface, which would have resulted in their ability to move robotic limbs. Their hypothesis is that if the thought is not followed by a concrete result, the thought itself does not come to mind, so to speak, which appears to show that thinking is directly related to the external environment. The consciousness of the self, then, also seems to depend on the ability of the subject to connect concretely with the outside world.

5. A set of questions that emerge from the possibility of direct brain-to-brain communication concerns how to measure brain activity, the interplay between the subjective and the objective, as well as the effect of the social world on how the brain data can be interpreted. As for the question about measurement, the paper, being a conceptual investigation, relies on the scientific methodologies and results offered by the scientists who are working on the topic. Conceptually speaking, brain data are measured in how much information is generated and transferred across the network. In that case, we could safely imagine a brain to be a network of data, with each neuron working as an individual contact point not unlike the transistor in a computer. The information within the brain is then measured by looking at the pattern of neurons firing or not firing within the brain, and as information can be copied and transmitted, the idea is that the information from one brain can be communicated to another brain, creating a network of brains. In a crude way, then, joining two brains together is comparable to joining two computers together in a network.

   The question about the interplay between the subjective and objective is a more interesting one as it pertains directly to the philosophical conundrum emerging from the research on brain-to-brain communication. Each brain is supposedly a locus for an individual subjectivity – an "I" having her own subjective experiences. Thus, when two brains are put in a network, then the line between the subjective and the objective would seem to crumble. As I shall make clearer in the body of the paper, joining two brains does not have to result in an elimination of individual subjectivity. On the contrary, subjectivity could be enhanced through the "merging of minds" that would be the result of the brain network. Instead of a kind of subjectivity that is limited to an individual body, it could be extended outside of that and engages in direct communication with the other subjectivity belonging to the other brain. This does not have to imply that each individual subjectivity is destroyed because the joining together is temporary, as it is in effect only when the brains are connected.

   The question and answer above leads directly to the last question concerning the social influence on how the data are interpreted. I take this question to mean something like "How should we understand these brain data when we are influenced by social influences?" Presumably, then, a person in China, say, might interpret these brain data differently from one in Germany. However, the focus of my paper is not on how external cultural factors influence, how data are perceived or interpreted, though that could perhaps be an interesting question to ask. My concern is more modest. Given the possibility given by current scientific research on brain-to-brain interfacing, what kind of ethical and metaphysical concerns could develop. In some cases, we find a rehearsing of old philosophical questions. That is to be expected, and I believe in any case that it would be a contribution to philosophical discourse that in the end it is found that brain-to-brain communication gives rise only to age old philosophical problems (that on the self, for example). Nonetheless, I also believe that the phenomenon gives a powerful new twist on the traditional problem. If it were possible that brains can be networked together, it would give a strong impetus in favor of the philosophical view that the self can be extended (as Clark and Chalmers have noted), as well as the view that the source of subjectivity and the person do not always have to coincide. Furthermore, the technology also

gives rise to a unique set of ethical implications centered on the possibility that the self of an individual could be totally absorbed into the network, resulting in a loss of her unique identity. This is potentially a very serious ethical concern.

6. Some further questions also emerge from this merging of brains through the direct interface. The first question concerns how the external data are going to be perceived. If the merging is such that there is an emergent consciousness which is neither any of the individual brain/mind that is part of the merging, then there would be no *external* data to begin with, as all the data within the network would be in a way internal to the emergent network. Secondly, some might ask whether it would be possible if the original consciousness, that is the individual consciousness before the merging, would know the origin of the thoughts, from which of the pre-merging individual do the thoughts come from, for example. This is a scientific question for which I am not qualified to provide a definite answer. Nonetheless, it might conceivably be possible to imagine that the consciousness belonging to one brain could recognize that a particular thought arising within its own "field of thought", so to speak actually comes from the other one he or she is communicating with. In this case, the situation would be analogous to normal conversation where we know that the idea that is received by us through the conversation belongs to the partner. However, once the two sets of consciousness merge, this singling out would not be possible during the merging. Furthermore, one might wonder whether it is possible for one brain to resist or over-ride the thought signal coming from the other brain. Again, I can only give a conceptual answer. If we imagine brains to be a huge network of data flowing through the neurons, interfacing two or more brains together would then be to expand those networks so that they become one larger network. In this case, it should still be possible for an individual brain to resist or even to refuse joining the network. This is an ethical issue. An individual should never be forced to join the brain network if she does not so wish. And once the brain merging is in effect, it should also be conceptually possible for a particular brain to split from the group, so to speak and break out of the joined network. The analogy would again be a team member who refuses to participate any longer and decides to leave the team.

