

Experiencia de estudio grupal. Primera introducción a las armas autónomas.

Federico Otarán¹, Lautaro De León¹, Joaquín Bogado¹, María Emilia Corrons¹, María Beatriz García², Francisco Javier Díaz¹

¹ Laboratorio de Investigación de Nuevas Tecnologías en Informática
LINTI - Facultad de Informática - UNLP

{ldeleon,ecorrns,federico.otaran}@cespi.unlp.edu.ar

{jbogado,jdiaz}@linti.unlp.edu.ar

<http://linti.unlp.edu.ar/>

² Sin afiliación

mariabeatriz.garcia@gmail.com

Resumen Este trabajo está dividido en dos partes, la primera donde presentamos una metodología que nos parece adecuada para encarar temas de ética profesional entre pares, tomada de otros grupos de investigación. Esta técnica, puesta en práctica, nos permitió elegir y estudiar un tema con cierta profundidad. La segunda parte de los objetivos es mostrar los resultados de dicha metodología al estudio de las cuestiones éticas relacionadas con las armas autónomas. Estos resultados son parciales y parte de un trabajo en progreso.

Keywords: ética informática, armas autónomas, máquinas morales

1. Introducción

“Definitivamente desencadenado, Prometeo, al que la ciencia proporciona fuerzas nunca antes conocidas y la economía un infatigable impulso, está pidiendo una ética que evite mediante frenos voluntarios que su poder lleve a los hombres al desastre” Así comienza el prólogo, fechado en New York en julio de 1979 del libro de Hans Jonas titulado “El principio de Responsabilidad: ensayo de una ética para la civilización tecnológica”. [11] Para Jonas es el encuentro entre la ciencia y la tecnología lo que caracteriza a la modernidad. El “progreso” tiende, en efecto, a cambiar la vida humana a una velocidad cada vez más intensa, esto hace necesario un replanteamiento de la ética. “¿Qué podrá servirnos de guía?...;El propio peligro que prevemos!” [11].

Con la idea de comenzar a estudiar temas de ética profesional, en mayo de 2016 conformamos un grupo de estudio, el cual a través de reuniones regulares, nos iba a permitir investigar la forma en la que la aparición de nuevas tecnologías influye en la sociedad y viceversa, cómo la sociedad influye en la aparición de nuevas tecnologías y en los nuevos usos que se hace de la ya existente.

La mecánica propuesta para el desarrollo del proceso de estudio fue basada en el uso de técnicas asociadas a las metodologías ágiles de desarrollo [20], aunque

también tomamos ideas de grupos de estudios de otra áreas como el psicoanálisis. Dicho proceso, sin embargo, fue adaptado a nuestras necesidades particulares sobre la marcha. Decidimos reunirnos una vez por semana durante unas 2 horas, en las cuales cada uno pondría en común sus ideas respecto a lo investigado durante la semana y si bien mantuvimos cierta flexibilidad en cuanto a horarios, que los temas fueran de interés para el grupo ayudó a no perder la constancia.

Como primer paso y durante los primeros encuentros, además de proponer y dar forma a la metodología, hicimos una primera puesta en común al estilo *brainstorm*[21] para seleccionar un tema de estudio. La decisión de encarar uno solo de manera grupal y no uno por cada integrante se basó, en principio, en que el tiempo que podíamos dedicar al proyecto era bastante reducido y era preferible enfocar los esfuerzos lo más posible. De entre los temas propuestos elegimos las Armas Autónomas por diversos motivos, entre los cuales podemos incluir que, sin haber ahondado en el tema, parecía más sencillo, principalmente debido a que la comunidad científica en su mayoría tiene una posición tomada al respecto[1][2][3]. Incluso Argentina se ha manifestado frente a organismos internacionales respecto a la necesidad de control de este tipo de armas[4][5].

A partir de esto, seguimos una serie de pasos que constituyen la columna vertebral de esta metodología. El primero se basó en obtener de las distintas fuentes bibliográficas disponibles, la información necesaria acerca de esta temática que nos permitiera tener una idea clara del estado del arte de las armas autónomas en la actualidad. Posteriormente, en base a la información acopiada e investigada, tratamos de comprender los argumentos a favor y en contra del desarrollo de las armas autónomas. A su vez, durante las reuniones presenciales, debatíamos entre nosotros sobre la contundencia o la debilidad de éstos y registrábamos las dudas que fueron surgiendo al contrastarlos y discutirlos por sobre nuestras ideas y subjetividades para con los sistemas de armas autónomas. Cabe destacar que muchas de ellas quedaron sin contestar.

