

ÉTICA E INTELIGENCIA ARTIFICIAL ¿NECESIDAD O URGENCIA?

Omar Danilo Castrillón
Gómez
Ingeniero de Sistemas,
Ph.D.
Departamento de
Ingeniería Industrial,
Universidad Nacional de
Colombia
Manizales, Caldas 0017,
Colombia

María del Pilar Rodríguez
Córdoba
Ingeniera Industrial, Ph.D.
Departamento de
Ingeniería Industrial,
Universidad Nacional de
Colombia
Manizales, Caldas 0017,
Colombia

Juan David Leyton
Castaño
Administrador de Sistemas
Informáticos
MSc en Administración
Universidad Nacional de
Colombia
Manizales, Caldas 0017,
Colombia

RESUMEN

La Inteligencia Artificial (IA), se ha definido como la habilidad que tienen los computadores de hacer tareas que hasta ahora hacen mejor los humanos. La IA, tiene por objetivo hacer que los computadores no sólo puedan procesar información, almacenarla y realizar acciones repetitivas, sino que sean capaces de entender la información, simular las competencias de un experto humano a partir de datos a priori cumpliendo tareas inteligentes, dejadas sólo para los humanos.

A través de los tiempos, la IA se ha desarrollado con gran empuje, llegando a estar presente en diversos ámbitos de la sociedad. Por tal motivo, se generan preguntas relacionadas con los impactos del uso de la IA, los cuales tienen, entre otras, una connotación ética.

En tal sentido, es importante considerar que en un futuro muy cercano la incorporación de la ética en los sistemas inteligentes como: Sistemas expertos, redes neuronales, sistemas evolutivos, etc., será una tendencia activa en el ambiente industrial de países con alto desarrollo tecnológico y con una gran inversión en investigación y desarrollo. Sin embargo, dicha incorporación puede traer altos sobrecostos en el diseño y construcción de estos sistemas, surgiendo así nuevas inquietudes.

A partir de una discusión sobre dichas inquietudes, la presente ponencia plantea que la ética es una necesidad y una urgencia para la IA en los años actuales y venideros.

PALABRAS CLAVES: Ética, Inteligencia Artificial, Sistemas Inteligentes, Agentes Éticos.

1. INTRODUCCIÓN

Hace 50 años era casi fantástico pensar que llegaran a existir máquinas que pudiesen pensar como los humanos y, más aún, realizar movimientos motrices como los humanos, era visto como solo cosas que pasaban en las películas. Ahora vemos que todo eso y muchas otras cosas pueden suceder en un futuro no muy lejano.

Las técnicas de la IA han tomado gran fuerza en los últimos tiempos. Actualmente son utilizadas en un sinnúmero de actividades relacionadas con la vida cotidiana y, especialmente, con organizaciones productoras de bienes y servicios. Algunos ejemplos son la seguridad, las comunicaciones, la programación, secuenciación de la producción y la toma de decisiones.

Estas técnicas han suscitado una variedad de debates, entre los cuales hay uno que llama poderosamente la atención y es el que se refiere a qué tan conveniente puede ser reemplazar a un ser

humano por una máquina. Es allí donde entra la ética a terciar en la discusión.

Por ello, en la presente ponencia plantea que la ética es una necesidad y una urgencia para la IA en los siguientes términos: Necesidad de crear una conciencia que logre cambiar los modos de pensar de mucha gente sobre la urgencia de la incorporación de la ética dentro de actividades de diseño y desarrollo de agentes o máquinas inteligentes. De esta manera se garantizan dos requisitos: Primero, que se tengan en cuenta las implicaciones del uso de técnicas de IA en la generación y mantenimiento del empleo. Segundo, que dichas máquinas no atenten contra la integridad de los humanos y cumplan con una serie de códigos de conducta, restringiéndoles sus acciones, y no se desvíen de las normas que rigen a los humanos [1].

