

ÉTICA Y SISTEMAS INTELIGENTES

Prof. Dr. Antonio Anselmo Martino*

Proposición 7 de Baruch Spinoza *Ética more geométrico*: El orden y conexión de las ideas es el mismo que el orden y conexión de las cosas.

Quinta parte: Del poder del entendimiento o de la libertad humana

Paso, por fin, a esta última parte de la *Ética*, que trata de la manera de alcanzar la libertad, es decir, del camino para llegar a ella. En esta parte me ocuparé, pues, de la potencia de la razón, mostrando qué es lo que ella puede contra los afectos, y, a continuación, qué es la libertad del alma, o sea, la felicidad.

Wittgenstein carta a Ludwig von Fricker: Mi libro consiste de dos partes: la aquí presentada, más lo que no escribí. Y es justamente esta segunda parte la más importante. Mi libro traza, por así decirlo, desde el interior los límites de la esfera ética, y estoy convencido que esta es la ÚNICA forma rigurosa de trazar sus límites. En pocas palabras, considero que ahí en donde tantos otros hoy peroran, me las arreglé para poner todo bien en su sitio callándome sobre ese asunto

Proposición 2.04 del *Tractatus*: *La totalidad de los eventos subsistentes es el mundo*

Proposición 6.421 Es claro que la ética no puede formularse. La ética es trascendental (ética y estética son una sola cosa)

Miembro de Academia Nacional de Derecho y Ciencias Sociales de Córdoba, Profesor emérito de las Universidades de Pisa, Italia y Salvador, Argentina

Introducción.

Lo primero que va dicho es que los sistemas inteligentes están llegando a casi todas las partes de nuestra vida y como han crecido en forma vertiginosa van creando mundos diferentes. Como toda invención o descubrimiento humano importante los sistemas inteligentes son peligrosos y por ello es necesario encontrar límites morales a su uso.

La inteligencia artificial se ha desarrollado en modo estrepitoso en los últimos años no solo desde el punto de vista teórico sino también práctico. No hay parte de las humanas actividades que no hayan sido alcanzadas por ellas desde los teléfonos celulares hasta las armas de la guerra cibernética.

Hay muchas definiciones de I.A. Una de ellas, la que se va afirmando, es decir proporcionar a una computadora capacidad de aprendizaje y aplicación de modelos estadísticos no programados explícitamente.

Prefiero hablar de sistemas inteligentes que es aún más amplio pero que voy a definir en modo sencillo como aquellos que presentan un comportamiento externo similar en algún aspecto a la inteligencia humana o animal. Se caracterizan por su capacidad para representar, procesar y modificar de forma explícita conocimiento sobre un problema, y para mejorar su desempeño con la experiencia. Esto les permite resolver problemas concretos determinando las acciones a tomar para alcanzar los objetivos propuestos, a través de la interacción con el entorno y adaptándose a las distintas situaciones.

Sistemas inteligentes se remiten a “pensar” pero el vocablo en el Diccionario de la Real Academia española tiene siete casos que van desde formar o combinar ideas o juicios. Examinar mentalmente algo con atención, opinar, tener intención de hacer algo, formar en la mente un juicio, recordar, tener en consideración algo o a alguien.¹ A esto podría añadirse recordar, deducir, querer, inducir, odiar, etc.

En este artículo vamos a dejar el tema de los sistemas inteligentes a lo que ya el lector conoce, con algunas precisiones y algo parecido haremos con la ética.

El aumento de la inteligencia es la capacidad de integrar nuestras facultades biológicas naturales de razonamiento con otras nuevas basadas en los avances culturales y tecnológicos.

Incluso algo tan fundamental como hablar y entender el lenguaje no es enteramente biológico. Desde el discurso, la escritura, la impresión y las comunicaciones electrónicas, el proceso ha sido largo e importante. Ha incrementado enormemente nuestra capacidad para recoger, analizar y actuar sobre la información.

El próximo salto cuántico en la mejora de las herramientas que nos permiten recoger datos sobre el mundo va a ocurrir a través de interfaces cerebro-máquina de varios tipos. Ahora se están probando y luego se aplican casi exclusivamente a personas con discapacidades, Pero al igual que ocurrió con el reconocimiento de voz por ordenador, estas interfaces se convertirán rápidamente en herramientas indispensables que todos queremos utilizar.

Cuando Elon Musk habla de la misión de Neuralink,² se refiere a la necesidad de mejorar nuestra capacidad, de aumentar nuestro ancho de banda de procesamiento de información, especialmente en un mundo cada vez más definido por la inteligencia artificial.

Cuando un mundo agitado te expone a grados de adaptación requeridos más allá de tus habilidades naturales, tienes dos opciones: rendirte, porque no eres capaz de superar esos límites o abrazar las adaptaciones a través de medios tecnológicos. A través de estos medios tecnológicos, puede extender el rango de su adaptabilidad al mundo cuya aceleración está aumentando. Como siempre se trata de un tema de fines y medios. No debes sentirte privado de tu capacidad de participar activamente sin ningún poder, puedes sentirte como un protagonista en el mundo del mañana, pero debes medir cuales son los riesgos y si moralmente estás de acuerdo.

¹ <https://dle.rae.es/pensar>. Personalmente no creo que se pueda decir de una máquina que piensa.

² *Neuralink de Elon Musk: el último avance del multimillonario empresario en su plan de conectar nuestros cerebros a computadoras*, 20 de agosto 2020 <https://www.bbc.com/mundo/noticias-53955394>.

La Inteligencia artificial se ha puesto de moda desde el punto de vista teórico y desde el punto de vista práctico y nadie quiere quedar fuera de ella. Ello supone algunos beneficios en el sentido que habiendo mucha producción hemos mejorado mucho nuestro conocimiento del tema dado que todos quieren hacer IA, pero también alguna dificultad sea porque se llama IA a cosas no lo son, y con este afán de llenar todo con IA impide que algunos casos que pueden resolverse en modo más sencillo se lo haga de manera complicada.

Las modas traen complicaciones. Como todos la quieren, muchos productores que no la tienen visten de IA lo que tienen y lo venden. Actualmente, por esa razón, se asocia a la inteligencia artificial productos que, interesantes en sus campos no tienen nada que ver con la IA. El primero es el blockchain. Más allá del uso siempre más frecuente de "blockchain" en forma de "buzzword" desde un punto de vista técnico se refiere a registros compartidos y distribuidos o "privados/centralizados" los registros están constituidos por nudos validados por técnicas de hash y criptografía. Una de las propiedades notables es la no modificabilidad de los "elementos/transacciones que forman parte de la cadena de bloques del registro. Que no se puedan modificar porque está de por medio el hash es algo así como el descubrimiento del agua tibia.

Las únicas reales diferencias entre las varias tecnologías para construir un blockchain se pueden resumir en el hecho que la distribución de los nudos (miners) es pública y libremente distribuida o privada/centralizada/controlada. O sea, cualquiera puede ser el minero o solo algunos. Obviamente se pueden "imprimir" monedas virtuales, notarizar contratos y transacciones o validar una bella producción musical o un film porno o cadenas de procesos y/o contratos entre empresas o dentro de una empresa

El haber podido superar (dicen) la amenaza de la computación cuántica y su aplicación en prácticamente cualquier actividad humana lo hacen muy atractivo. El verdadero límite de esta tecnología es el enorme dispendio computacional legado a la validación de bloques pues más se alarga la cadena, más se gasta.

Dicho esto, las competencias comunes, sobre todo de tipo criptográfico en un especialista en tecnología blockchain y uno de IA son iguales a cero.

Otro tanto puede decirse de una tecnología que se hará muy famosa próximamente que es el 5G, tecnología de redes que poco tiene que ver con la I.A.

Y que podemos decir de la ética La ética es una parte de la filosofía. Para hacer filosofía se requiere capacidad de asombro³.ética proviene el griego ethos, εθος que significa carácter, costumbre. En este sentido, la ética es considerada como una teoría del carácter y de las costumbres.

Detengámonos un momento en esta capacidad de asombro: no es sorpresa, ésta puede tenerla cualquier animal. El asombro en el diccionario de la Real Academia ⁴ es gran admiración y extrañeza y aun susto y espanto. Extrañarse es ponerse fuera y hay algo de eso en el asombro: se pone uno fuera, fuera de todo lo que le es familiar y por eso lo del susto o espanto. Y esta es una capacidad solo de la especie humana que le permite ir más allá de lo conocido de lo familiar y exponerse fuera a lo desconocido.

³Por eso recomiendo a quienes quieran adentrarse en la historia de la filosofía el libro de Jostein Gardener, *El mundo de Sofía*, Siruela, Madrid, 2008.

⁴ <https://dle.rae.es/asombro>

Es interesante notar que asombro trae también susto y espanto y esta es una característica de lo que en psicología se llama trauma que se define como un evento que amenaza profundamente el bienestar o la vida de un individuo. El trauma, la ruptura aparecen cuando uno abandona los conocimientos familiares y se aventura, exponiéndose, a saberes nuevos.

Y filosofía viene de amor a lo que no se conoce, pero se quiere tener, uno quiere lo que no tiene y quiere tener: conocimiento. Pero el conocimiento es peligroso y de allí también que pueda producir susto o espanto desde el asombro. No olvidemos que el árbol prohibido en el Paraíso era el del conocimiento⁵. Por eso la ética que es una actividad filosófica supone que estemos extrañados, asombrados, fuera de los conocimientos familiares, expuestos.

La pandemia ha traído muchas cosas, no todas malas. Pero el distanciamiento social es ciertamente entre los perniciosos, ya que el hombre es un ser "naturalmente político". El distanciamiento físico evita muchas efusiones afectivas y distancias mucho más. En esta distancia nace una nueva ética: no diferente y adversa a la tradicional, sino más precisa, más detallada y que prevé nuevas circunstancias. En una ética de deberes *prima facie* las nuevas circunstancias cambian los deberes. También la difusión de las redes es un nuevo elemento que debe ser considerado en lo social y en las costumbres.

Byung-Chul Han: "Nosotros, felices esclavos de la pandemia digital" En una entrevista concedida en estos días el filósofo sudcoreano afirma que estamos tan atrapados en la trama de las redes sociales y la IA que estamos contentos en la Pandemia porque tenemos Internet⁶. "La Red nos controla y nos guía como si fuera una dictadura". Describe nuestra "sociedad de supervivencia" en la que aquellos que no pueden permitirse el lujo de aislarse corren un mayor riesgo de enfermarse y donde uno está obsesionado con salvar su vida, olvidando lo que hace que valga la pena vivir. El miedo crece, el estado de emergencia es visto como la nueva normalidad y los llamados "hombres fuertes" que se inspiran en modelos autocráticos como el chino son los que se benefician de ello. China corre ahora el riesgo de convertirse en un modelo de régimen de vigilancia impulsado por una falta de humanidad alimentada por nuestro miedo a morir

2. Inteligencia artificial y riesgo

Lo cierto es que la inteligencia artificial llena casi todas las actividades humanas y comienza a mostrar sus riesgos.

Voy a obviar los riesgos que algunos autores atribuyen a la IA, como ser que sus algoritmos rápidamente sobrepasaran (si no lo han hecho ya) a la inteligencia humana porque, en general son debidos a miedos⁷. El robot provoca miedo y evoca miedos ancestrales. Desde el *golem* de la tradición hebrea, que aparece citado una sola vez en el Antiguo Testamento, esto

⁵ El árbol del conocimiento del bien y el mal(en hebreo, עֵץ הַדַּעַת טוֹב וָרָע *Etz ha-da'at tov va-ra*) en la historia del Jardín del Edén, en Génesis 2-3

⁶ https://rep.repubblica.it/pwa/robinson/2020/10/30/news/byung-chul_han_noi_schiavi_felici_della_pandemia_digitale_-272446064/?ref=nl-rep-a-bgr

⁷ Los seres humanos tenemos miedos ancestrales. Hay 6 que son universales: a la muerte, a la falta de autonomía, al hambre, a la soledad, a las pandemias, a la mutilación física o psíquica. También a la muerte

es, en salmo 139, verso 16, allí se dice “inconcluso o esbozado me vieron tus ojos. Tus ojos vieron mi embrión, y en tu libro se escribieron todos los días que *me* fueron dados, cuando *no existía* ni uno solo de ellos”. Ese embrión, según la interpretación, es el *golem*, pero para otros Adán. En hebreo, idioma original del Antiguo Testamento, se coloca el origen de la figura del *golem*, de donde la *Qabbal* (tradición) busca sus significados⁸.

En los orígenes hay una conexión etimológica entre Adán, creado por Dios y la tierra, concebida como metonimia en su componente de polvo. Adán, antes de que llegue su creador, el perfeccionamiento espiritual que lo vuelve una figura animada y dotada de intelecto, debidamente humano, es Adamah, tierra, *golem*. *Golem (golmi)* es una forma adjetival que en la Biblia aparece una sola vez en el salmo citado⁹.

De allí que las lecturas del *Midrash* cumplen la conjugación entre el aspecto humano (y paradigmático) y el aspecto no humano, inorgánico y material.

El estado golémico se atribuye al mismo Adán. Es importante percibir esta diferencia: en el estado orgánico, primitivo, hay materia, pero todavía no está el hombre, este va a venir luego, insuflado por el hálito divino. Y esta misma diferencia se cree ver en los robots, que son materia (que en algunos casos puede tener forma humanoide), pero le falta la acción que aparece luego a través de programas especiales, entre los cuales existen también los programas de aprendizaje. Y eso hace temer a quienes no conocen la robótica y tampoco los límites de IA que ese robot poco a poco se independice y llegue a ser más inteligente que el ser humano y termine dominándolo¹⁰.