Sin desalentarnos ni mucho menos ante la falta de respuestas, seguimos adelante convencidos que mostrando que podemos trabajar en responderlas, crearemos conciencia acerca de la necesidad de discutir estos temas en la comunidad.

2. De nuestros antecedentes

El tema de la Ética Profesional fue nutriendo la curiosidad de uno de los jóvenes que conforma este grupo y fue adquiriendo mayor importancia como eje de investigación con el paso del tiempo y las conversaciones informales con personas interesadas en esta problemática. En el contexto de CACIC 2011 con sede en la Facultad de Informática de la UNLP, se realizaba el primer ETHICOMP Latinoamérica, lo que mostraba el auge y la importancia del tema a nivel nacional.

En 2012, también en el ámbito del II ETHICOMP, se presenta un primer trabajo[22] y en 2013 en III ETHICOMP se presenta un segundo trabajo[23], que buscaba de alguna manera continuar algunas líneas de trabajo surgidas a raíz del trabajo presentados el año anterior, tomando como referencia el trabajo de

Matthew J. Sher. La participación y el avance en la investigación en estos temas siempre fue posible gracias al apoyo incondicional de la dirección del Laboratorio de Investigación de Nuevas Tecnologías en Informática (LINTI).

Fue necesario que transcurriese un tiempo para que alguien tomara la iniciativa de formar el grupo de investigación, que hoy estamos presentando en la comunidad científica, conformado por cuatro investigadores del LINTI, estudiantes de la carrera de Licenciatura en Informática y Licenciatura en Sistemas de la UNLP, dos de ellos egresados y dos en la etapa final de la misma.

Como bien señala Adela Cortina en su libro[10] “¿Para qué sirve realmente la ética?”: “El mundo de las profesiones tiene una larga historia, que se suele contar desde la tradición occidental, de la que se dice que nace con el célebre Juramento de Hipócrates, ligado a una profesión tan valorada como la médica. Otras dos profesiones le acompañarían en los orígenes, la de los sacerdotes y los juristas, de modo que entre las tres se ocuparían de cosas tan importantes para la vida de una sociedad como el bien del cuerpo, el bien del alma y el de la comunidad política. Quien ingresa en una profesión se compromete a proporcionar ese bien a su sociedad, tiene que prepararse para ello adquiriendo competencias adecuadas, y a la vez ingresa en una comunidad de profesionales que comparten la misma meta”.

Y un poco más adelante en este mismo libro podemos leer algo del espíritu que anida en este grupo de estudio e investigación y marca su norte: “Frente al *êthos* burocrático de quién se atiene al mínimo legal pide el *êthos* profesional, la excelencia, porque su compromiso fundamental no es el que le liga a la burocracia, sino a las personas concretas, a las personas de carne y hueso, cuyo beneficio da sentido a cualquier actividad e institución social. Es tiempo, pues, no de despreciar la vida corriente, sino de introducir en ella la aspiración a la excelencia.”

3. Las armas autónomas

Entendemos por Arma Autónoma (Autonomous Weapon), Sistema Letal Autónomo (Lethal Autonomous Weapon System, LAWS) o Arma Robótica Autónoma (Autonomous Robotic Weapons, ARW) a cualquier sistema de armamento capaz de tomar la decisión de disparar contra un blanco sin la intervención de un ser humano, de acuerdo con la definición de “Human-out-of-the-loop Weapons”[15]. No se conocen hasta el momento sistemas de este tipo que estén siendo usados en servicio, pero sí existen sistemas de defensa altamente automatizados como el sistema comercializado por Samsung, SGR-A1 utilizado en la zona desmilitarizada que separa Corea del Norte de Corea del Sur. Estos sistemas requieren pocos pasos para volverse completamente autónomos y países como China, Israel, Rusia y el Reino Unido ya han manifestado su interés en el desarrollo de esta tecnología[16]. Que un algoritmo, es decir, un programa de computadora tome la decisión de disparar terminando así con la vida de alguien en el otro extremo del arma, sin que sea otra persona la que aprieta el gatillo,

plantea problemas filosóficos incipientes acerca de diversos factores implicados en dicha decisión.