Con base en lo anterior, la presente ponencia se desarrollará en cuatro partes. Una primera parte conceptualizará sobre la ética; una segunda parte analizará la IA; un tercer punto se enfocará a la incorporación de la ética en la IA y; por último, se presentarán unas conclusiones generales sobre el tema, las cuales sirven de reflexión sobre el quehacer ético en el mundo industrial y más enfáticamente en la IA.

2. LA ÉTICA Y SUS INFLUENCIAS EN LOS DESARROLLOS TECNOLÓGICOS

El origen de la palabra ética son las expresiones griegas “ethikós” y “éthos” que significan carácter. De hecho, la ética refleja el carácter de los individuos y las sociedades a través de sus decisiones y comportamientos. En general, la ética puede entenderse como un cuerpo de conceptos filosóficos, un grupo de principios que involucran los comportamientos humanos, el mundo espiritual y la vida diaria. Adicionalmente, la ética estudia cómo nuestras decisiones afectan a otras personas; los derechos, obligaciones y normas morales que los individuos aplican en su toma de decisiones y; la naturaleza de las relaciones humanas. El objetivo de la ética es el mejoramiento de la vida humana, pero este objetivo debe tener en cuenta que las personas viven en sociedad. Por lo tanto, la ética incluye el respeto por los demás y el cuidado del medio ambiente [2].

La ética es una ciencia filosófica que se encarga de reflexionar sobre los comportamientos morales del

ser humano con el objetivo de realizar valoraciones genéricas que puedan ser universalizables. En este sentido, la ética, como reflexión filosófica, no se puede confundir con la moral aunque está directamente relacionada con ella, pues se encarga de estudiarla con la ayuda de un método científico, es decir, la moral es el objeto de estudio de la ética.

Como ciencia, los orígenes de la ética se remontan a la Grecia Clásica, siendo Sócrates, Platón y Aristóteles algunos de los representantes más conocidos [3].

Aquí, es importante tener en cuenta que la ética no se encarga de establecer lo que es bueno o malo, correcto o incorrecto, sino de reflexionar sobre las diferentes morales y de analizar por qué para éstas determinados comportamientos son catalogados como adecuados y otros como inadecuados. En este orden de ideas, la ética no dice cómo actuar sino que busca los criterios que justifican actuar de un modo o de otro

Una característica propia de la ética es su tendencia a la generalización o universalización de sus principios, basada en que a pesar de la diversidad de morales existentes a través de la historia y de las culturas humanas, existe homogeneidad en cuanto todos los seres humanos tenemos una competencia moral que nos permite hacer juicios de valor entre lo que consideramos bueno o malo.

En su carácter científico, la ética busca responder a preguntas relacionadas con el qué, el cómo, el por qué y el para qué de las realidades morales. Es decir, la ética parte de una descripción de la realidad moral (el qué), pero no se queda ahí, sino que continúa su búsqueda científica de respuestas, no sólo hacia el deber ser (ética normativa), sino también hacia la interpretación del cómo y por qué de las diferentes (e incluso opuestas) realidades morales, para llegar a un para qué (teleología) y a una generalización de los conceptos correspondientes.

La ética tiene una connotación histórica, pues reconoce la diversidad de morales en el tiempo, con sus correspondientes valores, principios y normas. Dichas morales son el objeto de estudio de la ética, pues ésta las estudia, las entiende, las explica y las trasciende por medio de conceptos, hipótesis y teorías, sin identificarse con una o un grupo de morales en particular. Para ello, la ética sostiene una relación dinámica y recíproca en donde nutre y se

nutre de otras ciencias humanas como la psicología, la sociología, la antropología, la historia, la política y la economía [2].

Como ciencia, la ética se sirve de los diferentes métodos científicos y concepciones metodológicas que permiten generar y proporcionar conocimientos sistemáticos, metódicos, válidos y confiables. Asimismo, se habla de ética aplicada cuando se incorpora la ética a la discusión de temas relativos a aspectos específicos de ciertas profesiones o áreas del conocimiento como la bioética, la ética empresarial, la ética jurídica, entre otras.