Indico un libro como representante de esta tendencia el de Yuval Noha Harari *21 lecciones para el siglo XXI*,¹¹ el autor es un historiador israelí que ha escrito por lo menos dos libros que han sido *best sellers*, pero de Inteligencia artificial no sabe mucho.

El Golem es un monstruo creado por el Rabi Levi en Praga y de allí hay mucha literatura y cine dedicad a él. A veces es un monstruo, a veces sirve para salvar a los judíos del gheto. Además del Golem hay otros monstruos como Frankenstein. Todos seres inventados.

Los signos son improntas que modelan la realidad a su imagen. El Golem es construido y vuelto vital del correcto uso de las letras (otot) del Nombre Divino. El análisis de los signos puede hacerse con un código y esta palabra aúna informáticos, juristas, lógicos, biólogos y lingüistas

Del Golem al ciborg, ajedrecista, tramite el expediente de reconducir la variada tipología de vida antropomórfica y artificial a la “mera” fabricación de autómatas: el robot indica solo la máquina que exonera del trabajo manual (trabajar, en checo *roboti*).

⁸ Dentro de estos cuatro métodos de comprensión de la Torá, existen un sin número de posibles vías de entendimiento. Por ejemplo: Hay muchas maneras diferentes de entender la Torá según Peshat. Es por eso que hay muchos comentaristas de la Torá que se concentran en Peshat (Rashi, Iben Ezra Rashbam y muchos más) y ellos estarán muy a menudo (al parecer, la mayoría) de acuerdo sobre el significado de un verso. De hecho, según la Cábala hay 600.000 formas de entender Peshat

⁹ Algunas de las referencias pueden verse en el excelente libro *Dal Golem al Cyborgs*, de Bárbara Henry, Salomone Belforte, Livorno, 2016

¹⁰ Hay en ciencia ficción un relato de unos robots muy sofisticados que van perfeccionando un modelo y, cuando llegan a la Tierra, satisfechos por su perfección lo dejan: han creado al hombre.

¹¹ Yuval Naoah Harai, *21 lecciones para el siglo XXI*, 2018, Penguin Random House Grupo Editorial, Buenos Aires

El maridaje entre orgánico e inorgánico aparecen en las discusiones socialmente y culturalmente acreditadas en el pasado de Occidente antes de Mary Shelley (autora de Frankenstein) y las inquietudes góticas del romanticismo de Byron.¹²

Sí, vale la pena ocuparse de los peligros serios que han aparecido ya y que están siendo tratados; podemos llamarlos riesgos y son debidos al hecho que se trata de una metodología nueva a la cual la sociedad no ha tenido tiempo de habituarse. Empecemos por el uso de armas autónomas: Las armas programadas para matar suponen un grave riesgo en el futuro de la IA. No solo son peligrosas por que puedan convertirse en autónomas completamente, cosa que veo poco probable pero sí que se la utilice sin mucha supervisión para hacerlas más eficientes. Son una gran tentación para cualquier vigía que quiere dominar el espacio circundante todo el tiempo y puede poner el arma en automático, lo que quiere decir que actúa sin control humano.

La manipulación de la sociedad ya apareció sea en el caso de las redes sociales pueden ser una gran fuente para obtener información sobre cualquier persona, y además de utilizarse para el marketing y brindar anuncios específicos para cada persona, como el caso de Facebook. Con el escándalo de Cambridge Analytica en relación con las elecciones presidenciales de EEUU y el Brexit en Reino Unido, se pudo ver el enorme poder que puede dar tener datos para manipular a las personas, con lo cual la IA al ser capaz de identificar algoritmos y datos personales, puede ser extremadamente peligrosa.¹³

La invasión de la privacidad para oprimir socialmente: Es posible ‘seguir’ las pistas de un usuario en la Red, y utilizar mucha información invadiendo su privacidad. Por ejemplo, en China se utilizará información como el reconocimiento facial de cámaras y la forma en la que se comportan, si fuman, o si ven muchos videojuegos para el sistema de crédito social¹⁴.

Divergencia entre nuestros objetivos y los de la producción de IA: Si nuestros objetivos no coinciden con los de la máquina, las acciones que pedimos llevar a cabo pueden terminar en un desastre. Por ejemplo, mandar una orden a la IA de llevarnos a un lugar lo más rápido posible, pero sin especificar que tiene que respetar las reglas de tránsito para no hacer peligrar las vidas humanas. O cualquier otra acción humana que comporta riesgos y que los humanos deberíamos calcular cuando hacemos el cálculo general de la acción.

Discriminación: Como las máquinas de IA pueden recoger tu información, analizarla y rastrearla, también pueden utilizar esta información en tu contra. Por ejemplo, una compañía de seguros puede negarte el seguro debido a las veces que las cámaras tienen imágenes de ti hablando por teléfono, o una persona que aspire a un puesto de trabajo puede perder la oportunidad debido a su baja red social de contactos en Internet. Sucede además que los humanos tenemos sesgos (prejuicios) y que los pasamos a la maquinas sin darnos cuenta. Un peligro extraordinario consiste en el aprendizaje de los programas de IA. Lo primero que hacemos es enseñarle a recoger toda la documentación existente sobre el tema (que no es solo

¹² Algunas de estas observaciones pude comprobarlas en el excelente libro *Dal Golem al Cyborgs*, de Bárbara Henry, Salomone Belforte, Livorno, 2016

¹³La resolución de la Cámara de representantes norteamericana del 7 de octubre de este año, con sus 415 páginas es una fuente valedera del riesgo y de lo que razonablemente harán las autoridades.

¹⁴Va dicho que el sistema chino es una forma autoritaria de gobierno, una feroz dictadura. Sobre dictadura ver Antonio A. Martino, *Osservazioni sulla definizione di "dittatura"*, Il Politico, XLII, n° 2, Bologna, 1978

escrita, sino fotos y videos) y aprender de allí. Pero como los humanos nos comportamos discriminatoriamente los programas aprenden a discriminar.

Dicho brevemente: si un invento humano es importante necesariamente será también peligroso, casi podría decirse que para medir la importancia de una invención humana hay que medir su peligrosidad. La desintegración del átomo y la fuerza que produce e Internet podrían decirse las dos invenciones más importantes del siglo XX, bueno son las más peligrosas. Un tema que le sigue es ¿podemos pensar una sociedad sin energía atómica? O ¿podemos pensar una sociedad sin Internet? Creo que la respuesta es negativa lo que comporta asumir los riesgos de la dos.

Los sistemas inteligentes no van a volver atrás, por lo tanto, la solución no es demonizarlos como si no existieran o fuéramos a desterrarlos sino ver cuáles son los riesgos que crean, diferenciarlos pues algunos son menores y otros mayores, colocarlo en los justos contextos, estudiar las medidas para contrarrestarlos y tener un ojo vigilante a una sociedad siempre más compleja, siempre más cruzada por todas las líneas de las nuevas tecnologías.

3. El ser humano

Para Aristóteles el hombre es un animal naturalmente social *ανθρωπος ζoon φυσει πολιτιχον*¹⁵. Estando con los otros hay maneras correctas de comportarse e incorrectas. Esa es la ética.

Creo que todos sabemos cómo comportarse éticamente y la mejor prueba es una cola. Es una regla social que nos dice cómo comportarnos Nos permite juzgar el comportamiento de los demás Y esperamos que otros en las mismas condiciones hagan lo mismo. Basta enunciar su nombre para convencer a alguien que está haciendo lo incorrecto: “esto es una cola”.

Pero no es tan fácil saber en cada caso cual es el comportamiento ético o moral (moral viene del latín *mores*, costumbre). Hay muchas discusiones sobre la diferencia entre moral y ética, en general no me convencen y por razones prácticas aquí las vamos a considerar sinónimos.

Justamente porque no es fácil distinguir el comportamiento moral hay muchas teorías éticas.

La primera distinción que conviene hacer es entre ética y metaética: mientras la primera se ocupa de las reglas morales la segunda se plantea el tema de saber qué tipo de enunciados son los éticos y aquí hay dos grandes corrientes metaéticas: los cognitivistas, quienes sostienen que los enunciados éticos son susceptibles de ser verdaderos o falsos y por lo tanto puede demostrarse una verdad ética y quienes sostienen que ello es imposible, los no cognitivistas metas éticos. El que no se crea que un enunciado ético sea demostrable empíricamente no quiere decir que haya una fe menor en cumplirlo. Simplemente no se lo considera demostrable por verdad y falsedad.

¹⁵ El hombre es un animal naturalmente político

Casi todas las ideas éticas que provienen de la religión son cognitivistas metas éticos porque sostienen que la obligatoriedad deriva del hecho que algún Dios lo ordenó y como creen, consideran que es empíricamente demostrable.

Entre las teorías éticas hay algunas que han conseguido una gran difusión. Por ejemplo, la teoría kantiana del imperativo categórico.

Dice Emanuel Kant en *Fundamentación de la metafísica de las costumbres* «*Obra sólo según aquella máxima por la cual puedas querer que al mismo tiempo se convierta en ley universal. Obra como si la máxima de tu acción pudiera convertirse por tu voluntad en una ley universal de la naturaleza*» (AA IV: 421).¹⁶ Y lo considera un imperativo categórico porque es una norma que no tiene ningún criterio de atenuación desde el punto de vista de las condiciones que se verifican en el momento de aplicarlo.

Obviamente es muy rígido y muy duro y muchos autores han tratado de desmentirlo o de falsificarlo. El caso más interesante es el del Marqués de Sade con su Filosofía del tocador.¹⁷ Vale la pena leerlo porque en la introducción el autor dice explícitamente que ha escrito ese libro para falsificar la teoría kantiana e iniciar a los franceses en la cultura del libertinaje. Pero en el transcurso de la obra, donde tratan e inculcar a Eugenia a divertirse y considerar a los demás como objetos o medios, nunca como fines, uno de sus maestros confiesa que ha sido muy grata la enseñanza porque vio que Eugenia gozaba. Pero eso, depravado o no, es contrario a los postulados del autor y reafirma la teoría kantiana de tratar a los demás como fines, nunca como medios: si el maestro estaba contento porque Eugenia gozaba se tuerce la finalidad primordial de pervertir, tratando al otro como un medio y renace la idea que el otro ha obtenido un beneficio y por lo tanto fue tratado como un fin y no un medio

Si alguien preguntase como resumir rápidamente las posiciones más famosas en materia ética creo que podría hacerse una breve lista que incluyera Sócrates, Platón, Aristóteles, el cristianismo, Kant, K. Marx, Pragmatismo, Utilitarismo y el Existencialismo

Entre los que han tratado de aliviar la dura ley kantiana hay muchos y Jean Jacob Rousseau, lo intenta en la Nueva Eloísa¹⁸.

El siglo pasado se caracterizó por el progreso científico, tecnológico y médico, entre otros, pero también por una inusitada crueldad que se tradujo en la pérdida de millones de vidas y una falta de libertad por culpa de numerosas dictaduras.

Los movimientos políticos de dicha época presumían de argumentar desde la lógica de la razón y se escudaban en interpretaciones interesadas de intelectuales como Friedrich Nietzsche o Martin Heidegger. Tanto el leninismo, como el fascismo y nazismo se aferraron a pensadores como estructura legitimadora, pero no son autores de una política destructiva. Hubo una lectura sesgada. En realidad, se puede sacar citas de todo.

¹⁶Emanuel Kant, *Fundamentación de la metafísica de las costumbres*, No Books, Buenos Aires, 2003, versión original *Grundlegung zur Metaphysik der Sitten*, Riga, 1785.

¹⁷Francois de Sade, *La filosofía nel boudoir ovvero i precettori immorali. Dialoghi per l'educazione delle fanciulle*, Garzanti, Milano 2004. Título original *La Philosophie dans le boudoir ou Les instituteurs immoraux* , Paris, 1795.

¹⁸ Jean Jacob Rousseau, *La nouvelle Eloise*, Paris, 1761 *Ce n'est pas assez, croyez-moi, que la vertu soit la base de votre conduite, si vous n'établissez cette base même sur un fondement inébranlable*".

Con la obsesión por el racionalismo y la falta de sensibilidad, el valor de la vida humana se depreció hasta tal punto que algunas personas dejaron de ser consideradas como ciudadanas. Este fenómeno derivó en la creación de guetos, campos de concentración y gulags. Esa deshumanización se traduce en un 'tú no eres de los nuestros' y, en su forma más radical, 'tú no eres humano'. Es vital reclamar la dignidad de las relaciones interpersonales.

Un pensador que ha envenenado la política y por consecuencia la moral es Karl Smith, creador de la disyuntiva “amigo – enemigo”.¹⁹ Estado es definido como el estatus político de un pueblo organizado sobre un territorio delimitado. En la época actual, el Estado se encuentra entremezclado con la sociedad, al punto que lo político no admite ser definido a partir de lo estatal. Por eso, para llegar a una definición de lo político se requiere el descubrimiento y la fijación de una distinción específica a la cual sea posible referir las acciones y los motivos políticos. Esa distinción es la distinción de «amigo y enemigo», la cual determina la esencia de lo político, y desde este criterio se puede determinar lo político y lo apolítico. En la medida que no es derivable de otros criterios, ella corresponde, para la política, a los criterios relativamente autónomos de otros ámbitos humanos: bueno y malo para la moral, bello y feo para la estética, etc.²⁰

Gracias a esa distinción aparecen “ellos” y “nosotros” y en sistemas populistas sirve para eliminar toda otra opción que no sea con el régimen o contra él, sin dejar lugar a otras elecciones. Se es fascista o disfattista, comunista o anticomunista, castrista o contrera.