Chamayou en su libro “A Drone Theory” [6], se pregunta quién sería el responsable en el caso hipotético de que un robot cometiera un crimen de guerra. ¿El general que lo empleó? ¿El Estado, que es su propietario? ¿Los informáticos que lo programaron?

Müller[17] por su parte, sostiene que los LAWS no plantean desafíos sustanciales a las leyes humanitarias ni plantean problemas a la hora de determinar responsabilidades, que las consecuencias de usar este tipo de armas serían en general positivas y que por lo tanto no deberían prohibirse sino regularse. Sin embargo concede que algunos aspectos de las guerras podrían empeorar¹ y que los principios de Guerra Justa (*Jus in Bello*) de distinción y proporcionalidad² podrían verse comprometidos.

Nosotros a nuestra vez nos preguntamos, como desarrolladores de software, si será posible diseñar una máquina incapaz de cometer un crimen de guerra³ pero que de todas formas continúe siendo útil a sus fines prácticos. Las ventajas que constantemente se remarcan en favor de las armas autónomas [7][8][9] incluyen entre otras:

- Reemplazan a los humanos en el campo de batalla, reduciendo así el posible número de víctimas.
- Una máquina es superior a un humano en varios aspectos. La máquina cuenta con un menor tiempo de reacción y con mayor precisión, no necesita descansar y puede funcionar con igual efectividad en ambientes hostiles y situaciones adversas.
- Una máquina no experimenta sensaciones de ira, enojo, miedo o ansias de venganza. Esto evitaría posibles violaciones a las reglas de enfrentamiento y leyes de guerra.

Respecto al tercer punto, para que un robot sea capaz de seguir las leyes de guerra, debería ser posible codificarlas de alguna manera, es decir, plasmarlas en un algoritmo. Dicho algoritmo diferenciaría para la máquina aquello que es moralmente correcto de lo que no lo es. Sin embargo, el trabajo de M. Englert *et al.*[12] propone límites al accionar moral de una máquina. En dicho trabajo, los autores proponen un Gedankenexperiment, una variante del problema del tranvía descrito por Thomson[13], en la cual:

- Solo hay una única acción correcta⁴ entre dos opciones.

¹En particular, Müller habla de que la reducción del costo de la guerra haría la guerra más probable.

²El principio de Distinción refiere a la distinción de Civiles de Militares y es crucial a la hora de seleccionar un blanco, mientras que el principio de Proporcionalidad refiere a que el uso de la fuerza debe estar en proporción al objetivo militar propuesto.

³Un crimen de guerra es una violación a las leyes de guerra y a las reglas de enfrentamiento descritas en diversos tratados internacionales.

⁴Los autores entienden como “correcta” a aquella opción moralmente aceptable y la diferencian de otra que claramente no lo es.

- Toda la información para resolver el problema está disponible.
- Todas las acciones ocurren de manera determinística.
- Y sin embargo, es fundamentalmente imposible reconocer la opción correcta de manera algorítmica.

El escenario propuesto se resume en que el programa debe elegir la alternativa moralmente más aceptable basándose en determinar si el código fuente de un programa contiene instrucciones maliciosas que efectivamente se ejecutarán al correr dicho programa. Usando métodos formales de teoría de la computación y verificación de programas, los autores demuestran que este problema no puede resolverse para todas las entradas. Dicho de otra manera, habrá casos, situaciones particulares para los cuales un algoritmo no podrá decidir cuál es la opción moralmente correcta a pesar de que se cumplan los puntos enumerados anteriormente.

Aquí nos parece razonable distinguir claramente dos conceptos que algunos autores como W. Wallach y C. Allen en su libro *Moral Machines*[14] utilizan de manera intercambiable, los conceptos de la Moral y la Ética. Parece posible, para dichos autores, diseñar una máquina que distinga el bien del mal siguiendo un algoritmo, es decir una máquina moral. Este comportamiento debería ser especificado de antemano por los programadores de dicho algoritmo. Esta tarea que al parecer es en extremo difícil, además estaría limitada, como queda demostrado por los métodos formales descritos anteriormente. No obstante, creemos que el actuar éticamente y aquí la diferencia con la Moral, es una actividad estrictamente humana e individual, que involucra otras actividades intrínsecamente humanas las cuales no son programables, como por ejemplo la búsqueda del buen vivir, el entendimiento y la deliberación[24].