Dentro del campo de la ética aplicada a la IA, se ha llegado a hablar de la “Roboética” [4], la cual trata de las situaciones que se pudiesen presentar entre los robots y/o máquinas inteligentes en la interacción con los humanos, la sociedad y el medio ambiente. Este nuevo campo de investigación considera a los robots, computadoras y máquinas inteligentes un tipo de agentes éticos e intenta implementar en ellos facultades morales de toma de decisiones [5].

En los agentes éticos citados por algunos autores [6] se distinguen dos clases, los agentes éticos implícitos, que son máquinas programadas para evitar conductas inmorales, comportándose éticamente y los agentes éticos explícitos los cuales son máquinas que calculan la mejor acción a seguir, usando principios éticos y algoritmos de solución de dilemas éticos.

Dentro de las facultades dadas a estas máquinas es donde se deben analizar las capacidades y/o las responsabilidades que se les asignan. Dado que se les asigne alguna responsabilidad, surgiría esta pregunta: ¿Hasta qué punto los humanos entraríamos a depender de ellas?

De otro lado, en la incursión de la IA en las empresas productoras de bienes y servicios (que son las más beneficiadas), se observa la reducción en costos al sustituir a las personas por máquinas que realizan funciones (generalmente repetitivas) de una manera rápida y eficiente. Esta sustitución de personas no sólo genera desempleo sino un temor o un estrés constante que ronda a diario en las poblaciones de trabajadores más vulnerables [7].

Por tanto, para estas máquinas que serán construidas, y a las cuales se les implantarán

emuladores del comportamiento humano, es necesario contar con unos constructores comprometidos en no afectar la especie humana, la sociedad, ni el medio ambiente, es decir, con unas fuertes bases éticas.

Y es que en el campo de los avances tecnológicos, siempre existe algún temor por parte de la sociedad, no solo por el avance tecnológico en sí, sino también por el tratamiento responsable que se le dé a las aplicaciones y desarrollos. Lo que preocupa no es tanto la tecnología en sí misma, sino mas bien la utilización que se le pueda dar a [8].

Algunos autores más futuristas trabajan en observar hasta qué punto estas máquinas ocuparán un lugar en la sociedad [9]. Por ejemplo, dichas máquinas y/o robots se podrían reunir y crear asociaciones o comunidades de agentes éticos. ¿Podrían estas comunidades exigir ser reconocidos legalmente? Si esto llegara a suceder ¿cómo debería juzgarse el maltrato, la obsolescencia y hasta la muerte de una máquina? o ¿será, entonces, simplemente como golpear una lata o un pedazo de plástico? [10].

Autores como Monopoli, [4] complementan lo anteriormente descrito y dan pie a la existencia de movimientos futuros que se congreguen en torno a la defensa de los derechos de las máquinas.

En síntesis, en este primer punto se afirma que los avances tecnológicos están ligados no sólo a la capacidad humana, sino también a las repercusiones que pudiesen tener en las condiciones de la clase trabajadora. Igualmente, estos avances están entrelazados con todas las sociedades donde estas máquinas formarán parte de la vida diaria e interaccionarán con las personas. Todo lo cual nos lleva a la siguiente pregunta: ¿Están nuestras leyes bien preparadas para asumir los retos que nos traen las nuevas tecnologías?

La conclusión de este punto es que el desarrollo de la ética aplicada a la IA es ya una necesidad urgente. Esta incorporación se debe hacer no solo con respecto a los robots actuales, sino también a las máquinas inteligentes que se desarrollen en el futuro. Para ello, se debería incluir un software que emule una serie de datos introducidos como base para cumplir los códigos o reglas definidas por los programadores para no ser nocivas ante una determinada situación con los humanos.

3. LA IA Y EL DESARROLLO TECNOLÓGICO

Con el nombre de IA se reúne un conjunto bastante diverso de disciplinas para el cumplimiento de unos objetivos específicos, resolución de problemas de manera automática y/o tareas que manejan un alto grado de complejidad e inteligencia [11].

