



Vitvælastofnun Íslands ses
Menntavegi 1
2.h. Úranus
101 Reykjavík

Umsögn um

tillögu til þingsályktunar um stuðning Íslands við alþjóðlegt bann á framleiðslu og beitingu sjálfvirkra og sjálfstýrðra vígvéla,

68. mál á 145. löggjafarþingi 2015–2016

Rannsóknnum í gervigreind og sjálfvirkni hefur fleygt fram síðasta áratuginn, sem er jákvæð þróun, enda gífurlegir nýtingarmöguleikar almenningi og heiminum öllum til heilla. Helsta ógnin við jákvæð áhrif slíkrar tækni er siaukin vígbúnaðarvæðing víðsvegar í heiminum og rannsóknir á nýtingu hennar í hernaðarlegum tilgangi. Vígbúnaðarkapphlaup á þessum sviðum mun ógna friði og auka líkurnar á misbeitingu tækninnar og valds gagnvart almennum borgurum. Alþjóðlegt bann getur hægt á, ef ekki stöðvað, slíka framþróun mála.

Til viðbótar gervigreindarsérfræðinganna, og öðrum sem stutt hafa bréf Future Life Institute um bann við sjálfvirkum drápsvélum¹ (yfir 20 þúsund manns, þ.m.t. undirritaður), hefur umræða um málið farið fram í Mannréttindaráði Sameinuðupjóðanna (fundir í apríl sl. í Geneva² með fulltrúa frá 90 löndum, og í október í New York), en á fundi ráðsins 2013 sem 20 þjóðir sátu, og svo aftur 2014, var lögð fram ályktun um bann slíkra vopna. Níu lönd hafa þegar stutt slíkt bann. Þá hafa alþjóðlegu samtökin ICRC (International Committee for Robot Arms Control) barist fyrir málinu síðan 2009. Þá má benda á að seint á síðasta ári varð Vitvælastofnun Íslands fyrsta rannsóknarstofnunin í heimi til að setja sér siðferðisstefnu um rannsóknir sínar í gervigreind (sjá Viðauka), þar sem fjármögnun frá hernaðarsjóðum er alfarið hafnað og þær hátæknirannsóknir fordæmdar sem ætlaðar eru til framleiðslu drápsvéla, skerðingar einstaklingsfrelsis og brota á alþjóðlegum mannréttindasáttmálum.

Vitvælastofnun Íslands fagnar heilshugar þingsályktunartillögu 68 um stuðning við alþjóðlegt bann á framleiðslu og beitingu sjálfvirkra og sjálfstýrðra vígvéla og hvetur Alþingi til að samþykkja hana.

Vitvælastofnun Íslands ses (Icelandic Institute for Intelligent Machines - IIIM) er eina sjálfstæða rannsóknarsetur landsins í gervigreind. Stofnunin hefur það markmið að brúa bilið milli atvinnuvega og háskóla á sviðum gervigreindar, hermunar, og annarrar hátækni. Stofnunin vinnur náið með fjölmörgum aðilum hér á landi við

¹ http://futureoflife.org/AI/open_letter_autonomous_weapons - sótt 15. mars 2016.

² [http://www.unog.ch/80256EE600585943/\(httpPages\)/6CE049BE22EC75A2C1257C8D00513E26?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/6CE049BE22EC75A2C1257C8D00513E26?OpenDocument) - sótt 15. mars 2016.



sameiginleg rannsóknar- og þróunarverkefni, ásamt hagkvæmnisathugunum og framleiðslu frumgerða á þeim sviðum.

Ísland stendur vel að vígi í stuðningi við tillögu um bann sjálfvirkra og sjálfstýrðra vígvéla, m.a. vegna herleysis og óvenju friðsællar fortíðar landsins. Vitvélastofnun Íslands býður hér með fram þekkingu sinna sérfræðinga við þau verkefni sem kunna að vera sett af stað í kjölfar samþykktar tillögunnar, svo sem við að skilgreina nánar hvers konar sjálfvirkni sé átt við, hvernig standa megi að eftirliti á þessu sviði, og frekari aðkomu Íslands að málaflokknum.

Virðingarfyllt,

Dr. Kristinn R. Þórisson
Framkvæmdastjóri og stofnandi
Vitvélastofnun Íslands ses



Vidauki

Siðferðisstefna Vitvélastofnunar Íslands ses (ágúst 2015).

IIM Ethics Policy

The Board of Directors of IIM believes that the freedom of researchers to explore and uncover the principles of intelligence, automation, and autonomy, and to recast these as the mechanized runtime principles of man-made computing machinery, is a promising approach for producing advanced software with commercial and public applications, for solving numerous difficult challenges facing humanity, and for answering important questions about the nature of human thought.

A significant part of all past artificial intelligence (AI) research in the world is and has been funded by military authorities, or by funds assigned various military purposes, indicating its importance to military activities. A large portion of the world's most advanced AI research is still supported by such funding, as opposed to projects targeting directly and exclusively for peaceful civilian purposes. As a result, a large and disconcerting imbalance exists between AI research with a focus on hostile applications and AI research with an explicit peaceful agenda. Increased funding for military research has a built-in potential to fuel a continuing arms race; reducing this imbalance may lessen chances of conflict due to international tension, distrust, unfriendly espionage, terrorism, undue use of military force, and unjust use of power.

Just as AI has the potential to enhance military operations, the utility of AI technology for enabling perpetration of unlawful or generally undemocratic acts is unquestioned. While less obvious at present than the military use of AI and other advanced technologies, the falling cost of computers is likely to make highly advanced automation technology increasingly accessible to anyone who wants it. The potential for all technology of this kind to do harm is therefore increasing.

For these reasons, and as a result of IIM's sincere goal to aim its research towards topics and challenges of obvious benefits to the general public, and for the betterment of society, human livelihood and life on Earth, IIM's Board of Directors hereby states the Institute's stance on such matters clearly and concisely, by establishing the following Ethical Policy for all current and future activities of IIM:

1 - IIM's aim is to advance scientific understanding of the world, and to enable the application of this knowledge for the benefit and betterment of humankind.

2 - IIM will not undertake any project or activity intended to (2a) cause bodily injury or severe emotional distress to any person, (2b) invade the personal privacy or violating the human rights of any person, as defined by the United Nations Declaration of Human Rights, (2c) be applied to unlawful activities, or (2d) commit or prepare for any act of violence or war.

2.1 - IIM will not participate in projects for which there exists any reasonable evidence of activities 2a, 2b, 2c, or 2d listed above, whether alone or in collaboration with governments, institutions, companies, organizations, individuals, or groups.

2.2 - IIM will not accept military funding for its activities. 'Military funding' means any and all funds that are designated to support activities of governments, institutions, companies, organizations, and groups, explicitly intended for furthering a military agenda, or to prepare for or commit to any act of war.

2.3 - IIM will not collaborate with any institution, company, group, or organization whose existence or operation is explicitly, whether in part or in whole, sponsored by military funding as described in 2.2 or controlled by military authorities. For civilian institutions with a history of undertaking military-funded projects a 5-15 rule will be applied: If for the past 5 years 15% or more of their projects were sponsored by such funds, they will not be considered as IIM collaborators.