4. Los drones y las armas autónomas

Notamos también un paralelismo entre algunos argumentos a favor del uso de Drones de Combate con fines militares con los argumentos a favor de las armas autónomas. Los Drones de combate o UAV (Unmanned Aerial Vehicles) se diferencian de las Armas Autónomas en que los primeros cuentan con un operador humano (dos en el caso del MQ-9 Reaper, un piloto y un operador de sensores). Por un lado, el dron permite proyectar fuerza sin proyectar vulnerabilidad[6], esto es, el dron es capaz de infringir fuerza letal a la vez que resguarda la vida de los operadores humanos que generalmente se encuentran en la sala de comando. De esta manera se entiende que el uso de drones permite “salvar vidas”. Sin embargo, las vidas salvadas son las de los combatientes del propio bando que las utiliza, mientras que las vidas que peligran son las de los militares y civiles del otro bando. De la misma manera, las Armas Autónomas permiten salvar vidas, a la vez que, como los drones, eliminan el combate y la posibilidad de defensa del enemigo. Así la guerra deja de ser tal para convertirse en una mera matanza, una aniquilación del enemigo.

La guerra, un estado de conflicto armado entre dos Estados o grupos dentro de un Estado[19], es el único en el cual se le está permitido a una persona tomar

la vida de otra sin que esto sea considerado un crimen. Mejorar las armas con el fin de hacer la guerra más barata bajaría el umbral para empezar un conflicto bélico en lugar de resolver las diferencias por otros medios.

El uso de los drones como dispositivos bélicos hecho por la administración de Obama ha sido duramente cuestionado[18] y sus diversos matices se plasman en la película “Eye in the Sky” dirigida por Gavin Hood y protagonizada por Helen Mirren, Aaron Paul y Alan Rickman. En el film puede verse cómo las reglas de enfrentamiento se ven modificadas en pos de acabar con la vida de cinco conocidos terroristas a pesar de los posibles daños colaterales a la población civil de un país aliado. A la vez que muchas de las situaciones del film plantean los mismos interrogantes si se reemplazara el dron por un arma autónoma, encontramos que en el caso del arma autónoma aparecen posibilidades todavía más siniestras. Está claro que, si bien estas armas podrán tomar decisiones de vida o muerte, actuarán siguiendo órdenes que podrían estar en conflicto con las leyes de guerra, pero si el arma no tiene mecanismos para rehusar una orden, podría estar siendo usada para cometer un crimen.

Supongamos, por otra parte, que se cometiera dicho crimen, adrede o de manera accidental, y que el arma tomara la decisión de disparar contra una persona terminando con su vida. Chamayou[6] argumenta que por un lado existe la posibilidad que los responsables del despliegue, la programación o los dueños de dichos sistemas (el Estado o una empresa privada, por ejemplo) puedan “pasarse la pelota”, y por otro lado, que el único responsable claramente identificado es la propia víctima, la cual por sus actos voluntarios o involuntarios, accionó los mecanismos para que el arma dispare.

5. La posición Argentina respecto a las Armas Autónomas

La posición oficial de Argentina respecto a las armas autónomas puede resumirse en el siguiente extracto tomado de un documento presentado en un debate ante el “Human Rights Council” en Ginebra el 30 de Mayo de 2013[4]: “Deseamos destacar nuestra preocupación por las referencias a que estos sistemas podrían llevar a una ‘normalización del conflicto’, y a una posible carrera de armamentos que crearía divisiones entre los Estados y debilitaría el derecho internacional; el posible fomento de las represalias, la retaliación y el terrorismo; y su impacto en materia de derechos humanos y del derecho internacional humanitario. Como forma de evitar estas consecuencias negativas, el informe concluye en que debe establecerse un órgano internacional encargado de examinar la situación y articular las opciones a más largo plazo.”

Entendemos que el desarrollo de armamento es un negocio global y que una censura total a la investigación y desarrollo de este tipo de dispositivos podría incomodar a muchos Estados. Pero los interrogantes al respecto de las armas autónomas siguen siendo demasiados como para alentar su desarrollo sin que haya garantías, entre las cuales, podemos destacar el acceso a la justicia para

aquellas víctimas del abuso de este tipo de armamento o la falta de certezas en cuanto a las ventajas que éstas brindan.