En el pasado siglo también apareció un autor inglés Charles Ross quien en su libro Fundamentos de Ética dice que tenemos un deber *prima facie* de ayudar a los demás, otro de mantener nuestras promesas, otro de devolver los actos de amabilidad anteriores y otro de no defraudar a las personas que confían en nosotros. Y habla de deberes *prima facie* porque nuestros deberes cambian con la información que obtengamos.²¹ En una cola lo elemental es respetar el orden de llegada, pero si aparece un señor con problemas cardíacos, nuestro deber cambia y es él quien debe ocupar el primer lugar en la fila. Algo parecido sucede con la llegada de una señora embarazada.²² La idea de los deberes *prima facie* es muy útil en nuestro tiempo donde permanentemente se van cambiando las condiciones del mundo en que vivimos: hay nuevos datos, hay nuevos deberes morales.

4. La llegada de los sistemas inteligentes

Con la llegada de los sistemas inteligentes hay situaciones nuevas. Se dice que en los últimos dos años se han hecho tantos avances como en los 18 anteriores y estamos hablando solo de este siglo. La velocidad del cambio tecnológico no tiene parangón en la historia de la humanidad y esto obliga a pensar una moral mucho más prolija y dedicada a cada una de las nuevas situaciones en las cuales se ha puesto a los humanos. No es que haya cambiado la moral, ha cambiado y mucho, el mundo.

¹⁹ Norberto Bobbio se carteaba con Karl Smith desde que lo había conocido en Alemania en 1933. Una vez me pregunto qué me parecía y le dije que estaba en desacuerdo a lo que él retrucó “el liberalismo obliga a escuchar a todos” Solo ahora me doy cuenta que no tenía razón porque el “todos” no puede incluir a gente que obra de mala fe.

²⁰ Karl Smith, *El concepto de lo político*, Madrid, Alianza Editorial, 2005. La obra original es de 1932.

²¹ David Ross, *Fundamentos de Ética*, Eudeba, Buenos Aires, 2003. El original inglés es de 1930

²² Una cosa curiosa es que los deberes *prima facie* de Ross dieron lugar a especulaciones de lógica difusa en campo de las ciencias duras. Y esto es debido a que todas las lógicas para consistentes disminuyen el valor de la negación. Lo opuesto a 1 no necesariamente es 0, lo opuesto a verdadero no necesariamente es falso sino que hay alternativas intermedias.

El progreso tecnológico y la masa extraordinaria de datos a disposición impulsó gobiernos, académicos y la sociedad civil a promover una reflexión sobre los desafíos éticos ligados a la gran progresión de sistemas inteligentes y comenzaron a surgir directivas y acuerdos.

Tenemos en claro los datos y la tecnología son herramientas utilizadas por los Estados para automatizar, predecir, identificar, monitorear, detectar, atacar y castigar. Actualmente son tantísimos, pero se destaca cómo deberían utilizarse principalmente para apoyar la protección social y los sistemas de bienestar y garantizar el bienestar.

4.1 Un caso importante

El domingo 18 de marzo de 2018, a las 21.59 hs en Tempe, Arizona, Elaine Herzberg fue embestida y muerta por un auto autónomo de prueba, de la firma Uber, conducido por los sistemas desarrollados para ese fin, con una conductora de seguridad llamada Rafaela Vásquez.

El automóvil disponía al momento del accidente, de cámaras con visión de toda la periferia, múltiples sensores, GPS, ubicación en mapas electrónicos, radar para la determinación de la velocidad relativa de cada objeto en movimiento, del LIDAR (Light Detection and Ranging) que permite una visión completa 300m de distancia y a 360°, con imágenes tridimensionales con muy alta definición debido a la utilización de láser. Todo esto controlado y comandado por un sistema inteligente de conducción.

La tecnología recibía los distintos datos y los procesaba con sus algoritmos, calculando continuamente y elaborando predicciones de las distintas alternativas de ocurrencia en los siguientes pocos segundos, en base a los objetos, personas, señalización y reglas de tránsito vigentes. El auto se desplazaba a 69 Km/hora, el sistema de detección de imágenes percibe 6 segundos antes de la colisión un objeto que no identifica en el medio de la calzada. Elaine Herzberg vestía un abrigo negro y llevaba a pie su bicicleta cargada.

El sistema detecta, un objeto no reconocido, luego interpreta un auto y finalmente una bicicleta, todo esto en concordancia con nuevas y sucesivas imágenes. Los sistemas de identificación necesitan varios miles de muestras de imágenes de un determinado objeto para aprender a identificarlo²³.

El algoritmo encuentra un objeto que desconocía, porque no estaba comprendido dentro los que habían aprendido, Estamos entonces en presencia de lo que en inteligencia artificial se denomina un “falso positivo”, es decir no identifica (falso) un objeto persona (verdadero). Los sistemas de reconocimiento de imagen expresan el resultado como una probabilidad que sea un objeto determinado, por ejemplo 95% casi certeza o un dudoso 15%.²⁴ Cuatro segundos después de la detección de un objeto no identificado, el sistema decide frenar para evitar el daño posible

²³Un niño con unas pocas fotografías puede detectar enseguida una bicicleta. Es la diferencia entre un humano y una maquina: el primero se ubica en un ambiente y reconoce cosas porque puede ponderar, la maquina puede ejercer muchos razonamientos muy rápidamente, pero carece de ponderación. Para decirlo en modo simple: la maquina puede ser racional, nunca razonable.

²⁴El programador debe evaluar cuál es el límite razonable de aceptación de la identificación, el caso del auto autónomo si el nivel de reconocimiento de la imagen es muy alto, cualquier objeto no reconocidos hará que el sistema frecuentemente frene y hasta se detenga, según las circunstancias, haciendo que el vehículo realice una conducción torpe y brusca.

en el caso que la consistencia del objeto sea contundente, esto ocurre a 1,3 segundos de la colisión, la velocidad al momento del impacto fue de 63 Km/hora.

Los sistemas inteligentes disponen de cantidad de información muy superior a la que tiene el ser humano, pero también tienen sus limitaciones por lo que exigen cuantificar los riesgos sin olvidarnos que se trata de pruebas encaradas para aprender de la realidad y corregir. En los dos años transcurridos desde el accidente, muchos de estos problemas fueron resueltos.

Se pudo verificar luego que la conductora de seguridad Rafaela Vásquez iba mirando un programa de TV en su celular, y recién levanta la vista en el último segundo antes del impacto, de acuerdo con la cámara interior del vehículo. Cuando hacen la autopsia de la persona fallecida, Elaine Herzberg, determinaron que su sangre contenía cristal y marihuana, además de una larga historia de drogas en su vida. El Estado de Arizona había permitido en el 2015 la circulación de autos autónomos de prueba, sin mayores indicaciones, es decir ni siquiera mencionaban la necesidad de un conductor de seguridad. Por esa razón las empresas se establecieron en ese Estado como centro de desarrollo y pruebas. Al momento del accidente la empresa Uber contaba con 400 autos autónomos en prueba circulando en Tempe. Un dato: el Gobernador del Estado de Arizona había manifestado antes del accidente, que las empresas podían tomar como empleados ex convictos. Rafaela Vásquez lo era porque había sido condenada a 5 años de prisión por intento de robo. Uber se encontraba en una situación complicada por una disputa con Google que le costó 245 millones de dólares.

Pido disculpas por el exceso de datos, pero son interesantes para mostrar como un cumulo de circunstancias muy difíciles de conocer con anterioridad pueden dar un rasgo fundamental a un hecho que luego puede ser considerado como disparador de reflexiones éticas que no existían hasta ese momento.

A partir de ese hecho comenzaron las preguntas que tal vez habría que haberse hecho antes del experimento sobre la eticidad de comportamiento y volvió al ruedo un viejo dilema que Filippa Foot había enunciado en los años sesenta del siglo pasado: el dilema del tranvía. Se escribieron libros como *Moral Machine* y hasta el famoso MIT (Instituto Tecnológico de Massachusetts) compuso una serie de 15 juegos sobre el famoso dilema que se puede ver (y jugar) en <https://www.moralmachine.net/>.²⁵

4.2. La cosa se pone más peligrosa

Los vehículos aéreos no tripulados (UAV) son ahora una herramienta central de antiterrorismo. Más específicamente, los vehículos aéreos no tripulados armados se han demostrado altamente valiosos para el ejército de los Estados Unidos al llevar a cabo asesinatos selectivos en regiones remotas del Oriente Medio y el Subsahara, África.

Los progresos recientes y en curso en la esfera de la inteligencia artificial (IA) abren la puerta a varias nuevas aplicaciones militares, particularmente en la asistencia del proceso de toma de decisiones de los operadores de vehículos aéreos no tripulados militares que llevan a

²⁵ Hay muchas críticas y escándalos a la manera de jugar de *moral machine* pero se olvida: 1. Que es un juego 2. Muchas de las críticas vienen de la visión meta ética de quien la formula, por ejemplo, pensando que los enunciados éticos pueden ser demostrados como verdaderos o falsos.

cabo asesinatos selectivos a distancia. Entre otras aplicaciones, el aprendizaje automático, a través de las redes neuronales y el aprendizaje profundo, puede ayudar a la identificación (y quizás incluso la selección) de los objetivos.

Las armas autónomas letales tienen algoritmos cognitivos que permiten identificar, seleccionar y eliminar un objetivo (humano o no) sin necesidad, o más bien con escasa, presencia humana. Drones, misiles de crucero, bombarderos y tanques están ya dotados de sensores, cámaras y radares que permiten el control remoto.²⁶

Va dicho que varios países, comenzando por EEUU tienen un sistema de espía permanente a través de satélites y que, apoyándose en estructuras autorizadas expresamente en el suelo de aliados, como Australia, le permiten verificar qué sucede en casi todo el mundo a un nivel de acercamiento digno de películas²⁷

En los Estados Unidos, desde 2001 existe una autorización para el uso de la fuerza militar dirigida desde espionaje satelital. Las convenciones que rigen el uso de la fuerza empezando por la de Ginebra de 1949, no regulan las diversas aplicaciones militares que con medio de IA podría llevar a cabo en la lucha contra el terrorismo. Se sabe un arma o un sistema de armas hoy se usa con fines antiterroristas, mañana en una guerra contra otro país, en algún momento como sistema de represión nacional.

Y estos usos son importantes en teatros de operaciones lejanos donde solo puede llegarse con medios modernos de observación y de transporte de armas, así como armas mismas que pueden operar sin necesidad de tripulación a distancias enormes y sin el riesgo de vidas amigas. Esto depende de la precisión de la observación y del ataque sucesivo.

De particular atractivo para los EE.UU., así como otras naciones que recurren a los UAVs, es la capacidad de "actualizar la violencia a distancia"²⁸ puede matar sin poner las botas en el suelo. Una operación clave.

La ventaja de esta realidad es que las vidas de los operadores de UAV no están en peligro de extinción.

Hay factores principales explican por qué los vehículos aéreos no tripulados están particularmente favorecidos cuando se trata de ataques dirigidos. En primer lugar, ya que los operadores encuentran se alejan del campo de batalla convencional y no tienen que temer por sus vidas, pueden flotar los UAVs sobre un lugar específico hasta que su objetivo previsto es donde tiene que estar para un ataque letal exitoso.²⁹

Lo importante es que la falta de un piloto humano en la cabina permite más la precisión de tales ataques porque un UAV puede ser volado a una altitud menor que su equivalente tripulado. En altitudes más bajas, los pilotos convencionales cada vez más vulnerables a los

²⁶Por ejemplo el Phalanx (MK15) US- Navy

²⁷De hecho hay una serie que habla del tema: Pine Gap de Netflix.

²⁸(17Junio de 2013), en línea: Institución Brookings <www.brookings.edu/articles/why-drones-work-el caso de las armas de elección de los lavaderos/> [perma.cc/Z9YC-W7FP].

²⁹Daniel L Byman, "Por qué funcionan los drones: El caso del arma elegida por Washington",

sistemas de defensa tierra-aire³⁰, lo que significa que deben volar rápidamente sobre su objetivo para evitar ser disparados hacia abajo si deciden volar bajo o permanecer a una mayor altitud y más lejos de su objetivo. En ambos casos, un ataque realizado por un hombre es probable que el avión sea menos preciso que el de un UAV

En segundo lugar, no tener un piloto en la cabina también significa que los UAVs pueden permanecer en el aire por largo tiempo períodos de tiempo en comparación en comparación con los aviones tripulados que tienen que aterrizar cuando la resistencia de sus pilotos se ha agotado. Por ejemplo, uno de los pilotos de los EE.UU., los UAVs más usados - el Predator - pueden volar hasta 35 horas sin tener que tomar tierra.³¹ Esto permite a múltiples operadores tomar turnos consecutivos y asegurar que un objetivo no se pierda porque un piloto necesite dormir. Esto también asegura una presencia casi permanente en el cielo. Tercero, el láser contemporáneo. Municiones guiadas, con las que se equipan los vehículos aéreos no tripulados armados, ahora permiten que un misil golpee su objetivo dentro de un margen de cinco pies de radio de la base de un régimen de gobierno para la proyección de la fuerza a través de Al...de los vehículos aéreos no tripulados asistidos.