References

Armstrong, D. and Ma, M. (2013), "Researcher controls colleague's motions in 1st human brain-to-brain interface", available at: www.washington.edu/news/2013/08/27/researcher-controls-colleagues-motions-in-1st-human-brain-to-brain-interface/

Baylis, F. (2013), "'I am who I am': on the perceived threats to personal identity from Deep Brain Stimulation", *Neuroethics*, Vol. 6, pp. 513-526.

Campbell, C.S., Keenan, J.F., Loy, D.R., Mathews, K., Wingograd, T. and Zoloth, L. (2008), "The machine in the body: ethical and religious issues in the bodily incorporation of mechanical devices", in Lustig, B.A., Brody, B.A. and McKenny, G.P. (Eds), *Altering Nature: Volume Two: Religion, Biotechnology and Public Policy*, Springer Science, pp. 199-258.

Clark, A. and Chalmers, D.J. (1998), "The extended mind", *Analysis*, Vol. 58 No. 1, pp. 7-19.

Clausen, J. (2013), "Bonding brains to machines: ethical implications of electroceuticals for the human brain", *Neuroethics*, Vol. 6, pp. 429-434.

Damasio, A. (2003), *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, Harcourt, Orlando, FL.

Damasio, A. (2012), *Self Comes to Mind: Constructing the Conscious Brain*, Vintage, New York, NY.

Dennett, D.C. (2013), *Intuition Pumps and Other Tools for Thinking*, Penguin, London.

Direct Brain-to-Brain Communication in Humans: A Pilot Study (2013), available at: http://homes.cs.washington.edu/~rao/brain2brain/

Erler, A. (2011), "Does memory modification threaten our authenticity?", *Neuroethics*, Vol. 4, pp. 235-249.

Hongladarom, S. (2007), "Analysis and justification of privacy from the Buddhist perspective", in Hongladarom, S. and Ess, C. (Eds), *Information Technology Ethics: Cultural Perspectives*, IGI Global, Hershey, PA, pp. 108-122.

Hongladarom, S. (2008), "Privacy, contingency, identity and the group", in Luppicini, R. and Adell, R. (Eds), *Handbook of Research on Technoethics (2 Volumes)*, IGI-Global, Hershey, PA, pp. 496-511.

Hongladarom, S. (2011a), "Personal identity and the self in the online and offline world", *Minds and Machines*, Vol. 21 No. 4, pp. 533-548.

Hongladarom, S. (2011b), "Pervasive computing, privacy and distribution of the self", *Information*, Vol. 2 No. 2, pp. 360-371.

Hongladarom, S. (2013), "Ubiquitous computing, empathy and the self", *AI & Society*, Vol. 28 No. 2, pp. 227-236.

Kircher, T. and David, A.S. (Eds) (2003), *The Self in Neuroscience and Psychiatry*, Cambridge University Press, Cambridge.

Kübler, A. and Birbaumer, N. (2008), "Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients?", *Clinical Neurophysiology*, Vol. 119 No. 9, pp. 2658-2666.

LeDoux, J. (2003), *Synaptic Self: How Our Brains Become Who We Are*, Penguin Books.

Liao, S.M. and Sandberg, A. (2008), "The normativity of memory modification", *Neuroethics*, Vol. 1 No. 2, pp. 85-99.

Merleau-Ponty, M. (1962), *Phenomenology of Perception*, Translated by Smith, C., Routledge, London.

Räikkä, J. (2010), "Brain imaging and privacy", *Neuroethics*, Vol. 3, pp. 5-12.

Rao, R.P.N., Stocco, A., Bryan, M., Sarma, D., Youngqist, T.M., Wu, J. and Prat, C.S. (2014), "A direct brain-to-brain interface in humans", *PLoS ONE*, Vol. 9 No. 11.

Schechtman, M. (2007), *The Constitution of Selves*, Cornell University Press, Ithaca, NY.

Witt, K., Kuhn, J., Timmermann, L., Zurowski, M. and Woopen, C. (2013), "Deep brain stimulation and the search for identity", *Neuroethics*, Vol. 6 No. 3, pp. 499-511.

Corresponding author
Soraj Hongladarom can be contacted at: s.hongladarom@gmail.com