6. Conclusiones y trabajo a futuro

La metodología descrita nos permitió encarar un tema de estudio como grupo. En futuras iteraciones debemos mejorar aspectos como sistematizar la toma de notas, referencias bibliográficas y extractos. También encontramos dificultades a la hora de fundamentar algunos argumentos, probablemente debido al desconocimiento de las principales corrientes filosóficas. Creemos que sería beneficioso un estudio de dichas corrientes para luego poder aplicarlas a temas específicos concernientes a la ética informática.

Además, consideramos que la postura que toma Argentina sobre la temática de armas autónomas es débil; no es suficiente contar con un organismo que regule su uso. Nos parece que Argentina debería desalentar el desarrollo de las mismas y no invertir fondos públicos en la investigación de estas tecnologías.

Como trabajo a futuro, creemos también necesario compartir nuestras inquietudes y reflexiones con nuestros colegas y con la comunidad en general, como parte integral de las actividades del grupo y a intervalos regulares. De esta manera, el grupo de investigación estaría contribuyendo a uno de los principales objetivos de la Universidad, que es concientizar a las futuras generaciones de profesionales. Estas presentaciones previenen trabajos como el presentado, pero también a través de charlas, paneles y foros de discusión.

Referencias

1. Stop Killer robots Campaigning - <http://www.stopkillerrobots.org>
2. International Committee for Robot Arms Control - <http://icrac.net/>
3. Autonomous Weapons: an Open Letter from AI & Robotics Researchers - <http://futureoflife.org/open-letter-autonomous-weapons>
4. Document delivered by Mr. Mariano Alvares Wagner on behalf of GRULAC. 2013 - http://www.stopkillerrobots.org/wp-content/uploads/2013/05/HRC\Argentina_09_30May2013.pdf
5. Country Statements about Killer Robots. 2014 - http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_CountryStatus_14Mar2014.pdf
6. *Grégoire Chamayou*. A drone theory. Kindle Edition, Ed Penguin 2015.
7. *William M. Fleischman*. Why We Should Not Build Autonomous Robotic Weapons. ETHICOMP 2013.
8. *Ronald Arkin*. Lethal Autonomous Systems and the Plight of the Non-combatant. 2013
9. *Ronald Arkin*. Ethical Robots in Warfare. www.cc.gatech.edu/ai/robot-lab/online-publications/arkin-rev.pdf
10. *Adela Cortina*. ¿Para qué sirve realmente la ética?. Ed.Paidós 2013.
11. *Jonas Hans*. Principio de Responsabilidad - Ensayo de una ética para la civilización tecnológica. Ed. Herder 1995.
12. *M. Englert, S. Siebert, M. Ziegler*. Logical Limitations to Machine Ethics with Consequences to Lethal Autonomous Weapons. CoRR 2014.

13. *Judith J Thomson*. The Trolley Problem. Yale Law Journal 1985.
14. *Wendell Wallach, Colin Allen*. Moral Machines: Teaching Robots Right from Wrong. Kindle Edition, Oxford University Press 2008.
15. Losing Humanity, a case against killer robots. Humans Right Watch 2012.
16. Shaking The Foundations: The Human Rights Implications of Killer Robots. Humans Right Watch 2014.
17. *V. Müller, T. Simpson*. Autonomous Killer Robots Are Probably Good News. PhilPapers 2015.
18. The Forgotten Victims of Obama's Drone War. The New York Times 2013
<http://www.nytimes.com/2013/05/23/opinion/the-forgotten-victims-of-obamas-drone-war.html>
19. Guerra. Diccionario de la real academia española, edición digital. - <http://dle.rae.es/?id=JoNx0nS>
20. *Martin Fowler*. The New Methodology. Blog post, 2005. - <http://martinfowler.com/articles/newMethodology.html>
21. *Alex Faickney Osborne*. Applied Imagination: Principles and Procedures of Creative Problem Solving. Ed. Scribner 1953.
22. *J. Bogado, B. García*. Reflexiones iniciales acerca de la validez ética de la utilización de técnicas de minería de datos sobre datos personales en la búsqueda de terroristas. ETHICOMP 2012.
23. *M. B. García, W. Fleischman, J. Bogado*. Una conversación con Matthew Sher sobre privacidad y la amistad. ETHICOMP 2013.
24. *Aristóteles*. Ética a Nicómaco. Ed Gredos 2010.