Estas técnicas, que son muy nuevas, pueden ser definidas como la capacidad que tiene una máquina para desarrollar alguna actividad tanto de procesamiento de datos como de almacenamiento y de toma de decisiones. Dichas técnicas se basan en actividades copiadas de los humanos, por eso se dice que emulan o simulan sus comportamientos. Actualmente, están muy presentes en las empresas de todo el mundo y son de gran utilidad a la hora de rendir cuentas sobre los costos y beneficios de los procesos productivos de bienes o servicios.

Técnicas como los sistemas expertos, las redes neuronales, las heurísticas, la visión por computador, el reconocimiento de voz, la lógica difusa, agentes inteligentes, algoritmos genéticos, colonia de hormigas entre otras, pretenden realizar un trabajo más eficaz y exacto que el que realizan los seres humanos.

Esta capacidad de emular o simular las actividades humanas, que en algunos momentos puede superar a los humanos por su alta capacidad de procesamiento y por ser máquinas que no necesitan de descanso, ha reemplazado a muchos seres humanos en puestos de trabajo alrededor del mundo. Este problema entra a formar parte de un proceso cultural, como lo explica Gorniak [12] en “La revolución de los computadores y los problemas globales éticos”.

La tecnología es, en el ser, un modo de existencia humana, entrando en lo más profundo de la vida de las personas. Por ejemplo, cualquier procedimiento quirúrgico en nuestros tiempos ya emplea técnicas de IA, siendo de utilidad en operaciones que necesitan alto grado de precisión y de serenidad. Sin embargo, esto no quiere decir que se pueda reemplazar al médico, cayendo en el error de perder las inspiraciones y las acciones humanas de tal manera que lleguemos a compararnos con máquinas.

Con esta postura queremos significar que la tecnología es, también, un proceso cultural que se debe estar dispuesto a aceptar y aplicar. Para no

pasar a ser parte del pasado, las personas se deben actualizar a diario, dependiendo de su campo de interés. No obstante, la tecnología por sí sola no resuelve los problemas de la sociedad actual; por ello, debe estar acompañada de un conocimiento de las ciencias humanas, entre las que destaca la ética.

De otro lado, la IA se puede explicar como el resultado de una interacción producida por millones de probabilidades de ocurrencia y de cálculos realizados por las máquinas para tomar la mejor decisión y saber cómo realizar el trabajo pedido de una mejor forma y más rápida. Además, estas máquinas pueden almacenar la información, transformarla y obtener nuevos resultados [13].

Como conclusión de este punto, volvemos a hacer énfasis en la necesidad urgente de que los agentes artificiales sean desarrollados de tal manera que los seres humanos puedan participar en las demandas morales que la sociedad actual requiere. De igual manera, surge el siguiente interrogante: ¿Bajo cuáles condiciones los agentes artificiales tendrían sus propias necesidades o intereses que exigen la preocupación moral de humanos? [14].

4. INCORPORACION DE LA ÉTICA EN LA IA

Como se ha observado en el escrito, a pesar de los aportes existentes, la IA aún no tiene unos límites éticos claros. Por tanto, es necesario ahondar en el estudio de este tema y acordar unos lineamientos o códigos de conducta que sean seguidos por desarrolladores, productores y consumidores. De esta manera podremos estar más conscientes de las consecuencias de nuestros actos y podremos establecer un contrato social o compromiso con una sociedad urgida de acciones éticas.

Preguntas como ¿cual deber ser el comportamiento ético que deban seguir los robots, máquinas y sistemas automatizados construidos en el futuro?, pueden generar ya sus propias conclusiones sobre la necesidad de incorporar desde el diseño de estas máquinas un componente ético y no sólo un programa que emule ciertos comportamientos o funciones humanas. Igualmente, significa que debemos ir más allá y fortalecer desde nuestros niños una enseñanza basada en el respeto por los demás, el respeto por el medio ambiente y el respeto por cada desarrollo que influya en el entorno del ser humano.