Previendo estos riesgos que aumentan con las armas no tripuladas se han realizado diversas conferencias como la de Buenos Aires en 2015 donde se publicó una carta abierta de problemas éticos ligados al uso bélico de la IA, en 2016 se creó un grupo de expertos gubernamentales sobre el tema y en 2019 se han activado conversaciones diplomáticas sobre el tema.

La CCW es la conferencia internacional sobre la revisión de armas convencionales para tratar los riesgos y efectos traumáticos excesivos que las nuevas armas ponen sobre la población civil y regularla a través de un derecho internacional humanitario. Se está tratando de aprobar un alegato adicional a la CCW, convocar una conferencia de expertos y aprobar una Robot Law o ley sobre el uso ético de sistemas inteligentes.

El art. 36 del Protocolo adicional a los acuerdos de Ginebra obliga a los estados a revisar si el uso de armas que producen o adquieren respetan las reglas del derecho internacional. Se empieza a considerar que las decisiones de herir, matar o traumatizar seres humanos no debe ser aprendido por sistemas autónomos bélicos.

Se va afianzando la idea que en la ciberguerra también el control humano significativo sea la piedra angular del todo el desarrollo.

Un nuevo programa de investigación de DARPA está desarrollando interfaces cerebro-computadora que podrían controlar "enjambres de zánganos, operando a la velocidad del pensamiento". ¿Y si tiene éxito?

¿Cómo las sospechas de espionaje amenazan a la ciencia transfronteriza?

³⁰James Der Derian, *Virtuous War: Mapping the Militar-Industrial-Media-Entertainment-Network* (Nueva York, NY: Routledge, 2009) en xxxi.

³¹Alcides Eduardo dos Reis Peron, "The "Surgical" Legitimacy of Drone Strikes? Issues

Una empresa de inteligencia advierte que China está explotando a los científicos cuánticos occidentales con fines militares. Las pruebas son escasas, pero las tensiones están aumentando. Siendo un sistema político opaco es difícil saber exactamente qué están haciendo y donde quieren llegar.

El ejército de los EE.UU. no tiene parangón en su capacidad de proyectar el poder en todo el mundo, y eso no va a cambiar.

Toda tecnología que cambia el juego es absorbida por los militares y la IA no es una excepción. Hasta cierto punto, máquinas como los drones utilizados en la guerra ya tienen una gran autonomía. Y eso plantea serias cuestiones éticas.

Esta semana, dice Weekend Reads, estamos explorando la inteligencia artificial militar y cómo se está utilizando tanto ahora como en el futuro. Y nos preguntamos: ¿qué pueden o deben hacer los investigadores para evitar que su trabajo sea convertido en un arma?³²

En su reciente libro David Kelcullen³³ pregunta cómo y qué han aprendido los oponentes de Occidente durante el último cuarto de siglo de conflicto. Aplicando una combinación de teoría evolutiva y detallada observación de campo, explica lo que pasó con las "serpientes" - amenazas no estatales incluyendo terroristas y guerrilleros - y los "dragones" - competidores estatales como Rusia y China. Explora cómo los enemigos aprenden en condiciones de conflicto y examina cómo el dominio occidental sobre una forma de guerra. Los agentes no estatales pueden ahora acceder a niveles de precisión y sistemas de armas letales que antes sólo estaban a disposición de los gobiernos.

El tema se lo planteaban ya en los años 70 en Alemania con las brigadas rojas y la banda Baader Meinhof: ¿era una nueva forma de guerra? Sí, pero desde las hordas bárbaras, pasando por Güemes, Joseph Broz, Tito que escribe un libro de guerra de guerrilla, Ernesto Guevara que teoriza sobre la guerrilla urbana, el tema no era nuevo.

Si los costos son bajos, cualquier grupo puede armarse de ciberguerra y hacer terrorismo. El ciberespacio es un dominio, tanto físico como virtual, muy complejo. Allí se desarrollan actividades de creación, procesamiento, almacenamiento, intercambio y visualización de información usando software y hardware interconectado. Lo constituyen tanto la Internet como todas aquellas redes aisladas que se utilizan con finalidades particulares.

A diferencia de los dominios terrestre, aeroespacial y marítimo -donde se pueden desarrollar operaciones militares- el ciberespacio, es de características artificiales. Por ejemplo, en los espacios tradicionales, aún si se interrumpiera el suministro de energía eléctrica o nadie estuviera en esos lugares, ese ambiente seguiría existiendo. Hemos consultado un militar que se ocupa del tema y sostuvo que los ciberataques son prácticamente diartios. ³⁴

³² La Weekend Reads , revista militar especializada, publica en su número del 31 de octubre 2020 *Military artificial intelligence* , donde hay artículos sobre el tema desde el miedo a futuros terroristas con sistemas inteligentes hasta el programa del ejército norteamericano de uso de la lectura (o averiguación) de la mente.

³³ David Kelcullen, *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One*, Chicago, 2016 y en el más reciente . *The Dragons and the Snakes: How the Rest Learned to Fight the West*, 2020,

³⁴ Aníbal Luis Intini, jefe del Comando de Ciberdefensa del Estado Mayor Conjunto argentino

Cambian las armas que se pueden usar y las despreocupaciones morales en aras de un resultado. *Nec novum sub sole*. Lo que tal vez falte y es encajar todo esto en el tema general de la guerra y sus grandes tratados: Carl von Clausewitz, Sun Tzu, Robert McNamara, para citar 3 excepcionales.

4.3 dando la cara

Los métodos de reconocimiento facial dedicados a vigilancia masiva tuvieron que esperar hasta la segunda década del siglo XX, en parte porque los primeros intentos mostraron que los resultados estaban lejos de ser los deseados. La primera vez que se utilizó en un acontecimiento de grandes proporciones fue en la Super Bowl de 2002, que resultó, en general, un fracaso. Numerosos falsos positivos mostraron que la tecnología no estaba aún preparada para las grandes multitudes. A partir del 2015, las fuerzas policiales del Reino Unido empezaron a probarla también para eventos públicos en vivo, pero un informe realizado por Big Brother Watch encontró que estos métodos seguían devolviendo resultados hasta un 98% imprecisos.

Una de las trabas para que funcione correctamente entre grandes multitudes puede ser dificultad de obtener una imagen de calidad. El Sistema de Reconocimiento Facial de la INTERPOL (IFRS) almacena imágenes faciales enviadas por más de 160 países, lo que la convierte en una base de datos única en el ámbito policial. Este sistema, puesto en marcha a finales de 2016 ha logrado identificar a más de 650 delincuentes, prófugos y desaparecidos. Pero su web ya advierte que la calidad de las imágenes es un aspecto esencial y que aquellas que solo posean una resolución media o baja pueden no conseguir o influir negativamente en la precisión de la búsqueda.

En dicha página se especifica que “lo ideal sería disponer de una fotografía de pasaporte conforme a la norma OACI, ya que se trata de una imagen frontal completa de la persona con iluminación homogénea en el rostro y un fondo neutro”.

Pero fuera del ámbito de la seguridad, ya vimos cómo la inteligencia artificial aplicada a este ámbito asomaba en nuestras redes sociales como algo inocente y original allá por el 2010, cuando Facebook lo incorporó para reconocer las caras de nuestros amigos en las fotografías que subíamos y las etiquetas que le proporcionábamos. Su uso se extendió rápidamente y hoy se encuentran en muchos de los smartphones y aplicaciones que utilizamos a diario.

Las polémicas sobre el uso de la biometría se han alzado principalmente al descubrirse que algunas empresas y organismos utilizaban la información recopilada para otro propósito diferente del autorizado. Este asunto ha estado en el centro del debate sobre ética y privacidad desde el 2001, propiciado a su vez por un vacío legal en la aplicación de las nuevas tecnologías de la información.

En los últimos meses, las controversias sobre su uso han aumentado, impulsadas en parte por el movimiento Black Lives Matter³⁵, por lo que diferentes organismos han empezado a recular acusando excesivo control y posible promoción del racismo y la injusticia social. Amazon o IBM dieron un paso atrás. San Francisco ya se había convertido el pasado año en la primera gran ciudad de EE UU en prohibir a todas las agencias locales, entre ellas la

³⁵Movimiento surgido en EEUU frente a abusos contra la comunidad afroamericana y que significa literalmente “la vida negra importa”.

policía, el uso de técnicas de reconocimiento facial, por miedo a que se las utilice para discriminar.

4.4 Las noticias falsas

El cuestionamiento del *ethos* comunicacional e informativo —o sea, de las normas, principios y valores que fundamentan las formas de comunicación y la información en una determinada sociedad— gira actualmente en torno a temas tan variados como la privacidad, la propiedad intelectual, el acceso libre al conocimiento, el derecho a la expresión en las redes digitales, la censura, las nuevas definiciones de género, la identidad digital, las comunidades digitales, el plagio digital, la sobrecarga informacional, la brecha digital y el control social digital

El filósofo italiano Luciano Floridi distingue entre una «ética de la comunicación global» (*ethics of global communication*) y una «ética de la información global» (*global information ethics*)³⁶ la primera se refiere al diálogo y las prácticas consensuales en la interacción entre diversas culturas y generaciones. Pero tratándose de una postura meramente pragmática deja de lado, según Floridi, preguntas más básicas como, por ejemplo: ¿cuáles son en este caso los principios éticos que se toman como base de dicho diálogo y en qué se fundamenta dicha elección? o ¿existe una «macroética» en el sentido de un cierto tipo de consecuencialismo o deontologismo o contractualismo?.³⁷

Estas preguntas son las que Floridi intenta responder con lo que él llama «ética de la información global». Todo aquello que soporta la segunda. Ética multicultural y una sociedad viva en las redes que trata de evitar el sufrimiento, así como toda forma de destrucción de los entes, puesto que todo ente, por el hecho mismo de ser, tiene derecho a una forma específica de respeto. *Ens et bonum convertuntur*, afirmaba el adagio escolástico. Floridi llama a esta ontología mínima «ontocéntrica», siendo aparentemente más radical que, por ejemplo, la «biocéntrica» y la «antropocéntrica».

La ontología ontocéntrica coloca en su centro no al agente sino al paciente de la acción, incluyendo los entes no vivientes, extensión que constituye una tesis muy loable y de gran relevancia ecológica actual, pero de difícil aplicación en los casos concretos. Floridi se opone tanto a una teoría metafísica que afirme algo respecto del ser de los entes —lo que sería una forma de «imperialismo ontológico»— como a un mero relativismo que es incapaz de promover una interacción eficaz a nivel global en referencia a problemas que afectan a todas las culturas. Esta «ética de la información global» no quiere imponer una jerarquía de valores comunes, sino permitir que estos se expresen en situaciones concretas con toda su situacionalidad.

Floridi opta por una ontología ligera y horizontal (*lite, horizontal ontology*) como condición de posibilidad de interacción pragmática entre culturas, las cuales, vistas en su densidad vertical o gruesa (*thick cultures*), son a menudo irreconciliables. La diferencia entre culturas ligeras y gruesas puede relacionarse con la distinción del filósofo estadounidense

³⁶Luciano Floridi, *Sextus Empiricus, The Recovery and Transmission of Pyrrhonism*. Oxford: Oxford University Press, 2002.

³⁷Luciano Floridi, *The Blackwell Guide to the Philosophy of Computing and Information*. (a cura di) Oxford: Blackwell, 2003.

Michael Walzer entre argumentos éticos finos (thin) y gruesos (thick), según se los analice o no en su profundidad cultural³⁸

Esta posición de Floridi es muy atractiva y útil a nivel pragmático, pero tiene el peligro de ignorar la necesidad de un análisis intercultural grueso o vertical.

Según el ético holandés Philip Brey, un diálogo ético intercultural tiene que tomar en serio las diferencias interculturales. Brey usa el concepto de ética de la información en el sentido ordinario de cuestiones ético³⁹. Desafíos teóricos y prácticos de la ética intercultural de la información y casos relacionadas con las TIC pero incluyendo también la ética de la computación (computer ethics), así como la de los medios de comunicación de masas (media ethics) y el periodismo.

Hay que distinguir entre un relativismo moral descriptivo y uno normativo llamado también meta ético. Este último, aplicado a la ética de la información, encara la pregunta de si existen conceptos y principios con validez universal o bien si la ética de la información es culturalmente relativa. Brey defiende la necesidad de un relativismo descriptivo en la ética de la información, dado que, si estas diferencias no existen, las discusiones en torno al relativismo se volverían abstractas.

La forma en la que se utilizan los sistemas cerrados no está contribuyendo a proporcionar una plataforma multicultural para el debate, sino que más bien permite que se restablezcan las viejas morales —más locales— de comunicación entre individuos. Por otro lado, los sistemas abiertos como The Well (y en cierta medida, Messenger de MSN), permiten que la gente sea más libre de discutir los temas que ellos mismos escojan, a su modo y con la profundidad que decidan.

En cualquier caso, nadie puede entender cómo serán utilizados realmente dichos sistemas por los usuarios finales, ya que —como se argumentará más adelante— su uso dependerá de factores involucran a un individuo o a un grupo muy limitado de personas, hablamos de ética privada. Un ejemplo de ética a este nivel es el aborto, cuando se debe tomar una decisión en torno a la supervivencia de un individuo en particular.

La responsabilidad corresponde únicamente a ese individuo que deliberadamente decide dentro de ese dilema; el escrutinio sobre dicha deliberación es realmente muy problemático, ya que son aquellos que analizan el dilema desde fuera quienes deben ponerse en los zapatos del otro individuo y conocer cada uno de los hechos y las contingencias emocionales en las que el otro se encuentra. Por otro lado, la dimensión pública de la ética involucra más a muchas personas y no solo grupos de ellas. De hecho, las decisiones públicas tienen un impacto generalizado no solo para las generaciones presentes, sino muy probablemente para generaciones futuras también.

Las decisiones públicas dependen de valores generales mediante los cuales queremos, por así decirlo, diseñar una buena sociedad. La discusión en torno a estas decisiones es (p. ej. Facebook) son instancias de lo que Karl Popper llamó «sociedades abstractas». De acuerdo con Popper: podríamos concebir una sociedad en la cual las personas prácticamente nunca se

³⁸Michael Walzer *Pensar políticamente*, Barcelona, Editorial Paidós, 2010, ISBN 978-84-493-2381-2

³⁹P Brey, JH Søraker - *Philosophy of technology and engineering sciences*, 2009 - Elsevier

conozcan cara a cara, donde todas las transacciones de negocios sean llevadas a cabo por individuos aislados que se comunican a través de escritura mecanografiada o a través de telegramas, y que se desplazan en autos motorizados ⁴⁰.

Hay aspectos moralmente negativos muy fuertes en torno a esto, pero Popper ve también los beneficios: por ejemplo, la posibilidad de nuevas relaciones personales, independientes del lugar de nacimiento o de su ubicación.

Popper entendía que las sociedades modernas son en muchos aspectos sociedades abstractas; sin embargo, la realización más próxima a dichas sociedades es posible solo a través de Internet. Hay una diferencia fundamental entre los Sistemas Medios por Computadora, tales como The Well y Facebook. La primera es un tipo de sistema construido de abajo hacia arriba, una sociedad cerrada se caracteriza por una jerarquía rígida inmutable, de hecho.

En ese mismo sentido las redes sociales diseñadas de arriba hacia abajo (Facebook y Twitter) son como las sociedades cerradas: al parecer no tienen reglas compartidas, que tienen las mismas funciones que los tabúes que son incuestionables puesto que son impuestos por quienes diseñaron el sistema. Sus valores fundamentales, basados en el credo dominante del fundamentalismo de mercado y el interés personal como bendición moral se refuerzan con la red: no es posible una apertura hacia otras culturas o individuos.

Las sociedades abiertas son exactamente lo contrario a las sociedades cerradas; son sociedades sensibles y tolerantes, donde los mecanismos y procedimientos se dice que son transparentes, flexibles y siempre negociables. Los países occidentales son sociedades abiertas, Facebook es una herramienta exitosa para el cambio, pero peligrosa por la dependencia que crea.

Con la promesa de libre consumo, Facebook ha ayudado a países cerrados y subdesarrollados —como los del norte de África— a deshacerse de sociedades jerárquicas y pasadas de moda mediante el cambio de una sociedad tribal a una basada en el individuo. Esa es la principal razón por la cual en países desarrollados y corruptos —aunque sean sociedades abiertas— Facebook no ha funcionado; estas ya eran sociedades individualistas, líquidas (en el sentido explicado por el filósofo social Zygmunt Bauman,⁴¹ en las cuales el código moral está dominado por los valores del mercado. En estas sociedades, únicamente herramientas comunitarias desarrolladas de abajo hacia arriba —como The Well— tienen más probabilidades de funcionar. Internet, en otras palabras, nos libera de los impedimentos de las sociedades a las que pertenecemos, es posible que sea por esta misma razón que podamos escapar de la influencia nefasta de este mismo ideal.

Con el fin de obtener una comprensión cultural y moral real entre las personas de todo el mundo, necesitamos un cambio comunitario, posible gracias a sistemas como The Well, para que ayude cambiar realmente los sistemas de comunicación enfocados en el individuo por sistemas de colaboración.

⁴⁰Karl Popper, *Conjeturas y Refutaciones: el Crecimiento del Conocimiento Científico*: Revisada por el autor en 1972, Ediciones Paidós, Barcelona, 1981, Título original: *Conjectures and Refutations. The Growth of Scientific Knowledge* Publicado en inglés por Routledge & Kegan Paul, Londres, 1972 (4.º edición)

⁴¹Zygmunt Bauman *Modernidad líquida*. México: Fondo de Cultura Económica, 2003.

Es interesante la resolución de octubre del 2020 de las autoridades escolares de Cataluña erradicando de las escuelas el uso de Facebook y otras redes sociales por no tener suficiente control de las noticias que en ellos aparecen. Y esto rige ya para el año lectivo en curso.

4.5 ¿Y el trabajo?

Un ejemplo concreto de conflicto entre las visiones económicas y éticas: la muy debatida cuestión de la pérdida de puestos de trabajo inducida por la automatización, incluida la IA. Un aspecto crucial, especialmente en el futuro, que puede conducir a un aumento sustancial de las desigualdades, a menos que se adopten medidas, principalmente políticas, pero no sólo, para evitarlo.

"¿Es correcto introducir nuevas tecnologías que conduzcan a la pérdida de puestos de trabajo?", pregunta el representante de la Oficina Nacional de Investigación Económica de Estados Unidos. Y trata de responder: "los economistas podrían estar tentados a responder inmediatamente 'sí' a mi pregunta. Pueden observar que en un mercado que funciona bien, los salarios reflejan perfectamente el valor social del trabajo; si a un nivel determinado de salarios, una empresa encuentra conveniente innovar de manera que ahorre costes de mano de obra, libera mano de obra para emplearla en otras actividades más útiles para la sociedad.

En la práctica, si un determinado trabajo ya no es necesario, se vuelve obsoleto, pierde su utilidad y, por lo tanto, su valor, por lo que es mejor no hacerlo más y cambiarlo. Después de todo, siempre ha funcionado así. "Pero también hay que tener en cuenta que la pérdida de puestos de trabajo es socialmente más cara de lo que sugiere una visión eficiente del mercado".

No es sólo el mercado el que puede decidir

En esencia, a menudo vale la pena que los empresarios y las empresas innoven (también) en términos de costos, mucho menos para los trabajadores que se quedan en casa, pero más generalmente para la sociedad en su conjunto, que tiene que intervenir con servicios públicos y dinero.

En los últimos decenios, aunque el crecimiento económico general ha continuado, la distribución de los beneficios económicos ha sido cada vez más desigual. En los Estados Unidos, la mitad inferior de la población, compuesta principalmente por trabajadores poco cualificados, apenas ha experimentado un aumento de los ingresos después de la inflación. Durante el mismo período, los ingresos reales de los ricos se duplicaron, los de los megamillonarios se triplicaron y los de los multimillonarios se cuadruplicaron.

A pesar de las diferentes tesis desde ambas perspectivas, los economistas y los éticos de la innovación pueden estar de acuerdo en varios puntos. En primer lugar, es conveniente garantizar que los trabajadores que pierden su empleo reciban ayuda no sólo en términos monetarios, por ejemplo, con medidas contra el desempleo, sino también en términos del valor más amplio que la sociedad atribuye a sus pérdidas. Después de todo, estos trabajadores han sido sacrificados en aras de la eficiencia económica: cuando sus puestos de trabajo fueron trasladados o cancelados, fueron el daño colateral para permitir que la economía adoptara procesos de producción más eficientes. Esas ayudas no económicas comienzan con formación para los nuevos desocupados y su familia, el acceso a lugares de lectura y recreación, el acceso

a una alimentación sana y sobre todo la cobertura médica mínima para su salud y la de su familia.

4.6 Los periódicos

Este año el ganador del premio más interesante es en la categoría "innovación tecnológica al servicio del periodismo". Ganó el proyecto realizado por un periódico canadiense, The Globe and Mail, cuyo sitio web ha sido dirigido por un sistema inteligente para elegir los artículos durante algún tiempo y los lectores nunca se han quejado de ello; juran, de hecho, que ni siquiera lo notarían. Digámoslo mejor: la disposición del 99% del contenido del Globe and Mail la decide Sophi, un motor de inteligencia artificial que lo actualiza según lo que hacen los lectores en tiempo real. Puede parecer el fin del periodismo y la victoria de los ordenadores, pero sólo puede parecerlo en una lectura muy superficial.

Sophi no escribe nada, simplemente se ocupa de la gestión de contenidos, de la administración de contenidos; ha sido capacitada durante algunos años por los periodistas de Globe and Mail para comprender qué temas son pertinentes, qué temas no son adecuados para la primera página, cuándo un artículo no reciente merece ser sacado, e incluso qué mezcla debe haber en las fotos de cada pantalla para reflejar los diferentes grupos étnicos del Canadá utilizando un software de reconocimiento facial. Los periodistas no han contribuido a que este sistema sea despedido sino, por el contrario, a hacer un mejor producto y al mismo tiempo tener más tiempo para hacer encuestas o entrevistas o análisis de calidad. Obviamente también en Globe and Mail hay un verdadero editor, que se reserva el derecho de elegir sólo el puesto principal: su nombre es David Walmsley y cree que el trabajo de los periodistas es encontrar y escribir grandes historias, el resto lo pueden hacer los algoritmos.

Esta innovación no viene por casualidad. En 2015, cuando la crisis se agudizaba, el editor, en lugar de recortar o impulsar supuestos contenidos virales, creó un laboratorio de innovación, el Lab351, con la tarea de apoyar las mejores innovaciones propuestas por los propios periodistas; para ello, además del dinero para cada proyecto, contrató a ingenieros informáticos y analistas de datos. "Tienes que gastar dinero si quieres ganar dinero" es una de las frases que repiten en el Globe and Mail cuando cuentan su historia de éxito. Pero tienes que gastarlo bien. Incluso en el periodismo, la tecnología no es un enemigo, no es una alternativa a los humanos. Bien utilizada, la tecnología debe darnos tiempo para volver a ser humanos.

Obviamente Sophi está sujeta a todos los riesgos de sesgos que significa preparar un sistema inteligente para que actúe en tareas humanas, más en esta tan delicada de elegir las noticias y diferenciarlas en principales y accesorias. Ya hay un problema de noticias falsas (*fake news*). Los periodistas deben cubrir o descubrir cuáles son los sesgos que provoca la recolección de datos de preferencia sociales y decidir si deben ser corregidos y como.

5. Documentos poniendo límites

La difusión de los sistemas inteligentes ha sido percibida y ya países, asociaciones, entidades universales y regionales han comenzado a redactar y difundir documentos de alarma y orientación sobre la necesidad de límites éticos.

3.1. la Unión Europea fue la primera en plantear el problema de la inteligencia artificial no sólo desde el punto de vista ético (punto de vista que ya ha seguido el Reino Unido en 2016), sino también desde el punto de vista reglamentario, especialmente desde los mecanismos de imputación de la responsabilidad civil.

En su resolución de 16 de febrero de 2017, que contiene recomendaciones a la Comisión Europea sobre las normas de derecho civil relativas a la robótica, en las que también se tienen en cuenta los sistemas de inteligencia artificial, el Parlamento Europeo ya había indicado como temas de atención los relativos a las repercusiones éticas y sociales, subrayando que el desarrollo de la robótica y la inteligencia artificial debe tener por objeto integrar las capacidades humanas y no sustituirlas.

Esta resolución, si bien tiene por objeto proporcionar directrices en el ámbito civil, ofrece muchas ideas para comprender el alcance de los problemas jurídicos y sociales que la inteligencia artificial trae consigo.

Desde el punto de vista de la responsabilidad de los agentes inteligentes, el desafío surge cuando consideramos que el criterio tradicional de imputabilidad está vinculado a una conducta del agente sujeto ¿pone una inteligencia artificial su conducta autónoma en el mundo físico? Sin embargo, los sistemas de inteligencia artificial parecen recordar los principios de la responsabilidad del productor, que deben adaptarse al hecho de que esos sistemas asumen una capacidad de decisión autónoma, y podrían también implicar a sujetos distintos del productor únicamente, como los programadores o los que elaboran los algoritmos de decisión.

Otro criterio de imputación de responsabilidad podría encontrarse también en la culpa en la vigilancia de la persona que utiliza el sistema, cuando puede tener conocimiento de la adopción de decisiones erróneas por parte del sistema.

El documento del Parlamento Europeo enumera una serie de puntos de atención y también sugerencias para una correcta solución normativa de los mismos.

Desde un punto de vista jurídico, están apareciendo trabajos sobre algunos casos concretos: La IA y la competencia, la IA y los derechos de autor directos, la IA y las armas letales, la IA y los coches que se conducen solos, la IA y el mercado financiero, la IA y el reconocimiento facial, la IA y el mercado laboral, la IA y la asistencia sanitaria.

5.2 El Vaticano

El pasado 28 de febrero se firmó en el Vaticano un documento que enfatiza los aspectos éticos de la inteligencia artificial y que coloca otra pieza en la base de la regulación. Firmaron representantes de las grandes empresas de distribución de datos Amazon, Google, Facebook, Microsoft, Apple, representantes vaticanos, del gobierno italiano y de la Unión Europea. Concretamente firmaron Paola Pisano, Ministra de Innovación Tecnológica de Italia, Monseñor Vincenzo Paglia, Presidente de la Academia Pontificia para la Vida, el Director General de Fao Dongyu Qu y los presidentes de Microsoft e IBM. David Sassoli, Presidente del Parlamento Europeo, también estuvo presente en el evento⁴²

5.3. El Parlamento Europeo (PE),

⁴²https://www.wired.it/attualita/tech/2020/02/28/intelligenza-artificiale-etica-vaticano-microsoft-ibm/?refresh_ce=

Tras una propuesta de borrador (junio 2016) y un informe razonado (febrero 2017), el PE ha aprobado un informe sobre Robótica en el que se establece un Código Ético de Conducta⁴³.