5. CONCLUSIONES

La evolución de la vida humana es una constante, los conocimientos crecen a un ritmo exponencial y con ellos las necesidades individuales y sociales. En consecuencia, todo evoluciona y las máquinas que tenemos también evolucionan a una velocidad pasmosa. Dentro de todo ello está la IA, cuyo rol protagónico tiene una inmensa capacidad para ayudar a mejorar o para lo contrario.

Por ello reiteramos, junto con Mowshowitz [15] que en la interacción de los humanos y los robots y/o máquinas inteligentes que empiecen a mostrar una conducta autónoma, se deben instaurar o crear códigos éticos que rijan o limiten las acciones de dichas máquinas. De igual manera, se deben implantar dichos códigos en los procesos de producción de otras máquinas inteligentes.

Por otra parte esta la ética positiva, encargada no de la prohibición de conductas potencialmente nocivas, sino de la motivación de conductas tendientes a generar bienestar individual y social. Esta ética positiva se debe aplicar, con el fin de mejorar la calidad de las relaciones que se creen con las máquinas, de tal manera que éstas no atenten contra las personas y conviertan el mundo en una guerra de poderes entre máquinas y humanos, como lo ilustran algunas obras literarias y cinematográficas de ciencia ficción. En este sentido, debe quedar siempre claro que las personas son las que dominan y las máquinas son herramientas que ayudan a crear mayores opciones de bienestar, a mejorar la calidad de vida y a dignificar los trabajos que se realizan. Finalmente, tal vez en un futuro veamos sistemas tan automatizados que la vida activa de los seres humanos sea reducida a la toma de aquellas decisiones de suma importancia. Por ello, al pensar en cómo serán los próximos 50 años, no debemos olvidar que los avances tecnológicos son una constante, enmarcada en los confines de la vida humana. No sabemos si esta vida pueda llegar a ser destruida por la tecnología. De lo que si estamos seguros es que esas preguntas podrán ser resueltas a favor del género humano en la medida en que se incluya, como una necesidad urgente, el debate ético en todo avance de las técnicas de IA. Como lo mencionan Mulej et al. [16], la ética se convierte en códigos a seguir por las personas, la sociedad y la comunidad, siendo igual o de mayor importancia que la creatividad, las habilidades y el conocimiento.

La ética es una ciencia práctica que es fundamental en todos los aspectos de la vida diaria. Esta ciencia filosófica nos ayuda a reflexionar sobre los comportamientos correctos e incorrectos y sobre cómo llegar a tener una buena vida en interacción con otros seres humanos y en armonía con el medio ambiente.

La ética es un componente clave para el desarrollo actual y futuro de la IA, pues es la base para mantener la integridad y la legitimidad de de sus aplicaciones y para salvaguardar la misma vida.

Por lo anterior, los desarrollos tecnológicos deben estar sujetos a la aprobación y verificación de la sociedad no sólo con criterios técnicos, sino también con criterios éticos.

La IA es un campo nuevo y en continua evolución, en el que todo está por investigar y desarrollar. De ahí la necesidad de generar en la sociedad la confianza de que lo que se produzca bajo sus técnicas está elaborado con una serie de principios y códigos éticos que guiaran el comportamiento de dichas máquinas.

Avizoramos que la creación de valor a partir de la incorporación de la ética en los nuevos productos llevará a marcar la pauta entre las organizaciones dedicadas al desarrollo y producción de tecnología. Esto se podría convertir un estándar internacional, mejorando tanto la calidad de vida de los empleados de estas organizaciones, como la de sus clientes.

Dichos estándares crearían entre los países una gran diferencia, llevándolos a que trabajen en la vinculación de la ética dentro de cada proceso que se desarrolle en sus organizaciones. Esta sería una gran opción para la mejora de países con altos niveles de corrupción y de tecnocracia en la automatización industrial.