La propuesta de resolución del PE establece que es necesario establecer “un marco de guía ético para el diseño, producción y uso de los robots” que sirva de complemento a las diferentes recomendaciones puramente legales que se realizan. Es decir, profundizar en una nueva disciplina que aborde los problemas éticos de la Inteligencia Artificial, la “roboética”.

La idea de fondo es que los estándares éticos deberían ir dirigidos a la humanidad –esto es, los diseñadores, productores y usuarios de los robots– y no tanto a los robots en sí mismos. Como indica la profesora Nathalie Nevejans, responsable del informe encargado por el propio PE, no hay que confundir la ética en la robótica con la ética en las máquinas, es decir, una ética que obligue a los propios robots a adherirse a reglas éticas. Existen varios principios fundamentales que han sido recogidos por la resolución que incluyen la protección de la dignidad humana, la privacidad, la libertad, la igualdad de acceso o los efectos sociales, entre otros.

Enumeración de los principios sancionados por el Parlamento Europeo:

Proteger a los humanos del daño causado por robots: la dignidad humana.

Respetar el rechazo a ser cuidado por un robot.

Proteger la libertad humana frente a los robots.

Proteger la privacidad y el uso de datos: especialmente cuando avancen los coches autónomos, los drones, los asistentes personales o los robots de seguridad.

Protección de la humanidad ante el riesgo de manipulación por parte de los robots: Especialmente en ciertos colectivos –ancianos, niños, dependientes– que puedan generar una empatía artificial.

Evitar la disolución de los lazos sociales haciendo que los robots monopolicen, en un cierto sentido, las relaciones de determinados grupos.

Igualdad de acceso al progreso en robótica: Al igual que la brecha digital, la brecha robótica puede ser esencial.

Restricción del acceso a tecnologías de mejora regulando la idea del transhumanismo y la búsqueda de mejoras físicas y/o mentales

5.4 La visión de los expertos: los 23 principios de la IA de Asilomar

En febrero de 1975 un grupo de genetistas se reunió en un pequeño pueblo de California, Asilomar, para decidir si su trabajo podría destruir el mundo. Estábamos al inicio

⁴³ P8_TA (2017)0051 Normas de Derecho civil sobre robótica Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

de la ingeniería genética y la manipulación del ADN, y de esa reunión surgieron una serie de principios y un estricto marco ético para la biotecnología.

Cuatro décadas después –organizado por el Future of Life Institute– otro grupo de científicos se reunió en el mismo lugar y con el mismo problema. Pero esta vez se trataba de analizar las posibles consecuencias de la Inteligencia Artificial. La idea de fondo fue clara y compartida: un profundo cambio está llegando y afectará a toda la sociedad y las personas que tengan algún tipo de responsabilidad en esta transición tienen tanto una gran responsabilidad como la oportunidad de darle la mejor forma posible, se plantean una serie de principios agrupados en tres consideraciones generales: a) Principios relativos a la investigación; b) Ética y valores; y c) Temas a largo plazo.

1) Meta de la investigación: el objetivo de la investigación de la IA no debería ser crear inteligencia sin dirigir, sino inteligencia beneficiosa.

2) Financiación de la investigación: la inversión en IA debería ir acompañada de fondos para investigar en asegurar su uso beneficioso, incluyendo cuestiones espinosas sobre ciencias de la computación, economía, legislación, ética y estudios sociales.

3) Enlace entre ciencia y política: debería haber un intercambio constructivo y sano entre los investigadores de IA y los legisladores.

4) Cultura de la investigación: una cultura de cooperación, confianza y transparencia debería ser fomentada entre los investigadores y desarrolladores de IA.

5) Evitar las carreras: los equipos que estén desarrollando sistemas de IA deberían cooperar activamente para evitar chapuzas en los estándares de seguridad.

6) Seguridad: los sistemas de IA deberían ser seguros a lo largo de su vida operativa, y verificables donde sea aplicable y posible.

7) Transparencia en los fallos: si un sistema de IA causa daño debería ser posible determinar por qué.

8) Transparencia judicial: cualquier intervención de un sistema autónomo en una decisión debería ir acompañada de una explicación satisfactoria y auditable por parte de una autoridad humana competente.

9) Responsabilidad: los diseñadores y desarrolladores de sistemas avanzados de IA son depositarios de las implicaciones morales de su uso, mal uso y acciones, con la responsabilidad y oportunidad de dar forma a dichas implicaciones.

10) Alineación de valores: los sistemas de IA altamente autónomos deberían ser diseñados para que sus metas y comportamientos puedan alinearse con los valores humanos a lo largo de sus operaciones.

11) Valores humanos: los sistemas de IA deberían ser diseñados y operados para que sean compatibles con los ideales de dignidad humana, derechos, libertades y diversidad cultural.

12) Privacidad personal: la gente debería tener el derecho de acceder, gestionar y controlar los datos que generan, dando a los sistemas de IA el poder de analizar y utilizar esa información.

13) Libertad y privacidad: la aplicación de la IA a los datos personales no puede restringir de forma poco razonable la libertad, real o sentida, de las personas.

14) Beneficio compartido: las tecnologías de IA deberían beneficiar y fortalecer a tanta gente como sea posible.

15) Prosperidad compartida: la prosperidad económica creada por la IA debería ser compartida ampliamente, para el beneficio de toda la Humanidad.

16) Control humano: los seres humanos deberían escoger cómo y si delegan decisiones a los sistemas de IA para completar objetivos escogidos previamente.

17) Sin subversión: el poder conferido por el control de sistemas de IA altamente avanzados debería respetar y mejorar, más que subvertir, los procesos sociales y cívicos de los que depende la salud de la sociedad.

18) Carrera armamentística: debería ser evitada cualquier carrera armamentística de armas autónomas letales.

19) Capacidad de precaución: al no haber consenso, deberíamos evitar las asunciones sobre los límites superiores de las futuras capacidades de la IA.

20) Importancia: la IA avanzada podría representar un profundo cambio en la historia de la vida en la Tierra, y debería ser planificada y gestionada con el cuidado y los recursos adecuados.

21) Riesgos: los riesgos asociados a los sistemas de IA, especialmente los catastróficos o existenciales, deben estar sujetos a planificación y esfuerzos de mitigación equiparables a su impacto esperado.

22) Automejora recursiva: los sistemas de IA diseñados para auto mejorarse recursivamente o auto replicarse de una forma que pudiera llevar al rápido incremento en su calidad o cantidad deben estar sujetos a unas estrictas medidas de control y seguridad.

5.5. La OCDE

La OCDE y los países socios han adoptado formalmente hoy el primer conjunto de directrices de políticas intergubernamentales sobre Inteligencia Artificial (IA), y convenido en someterse a unas normas internacionales que velen por que el diseño de los sistemas de IA los haga robustos, seguros, imparciales y fiables.

Los 36 países miembros de la OCDE, junto con Argentina, Brasil, Colombia, Costa Rica, Perú y Rumanía han suscrito hoy en París los **Principios de la OCDE sobre la Inteligencia Artificial** en el marco de la Reunión del Consejo de Ministros de la Organización, con el lema “*La transición digital al servicio del desarrollo sostenible*”.

Los Principios, elaborados a partir de las orientaciones proporcionadas por un grupo de expertos integrado por más de 50 miembros procedentes de gobiernos, instituciones académicas, el mundo empresarial, la sociedad civil, organismos internacionales, la comunidad tecnológica y sindicatos, comprenden cinco principios basados en valores para el despliegue responsable de una IA fiable y cinco recomendaciones en materia de políticas públicas y cooperación internacional. Su objetivo es guiar a los gobiernos, organizaciones e individuos para que, en el diseño y la gestión de los sistemas de IA, prioricen los intereses de las personas, así como garantizar que quienes diseñen y gestionen sistemas de IA respondan de su correcto funcionamiento.

«La inteligencia artificial está revolucionando nuestra forma de vivir y trabajar, y ofrece unas ventajas extraordinarias a nuestras sociedades y economías. Ahora bien, también plantea nuevos desafíos y siembra incertidumbre y preocupaciones de carácter ético. Compete, por tanto, a los gobiernos asegurarse de que el diseño de los sistemas de IA respete nuestros valores y leyes, de forma que las personas puedan confiar en que su seguridad y privacidad serán objeto de una consideración prioritaria», manifestó el Secretario General de la OCDE, Angel Gurría. «Estos Principios constituirán un referente global para una IA confiable, de modo que las oportunidades que brinda redunden en los mejores resultados para todos», terminó de presentarlos⁴⁴

5.6 La Unesco

El documento que más impresiona es el de Unesco que luego de proponer un borrador al cual se podía intervenir hasta el 31 de julio de 2020 se prevé una revisión por zonas geográficas que en dos años determinará primer instrumento normativo mundial sobre la ética en la inteligencia artificial en forma de recomendación, de conformidad con la decisión adoptada por la Conferencia General de la UNESCO en su 40º período de sesiones, en noviembre de 2019.

Expertos de 155 países, sociedad civil (mediante una encuesta mundial en línea), organismos de las Naciones Unidas, grandes empresas mundiales, como Google, Facebook y Microsoft, así como el sector académico, desde la Universidad de Stanford hasta la Academia de Ciencias de China, pudieron compartir sus impresiones y enriquecer las conclusiones del proyecto.

Este proyecto de recomendación se ha enviado recientemente a los 193 Estados miembros de la UNESCO y será objeto de una serie de negociaciones con miras a su adopción definitiva en la Conferencia General de la UNESCO en noviembre de 2021.⁴⁵

La estructura del documento es de gran interés porque establece por un lado los valores que deben ser protegidos, los principios que deben ser seguidos y las políticas concretas que se pueden adoptar para lograr los fines perseguidos. Esta estructura se va a ver reflejada en otros instrumentos que se han dado posteriormente, como el de Colombia.

5.7 Libro blanco de Inteligencia artificial de la UE

⁴⁴ <https://www.oecd.org/centrodemexico/medios/cuarentaydospaísesadoptanlosprincipiosdelaocdesobreinteligenciaartificial.htm#:~:text=Cuarenta%20y%20dos%20pa%C3%ADses%20adoptan%20los%20Principios%20de%20la%20OCDE,Par%C3%ADs%20mayo%202022%20de%202019.&text=%C2%ABLa%20inteligencia%20artificial%20est%C3%A1%20revolucionando.a%20nuestras%20sociedades%20y%20econom%C3%ADas>

⁴⁵ <https://es.unesco.org/news/unesco-da-gran-paso-primer-instrumento-normativo-etica-ia>

Lleva como subtítulo “sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza” es de febrero de 2020.

Dice “la Comisión respalda un enfoque basado en la regulación y en la inversión, que tiene el doble objetivo de promover la adopción de la inteligencia artificial y de abordar los riesgos vinculados a determinados usos de esta nueva tecnología. La finalidad del presente Libro Blanco es formular alternativas políticas para alcanzar estos objetivos; no aborda ni el desarrollo ni el uso de la inteligencia artificial para fines militares. La Comisión invita a los Estados miembros, a otras instituciones europeas y a todas las partes interesadas, como la industria, los interlocutores sociales, las organizaciones de la sociedad civil, los investigadores, el público general y demás personas con interés en la materia, a que presenten sus opiniones con respecto de las opciones que se muestran a continuación y a que contribuyan a la futura toma de decisiones de la Comisión en este ámbito.

Establece siete requisitos esenciales contemplados en las directrices del grupo de expertos de alto nivel, a saber: • acción y supervisión humanas; • solidez técnica y seguridad; • gestión de la privacidad y de los datos; • transparencia; • diversidad, no discriminación y equidad; • bienestar social y medioambiental; • rendición de cuentas.

Se ocupa de los riesgos que produce la IA y distingue entre riesgos bajos y altos riesgos, cosa muy elemental pero que vale la pena reivindicar: no es lo mismo usar un Sistema automático como el parquímetro para relevar una prolongación indebida del tiempo de un estacionamiento y la sentencia en un juicio penal con muchos imputados, varios delitos y particulares circunstancias.

Se hace cargo de un tema común entre informáticos y organizadores y casi desconocido para el gran público: los prejuicios (sesgos) que acompañan la toma de decisiones y como ellos se transmiten a los sistemas inteligentes que creamos ⁴⁶

5.8 Marco ético para la Inteligencia artificial en Colombia.

Es interesante ver lo que establece el marco del país latinoamericano sobre el tema: habla de riesgos del uso de sistemas inteligentes, principios que son más o menos los que ya hemos expuesto y herramientas entre las cuales aparecen auditoria de algoritmos, limpieza de datos, evaluación de la legitimidad, definición y gestión de riesgos, modelos de gobernanza para asegurar la ética en el uso de la I.A. No es particularmente original, pero lo ponemos a la atención porque es el primero en América Latina en ocuparse en forma sistemática del tema

⁴⁶Los prejuicios y la discriminación son riesgos inherentes a toda actividad social o económica. La toma de decisiones de las personas no es ajena al error ni a la subjetividad. No obstante, en el caso de la IA, esta misma subjetividad puede tener efectos mucho más amplios, y afectar y discriminar a numerosas personas sin que existan mecanismos como los de control social que rigen el comportamiento humano³⁵. Puede suceder también que el sistema de IA «aprenda» mientras está funcionando. En tales casos, cuando los resultados no puedan preverse ni anticiparse en la fase de diseño, los riesgos no se deberán a fallos en el diseño original del sistema, sino más bien a las repercusiones prácticas de las correlaciones o de los modelos que reconozca el sistema en un gran conjunto de datos.