En esta ponencia han surgido muchas más preguntas que respuestas. Por ejemplo ¿hasta dónde podemos llegar con la tecnología?, ¿hasta dónde quieren las personas reemplazarse por máquinas capaces de realizar infinidad de procesos matemáticos y de razonamiento en un menor tiempo? Y aún más lejos, ¿hasta qué punto estas máquinas tendrán la capacidad de sentir, demostrar sentimientos, sobrepasar frustraciones, alegrarse por sus logros?,

¿llegarán a sentir frío, calor, la calidez de una mañana, el rocío al amanecer y otras sensaciones que solo pueden ser percibidas por los humanos? Si lo anterior llegara a cumplirse, ¿Podrán los robots o las máquinas inteligentes llegar a tener los mismos derechos que tienen los humanos?, ¿llegarán a tener estas máquinas procesos cognitivos como los seres humanos?, ¿podrán heredar las capacidades de sus padres?

Consideramos que el sólo hecho de imaginar respuestas a estos interrogantes significa un avance en la aplicación de la ética a la IA. Igualmente, esperamos que esta ponencia sirva de motivación para aquellas personas interesadas en la investigación en esté difícil pero necesario y urgente desafío para los profesionales del área.

6. REFERENCIAS BIBLIOGRAFICAS

- [1] Proença, Ana (2001). Conversando sobre inteligencia artificial. Lisboa, Revista Intelecto No. 5, febrero. Disponible en http://www.geocities.com/revistaintelecto/ia_e.html, Consultado el 1 de mayo de 2008.
- [2] Rodríguez Córdoba, M. P. (2005). Comportamiento Ético Gerencial. Manizales: Centro de Publicaciones Universidad Nacional de Colombia Sede Manizales.
- [3] Sánchez, A. (2005). Ética. México D.F: Random House Mondadori.
- [4] Monopoli, Antonio (2005). Roboetica. Disponible en <http://www.roboetica.it/>. Consultado el 3 de Abril de 2008.
- [5] Moor, James (2006). The Nature, Importance, and Difficulty of Machine Ethics. IEEE Intelligent Systems, julio/agosto, Vol. 21, N° 4, p. 18-21.
- [6] Anderson, Michael & Leigh Anderson, Susan (2007). Machine Ethics: Creating an Ethical Intelligent Agent. AI Magazine. P. 15 – 25.
- [7] Cordeiro, William P (1997). Suggested Management Responses to Ethical Issues Raised by Technological Change. Journal of business ethics, Vol. 16, p. 1393 – 1400.
- [8] Moriello, Sergio (2007), La Ética para las máquinas, nuevo campo de investigación. Disponible en http://www.tendencias21.net/La-etica-para-las-maquinas,-nuevo-campo-de-investigacion_a1557.html, consultado el 30 de abril de 2008.
- [9] Lucas, Chris (2006). Alien Consciousness - Philosophical Problems Version 4.83, mayo. Disponible en <http://www.calresco.org/alien.htm>, consultado el 2 de mayo de 2008.
- [10] Storrs Hall, John (2000), Ethics for Machines. Disponible en <http://autogeny.org/ethics.html>, consultado el 28 de abril de 2008.
- [11] De Cuadra Garcia, Fernando (2002). Límites Éticos de la Inteligencia Artificial. Revista hispanoamericana de cultura Razón y Fe, enero/febrero 2002, p. 10-15.
- [12] Gorniak-Kocikowska, Krystyna (1996). The Computer Revolution and the problem of Global Ethics. Science and Engineering Ethics. Vol. 2. P. 177-190.
- [13] Zhong-Zhi Shi & Nan-Ning Zheng (2006). Progress and Challenge of Artificial Intelligence. Journals Computer, Science & Technology, Sep, Vol. 21, No. 5, p. 810-822.
- [14] Torrance, Steve (2008). Special Issue on Ethics and Artificial Agents. AI & Soc Vol. 22 p.461–462.
- [15] Mowshowitz, Abbe (2008). Technology as excuse for questionable ethics. AI & Soc Vol. 22 p.271–282.
- [16] Mulej, Matjaž; Kajzer, Stefan; Ženko, Zdenka y Potocan, Vojko (2002): Ethics of interdependence and systems thinking. Proceedings on IDIMT-2002, p. 97/117.