6. Economía y ética

Siempre hay una justificación económica para un desastre social, pero los economistas serios no están de acuerdo. Lo que no se puede negar es que el índice Gini ha tocado los límites más altos.⁴⁷

Observando la miseria creada por la pérdida de puestos de trabajo, la ética puede verse tentada a responder 'no' a la pregunta formulada supra. Pueden ver los daños y sufrimientos tangibles infligidos a los trabajadores despedidos y observar que no es ético imponérselos a los trabajadores, quienes no pueden apreciar los efectos a largo plazo del progreso económico en el bienestar humano.

No sólo eso, cuando los mercados son tan perturbadores, puede ser socialmente indeseable que sean la única guía de la toma de decisiones humanas. Por eso hay que hacer compromisos entre la ética y el mercado.

Las pérdidas de puestos de trabajo causadas por la automatización suelen reducir la demanda de mano de obra y los ingresos de los trabajadores, y pueden aumentar las ganancias de los empresarios y los accionistas. Los resultados pueden generar una distribución más desigual de los ingresos, que la sociedad percibe como menos deseable. "No hay ninguna razón teórica para creer que el libre mercado dirigirá los esfuerzos de innovación hacia las innovaciones más deseables socialmente", señalan Zachary, Bthune y Korinek ⁴⁸

El primer teorema fundamental de la riqueza en la economía, comúnmente llamado "teorema de la mano invisible", establece que, bajo ciertas condiciones idealizadas, el mercado generará una distribución eficiente de los recursos existentes en la economía. Sin embargo, este teorema no se aplica al progreso tecnológico, por lo que el mercado puede impulsar la innovación en la dirección equivocada.

Si las innovaciones generan importantes perturbaciones sociales que la 'mano invisible' del mercado no puede manejar, entonces los innovadores tienen el deber moral de internalizar los efectos perturbadores que generan; las cuestiones de la eficiencia del sistema y la distribución de los ingresos no pueden separarse. Sin embargo, es demasiado común que los perdedores del progreso tecnológico sean abandonados a su suerte, o que reciban un apoyo mínimo de los ganadores o de la sociedad para compensar sus pérdidas.

Los economistas y los éticos de la innovación pueden estar de acuerdo en varios puntos. Vale la pena garantizar que los trabajadores que pierden su empleo reciban ayuda no sólo en términos monetarios, sino también con políticas contra el desempleo y atendiendo a las exigencias del desempleado y su familia que exceden el campo económico. Preservar a los trabajadores y sus familias es una forma de evitar daños colaterales del mejoramiento de la producción de bienes con nuevas tecnologías. El mejoramiento de la vida, en general, no puede, no debe traer aparejado exclusivos reducidos a la miseria. Es una película ya muchas veces vista que hay que evitar.

⁴⁷El coeficiente de Gini es una medida de la desigualdad ideada por el estadístico italiano [Corrado Gini](#). Normalmente se utiliza para medir la [desigualdad en los ingresos](#), dentro de un país, pero puede utilizarse para medir cualquier forma de distribución desigual.

⁴⁸Zachary A. Bethune & Anton Korinek, 2020. "Covid-19 Infection Externalities: Trading Off Lives vs. Livelihoods," NBER Working Papers 27009, National Bureau of Economic Research, Inc.

7. Siempre teniendo cuidado

Si las normas son demasiado débiles, existe un riesgo demasiado grande de que nuestros derechos y libertades se vean socavados: Esto se aplica actualmente a todas las aplicaciones de la inteligencia artificial, que hasta ahora sólo se han basado en principios y valores éticos no vinculantes. Con la nueva esta legislación, Europa tiene la posibilidad de adoptar un marco jurídico para la IA con normas claras. Necesitan instrumentos sólidos para proteger los derechos fundamentales y la democracia.

En los últimos años, los gobiernos de todo el mundo se han ocupado de crear comités, consejos y grupos de expertos especiales para debatir la ética de la inteligencia artificial: el Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial nombrado por la Comisión Europea, el Grupo de Expertos sobre la IA en la Sociedad de la Organización de Cooperación y Desarrollo Económicos (OCDE), el Consejo Asesor sobre el Uso Ético de la Inteligencia Artificial y los Datos de Singapur y el Comité Selecto sobre Inteligencia Artificial de la Cámara de los Lores del Reino Unido o, en los Estados Unidos, la administración Obama elaboró una Hoja de Ruta para la Política de la IA.

La ciudad de Nueva York, que fue una de las primeras en crear un consejo especial llamado "Grupo de Tareas del Sistema de Decisión Automatizada", tardó más de un año en acordar incluso una definición de la toma de decisiones automatizada. Hasta la fecha, todavía no ha logrado elaborar un panorama general de todos los sistemas de IA que se utilizan en la ciudad.

Al mismo tiempo, muchas empresas han aceptado con gratitud el libre paso a la autorregulación estableciendo juntas de ética, escribiendo directrices y patrocinando investigaciones en temas como el sesgo algorítmico o la "imparcialidad en la inteligencia artificial". No hace falta molestar a Marshall McLuhan para saber que el medio condiciona el mensaje, desde ese punto de vista no hay medios neutros.

El debate sobre el cuestionamiento moral a la acción de megaempresas que manejan el 75 % de los datos que soy muchos puede ser y ha sido usada como pantalla para distraer la atención del daño que están infringiendo a la sociedad y de los extraordinarios beneficios que obtienen con prácticas monopólicas.

En los últimos años, el sector empresarial ha estado construyendo una "lavadora de ética" tras otra. Facebook ha financiado a los especialistas en ética de la IA en la Universidad Técnica de Múnich en Alemania, mientras que Google, SAP y Microsoft han adoptado principios rectores y códigos éticos.

Sin cuestionar estas supuestas buenas intenciones, es obvio que el enfoque en los debates éticos ha sido desde hace mucho tiempo una manera eficiente para la industria de ganar tiempo y evitar una dura regulación. Con demasiada frecuencia, confiar en la ética y la autorregulación ha sido insuficiente para que las empresas no rindan cuentas, si no se dispone de una aplicación estricta y de mecanismos de supervisión independientes proporcionados por un marco institucional.

En los últimos años hemos visto abundantes pruebas que ponen de relieve el hecho de que nuestras leyes actuales son insuficientes para proteger contra la discriminación, cuando la carga de la prueba sigue recayendo en la víctima, que puede ni siquiera ser consciente de que está siendo discriminada por un algoritmo, y cuando la responsabilidad por los daños no siempre puede establecerse a lo largo de la compleja cadena de suministro de un sistema de inteligencia artificial.

Las personas que tienen más probabilidades de ser discriminadas por los sistemas de inteligencia artificial también tienen más probabilidades de no tener los medios financieros o la confianza en sí mismas para entablar una demanda de resultado incierto. Las herramientas automatizadas de toma de decisiones pueden a menudo exacerbar el racismo y el sexismo anclado en nuestras sociedades.⁴⁹

Piensen en el sistema COMPAS que se utilizó en los Estados Unidos para predecir si los acusados volverán a cometer delitos - se descubrió que discriminaba a los acusados negros. En un estudio reciente del Instituto Alan Turing de Londres y la Universidad de Oxford, un equipo de investigación demostró que las leyes actuales en Europa son insuficientes para proteger a las personas contra el daño causado por los algoritmos defectuosos.

Hay muchos ejemplos que ilustran que un enfoque de derecho no vinculante a través de la ética, la autorregulación y la responsabilidad social de las empresas fracasa drásticamente casi siempre. Por ejemplo, The Intercept reveló en 2018 que Google estaba desarrollando una versión censurada de su motor de búsqueda para el gobierno chino, en violación directa de sus principios éticos de IA.

No será fácil para los responsables políticos europeos redactar una ley que vaya al grano, que contenga medidas específicas, que siga el ritmo de la tecnología y que no cause daños colaterales a nadie.

Pero este desafío también es una oportunidad en Europa. Tienen la oportunidad de ser el primer continente en poner a la humanidad en el centro de las políticas digitales y tener en cuenta que poderosos intereses intentan impulsar la "agenda ética". Necesitamos derechos fuertes y ejecutables para los usuarios de los futuros sistemas de inteligencia artificial porque tienen una influencia cada vez mayor en todos los aspectos de las vidas de sus ciudadanos.

Por ello, el proceso legislativo debe ser lo más participativo e inclusivo posible. La Comisión de la UE acaba de celebrar una consulta pública con el fin de preparar su propuesta para el enfoque europeo de la IA. En teoría, los ciudadanos podrían responder, pero el proceso de consulta de la Comisión de la UE no es, lamentablemente, muy accesible, debido a que se le pide que cree una cuenta antes de poder responder a un cuestionario que contiene una larga lista de preguntas muy técnica

Por consiguiente, los encargados de la formulación de políticas de la UE deberían encontrar otros foros y hacer participar proactivamente a los grupos más afectados por la IA y con menos probabilidades de sentarse a la mesa donde se toman las decisiones. Esto también ayudaría a aumentar la transparencia y a crear confianza en la toma de decisiones de la UE.

⁴⁹Ver el informe de Juan G. Corvalán y Enzo María Le Fevre Cervini; *El sistema de IA discrimina las personas?* Hecho en el Laboratorio de Ialab , www.ialab.com.ar, IA confiable <https://ialab.com.ar/wp-content/uploads/2020/07/Evaluacion-IA-confiable-pdf>

Sólo así podrán asegurarnos de que adoptan normas que hagan aplicables derechos y libertades en virtud de los Tratados y la Carta, en lugar de adoptar simplemente una orientación ética en nombre de "principios y valores".

Es de señalar la existencia de una roboética, primero como disciplina universitaria y ahora prevista directamente por la Unesco. Es curioso, pero a las 3 leyes robóticas de Asimov se agregó una cuarta: "El robot debe siempre respetar la continuidad de la especie humana, aunque esto colisione con las tres reglas anteriores" y esto por dos razones: la primera es que algunos temen que la inteligencia artificial supere a la natural y la segunda porque, de cualquier modo, en un mundo de mucha tecnología es mejor defender y proteger a los seres humanos que vendrá, o sea defendamos la especie. Que es una ley de Darwin adaptada a los desarrollos tecnológicos.

Se van perfilando éticas consecuencialistas, basadas en resultados, que tienden a proteger el bienestar de los usuarios y éticas deontológicas basadas en los procesos de toma de decisiones donde el valor moral es el parámetro para decisiones individuales o colectivas que atañen el bienestar de los demás.

Una cuestión fundamental es si los procesos de decisiones se pueden trasladar en todo o en parte a sistemas algorítmicos que usan miríadas de datos. La idea que se va imponiendo es no limitar el uso de sistemas inteligentes sino calcular, detectar y perfeccionar el modo de prever los riesgos, encontrar formas universales de calificarlos y medirlos y mitigarlos a través de un uso serio y eficaz de sistemas de derechos humanos.

Los sistemas de inteligencia artificial deberían empoderar a los seres humanos y ser resistentes y seguros. Capaces de asegurar respeto total por la privacidad y la protección de datos, y mecanismos adecuados de gobierno de datos. Los modelos de negocio de datos, los sistemas expertos y la IA deben ser transparentes y garantizar diversidad, no discriminación y equidad, evitando el sesgo injusto. También es necesario que los sistemas de IA beneficien a todos los seres humanos

El desarrollo tecnológico ha dado un gran salto adelante en este siglo, pero no ha sido seguido por su desarrollo social, político e informativo. Las fantasías, por un lado, la ignorancia no inocente y los miedos atávicos, por el otro, están dando una considerable contaminación de información. Todas nuestras vidas se ven afectadas por la IA, desde los teléfonos móviles hasta los vuelos espaciales, pero cinco compañías poseen el 75% de todos los datos del planeta. Nos encontramos con nuevas situaciones y no tenemos ni ética ni cultura, ni siquiera normas para tratarlas. Y es urgente hacerlo.

8. Ética y estética

Siempre pensé que había una relación entre ética y estética pero siempre me resulto difícil explicar por qué. Recuerdo la anécdota que contaba Giménez de Azua, era que cuando el régimen había echado al profesor de estética el de ética dijo "entonces yo también debo irme"

Revisando el *Tractatus* de Wittgenstein en la proposición 6.421 dice “ética y estética son la misma cosa”⁵⁰. Este argumento tiene dos partes. La primera es negativa, y se desprende directamente de la operación de someter el lenguaje humano al canon de esa suerte de lenguaje ideal que se exhibe a través de la semántica que Wittgenstein pone en marcha en el *Tractatus*, la horma de su atomismo lógico, mientras que la segunda, de carácter afirmativo, es deudora de una distinción fundamental que subyace al tratamiento del complejo lingüístico, la distinción entre decir (*sagen*) y mostrar (*zeigen*).⁵¹

Con respecto a “trascendental” parece que Wittgenstein se remita a la idea de Kant en la *Crítica de la razón pura*, cuando ya en la Introducción dice textualmente: “Llamo trascendental todo conocimiento que se ocupa, no tanto de los objetos, como de nuestro modo de conocerlos, en cuanto que tal modo ha de ser posible a priori” (A12/ B25)

Pero no solo los filósofos del pasado nos ayudan en esta tarea, también algunos del presente como Marcus Gabriel, quien sostiene que si bien la reducción puramente cibernética de la mente humana es rechazable por muchas razones (por ejemplo, la tontería de que nuestro cerebro nos manipula y por tanto no hay libre albedrío), también lo es minimizar la importancia del cerebro para la mente. Basta con reparar en las devastadoras consecuencias de la carencia de empatía, del Alzheimer o de la transmisión hereditaria de patologías como la esquizofrenia.⁵²

“El arte es aquella super inteligencia temida por los críticos de las tecnologías digitales”⁵³ y defiende la de que hoy es el arte quien controla al poder y no al revés. Según él vivimos en una época estética, por lo que se encuentra presente en todas partes, como herramienta de persuasión o disfraz; fusionado con el diseño, en la publicidad y en las gráficas de las páginas web, películas, ropa.

El catalán Alex Hinojo formula críticas al tecno consumismo y dice que con la multiplicidad de fuentes lo que hemos logrados son las noticias falsas *fake news*, mientras que del lado de la producción se ha mantenido un nivel y una moral gracias al pedido de trazabilidad “La ética ha exigido trazabilidad a los sistemas de producción. Hoy, cualquier producto del supermercado incorpora un número de lote que nos permite, en caso de incidencia, detectar su origen. En este caso, la tecnología ha servido para mejorar la calidad de los sistemas de producción y de distribución, así como del servicio al consumidor”⁵⁴

En su libro *On the Existence of Digital Objects*⁵⁵ el chino Yuk Hui, se apoya en sus antecedentes como ingeniero en computación y programador para investigar entidades digitales como virus de computadora, videoclips, algoritmos y redes. En el prólogo de dicho

⁵⁰Ludwig Wittgenstein, *Tractatus logico-philosophicus* proposición 6.421 “Está claro que la ética no resulta expresable. a ética es trascendental. (Ética y estética son una y la misma cosa.)Ethik und Aesthetik sind Eins.

⁵¹ Ver también la cita de la carta a Ludwig von Fricker puesta en el acápite.

⁵²Markus Gabriel, *Yo no soy mi cerebro. Filosofía de la mente para el siglo XXI*/Barcelona: Pasado y Presente, 2016 ISBN 978-84-944950-7-6

⁵³Markus Gabriel, *El poder del arte*, Roneo, Buenos Aires, 2019.

⁵⁴Alex Hinojo *Hacia una nueva ética informativa*, CCCBLab, 16 de enero del 2020, <http://lab.cccb.org/es/hacia-una-nueva-etica-informativa/>

⁵⁵Yuk Hui, *On the Existence of Digital Objects*, (Volume 48) (Electronic Mediations), Minnesota, 2020

libro, el filósofo francés Bernard Stiegler describe el pensamiento de Hui como un “generoso y abierto entorno teórico para la exploración de la experiencia humana en conexión con la infoesfera”.

En este libro Hui, realiza un examen filosófico de los objetos digitales y su esquema de organización creando un diálogo entre Martin Heidegger y Gilbert Simondon, que Yuk Hui contextualiza dentro de la historia de la informática. ¿Cómo pueden entenderse los objetos digitales según la individualización y la individuación? persigue esta pregunta a través de la historia de la ontología y el estudio de los lenguajes de marcas y las ontologías de la Web; investiga la estructura existencial de los objetos digitales dentro de sus sistemas y entornos. Con este enfoque relacional hacia los objetos digitales y los sistemas técnicos, el libro aborda la alienación, descrita por Simondon como la consecuencia de ver erróneamente las técnicas en oposición a la cultura. Y llega a la cosmotécnica.

Y hablando de la cosmotécnica “Para los griegos, “cosmos” significa un mundo ordenado. Al mismo tiempo, el concepto apunta a lo que está más allá de la Tierra. La moralidad es ante todo lo que concierne al reino humano. La cosmotécnica, como yo la entiendo, consiste en la unificación del orden moral y el orden cósmico a través de actividades técnicas. Si comparamos Grecia y China en la antigüedad, descubrimos que tienen una comprensión muy diferente del cosmos y también concepciones muy diferentes de la moralidad. El arbitraje entre ellas tiene lugar de diferentes maneras con diferentes tecnologías. Una cosmotécnica del tipo de la tianxia no es posible en un tiempo que ya no tiene una concepción de “Cielo”, como la tuvo la gente del pasado. Como otras grandes naciones, China tiene satélites orbitando la tierra. Los cielos se han convertido en un lugar secular, utilizado por los humanos, y no pueden jugar un papel como poder legitimador moral”.⁵⁶

Hay relaciones entre la estética y sistemas inteligentes: artistas que, gracias a sus conocimientos matemáticos e informáticos y a su relación con los grandes centros de computación de Google, NVIDIA, etc., exploran las posibilidades estéticas y creativas que ofrece la IA e incluso crean sus propios algoritmos, produciendo obras originales, a menudo muy escénicas. Este es el caso de la *Historia latente* de Refik Anadol presentada en @fotografiska en Estocolmo, para la cual el artista ha reunido un millón de datos fotográficos de materiales de archivo: "A través de la exploración de las memorias fotográficas de los últimos 150 años, esta exposición pretende investigar y reimaginar la memoria colectiva, las capas ocultas de la historia y la conciencia de una ciudad [Estocolmo] que de otra manera podría ser reinterpretada mal o quedar a oscuras"⁵⁷.

La inteligencia artificial juega un papel crucial en el ecosistema cultural mundial. Aconseja lo que debemos ver, oír, leer y comprar. Decide cuántas personas verán nuestro contenido compartido. Nos ayuda a tomar decisiones estéticas cuando creamos contenido digital. En la producción cultural, la IA se ha utilizado durante mucho tiempo para producir trailers de películas, álbumes de música, artículos de moda, diseño de productos y de páginas web, arquitectura, etc. Lev Manovich es un creador y al mismo tiempo un crítico que se ocupa

⁵⁶Edwin Lo, *Entrevista: sobre Tecnodiversidad: una conversación con Yuk Hui*, 27 de julio de 2020 , Seminario de Tecnologías Filosóficas, <http://philosophyandtechnology.network/3939/entrevista-sobre-technodiversity-una-conversacion-con-yuk-hui/>.

⁵⁷Refik Anadol: "*Melting Memories*" esculturas digitales en movimiento, exposición en San Francisco, 1019.

de los límites que pone la estética a los sistemas inteligentes⁵⁸. Nos presenta un marco sistemático para ayudarnos a pensar en los usos culturales actuales y futuros de la IA. Y nos ofrece la posibilidad analítica de entender los medios, el diseño y la estética en la era de los sistemas inteligentes.

El arte se está planteando temas de conocimiento y como las nuevas cogniciones llevan a otra estética o por lo menos a una estética que tiene en cuenta los sistemas inteligentes y va observando sus límites. Algo que hace Gianluca Consoli y escribe sobre estas experiencias ⁵⁹.

La Inteligencia Artificial, con los actuales sistemas de Aprendizaje Automático y Aprendizaje Profundo, está ganando cada vez más interés por parte de los artistas que se dedican a investigar las relaciones entre el arte, la ciencia y la tecnología, por lo que importantes centros de exposición también están proponiendo y discutiendo sus obras: Entre los que están en evidencia, el Barbican Center de Londres presenta *AI: More than Human* (1), el MAK de Viena *Künstliche Intelligenz & Du* (2), la Kate Vass Galerie de Zurich *Automat und Mensch* (3), la Transfer Gallery de Nueva York *Forging the Gods* (4) y finalmente el HEK de Basilea *Entangled Realities, Living with Artificial Intelligence* (5).

Es muy difícil obtener una idea acabada de lo que está sucediendo dado que del conjunto de propuestas es difícil, pero generalizando, sin embargo, es posible identificar dos líneas de tendencia algo opuestas: de una parte los artistas usan los sistemas inteligentes para realizar obras que consideran más actuales de las que otros artistas que no usan tales medios obtienen y del otro lado muchos artistas consideran que algunos resultados de la IA son feos, poco atractivos y deben ser reformulados. En ambos bandos la ética tiene que ver con la estética,

En una entrevista que saldrá en la edición de noviembre de 2020 de *Icon*, Pedro Almodóvar, el cineasta confiesa que le parece "un milagro" llegar a los cines con una experiencia cinematográfica en este momento de tantas dificultades. Se refiere a la película *La voz humana* sobre el texto de Cocteau "Al principio, mi idea de la exhibición era otra. Pero después de la proyección en Venecia, los distribuidores nos llamaron para estrenarla en cines de España y del extranjero, y sólo puedo decir que es maravilloso que suceda y más aún en un momento como éste". Y agrega: "El algoritmo me aterroriza y me horroriza". Vaya uno a saber qué entiende Almodóvar por algoritmo.

Dicho de otro modo, arte y moral, ética y estética dialogan sobre los sistemas inteligentes y van dibujando recorridos posibles y también límites.

9. Conclusiones

Sinceramente creo que los humanos tenemos una noción bastante clara de lo que está bien y lo que está mal de la vida en sociedad y además de su propia vida. Dicho de otro modo,

⁵⁸ Lev Manovich *L'estetica dell'intelligenza artificiale: Modelli digitali e analitica culturale* (Numerus Kindle Edition,

Luca Sossella editore (25 de junio de 2020)

⁵⁹ Gianluca Consoli, *Arte e cognizione. Rapporti tra estetica e intelligenza artificiale*, Bulzoni, Milano, 2006

que la moral social y la ética individual han acompañado al hombre desde su aparición en la Tierra.

Siendo tanto la moral (mores) como la ética (ethos) dos palabras que significan costumbre, necesariamente tienen que adaptarse a los tiempos. Pero no es que haya una gran diferencia en el transcurso de la historia. Ciertamente, la moral de un *pithecanthropus erectus* es distinta de la del cosmonauta Yuri Gagarin, porque el mundo a su alrededor es distinto, pero si se va a excavar que es lo que está bien y que lo que está mal se pueden enumerar las acciones que luego las religiones habrían de establecer como virtuosas o pecaminosas.

En el Antiguo Testamento se establecen las reglas de las religiones monoteístas y Dios tiene su oponente: Satán ⁶⁰ En el nuevo testamento se llamará el Diablo y se le aparece a Jesús, como tentación en el desierto.

Satán o el diablo no aparecen, por lo que se, en el budismo ni en hinduismo y hasta algunos creyentes famosos tuvieron problemas con esta figura. Leibniz ⁶¹, cuando escribe sobre la presencia de Dios en el mundo, dice que es fácil verla: la bondad, pero cuando quiere encontrar la maldad dice simplemente que es la ausencia de Dios en el mundo. A punto tal que puede indicar lo mismo apelando a números. La presencia de Dios en el mundo 1 y su ausencia 0. Y agrega que con esos dos elementos puede nombrar cualquier cosa y comienza con la sucesión fundamental de los primeros ocho números que se convierten en una tira de 0 y 1. ⁶²

Leibniz de este modo se libera del diablo en el cual no creía.

Pero sin tropezarnos con el diablo, la vida virtuosa y la maldad (o ignorancia en el budismo) parecen acompañar al hombre desde su aparición en Tierra. Es cierto que algunos autores han concebido y expresado las máximas morales en modos muy estrictos como Kant y su imperativo categórico y otros de una manera más contextualizada como David Ross con sus deberes *prima facie*. Pero ahí estamos: la moral parece no cambiar mucho, lo que cambia es el mundo y entonces debemos adecuar la ética al mundo nuevo.

Esto es lo que nos sucede con los sistemas inteligentes: personalmente no creo que haya otra moral, pero no se puede desconocer que hay otro mundo y por lo tanto la ética debe explicarse y expresarse en este nuevo mundo. Un mundo nuevo acelerado por los cambios tecnológicos como nunca lo fue. Antes los cambios tecnológicos duraban algún siglo y permitían a la sociedad ir adaptándose. En el último cuarto del siglo pasado todo se aceleró de manera vertiginosa: la aparición de Internet significó un antes y un después. Se puede pensar que eso sucedió porque tanto la teoría como la práctica estaban preparadas y por supuesto que es así, siempre es así, pero la aceleración que se desató fue inconmensurable con cualquier cambio del

⁶⁰*Satán* es una palabra hebrea que significa “acusador o adversario, alguien que resiste”. El término se utiliza 19 veces en el Antiguo Testamento, 14 de las cuales se encuentran en Job 1 y 2. También se menciona a *Satán* en 1 Crónicas 21:1, Salmos 109:6 y Zacarías 3:1-2

⁶¹Gottfried Wilhelm, Freiherr von Leibniz, *Ensayo de Teodicea. Acerca de la bondad de Dios, la libertad del hombre y el origen del mal.*, Abada Editores, Madrid, 2019, ISBN: 9788416160112

⁶²El lector dirá, “pero eso es el sistema binario” y tiene razón. A tal punto que le escribió a un jesuita que enseñaba en China sobre el tema y este le contestó alborozado que era una manera de explicar a los chinos el catolicismo pues el sistema binario coincidía con el IChin compuesto por 64 hexagramas. Ambos le escribieron al Papa quien sentenció que la religión católica debía enunciarse en latín.

pasado. La intuición que la lógica era sintáctica acercó la lógica que se estaba acelerando con la matemática y con la informática⁶³

Llevábamos veinticinco siglos, desde que Aristóteles sostuvo que la lógica se ocupaba de los enunciados verdaderos o falsos⁶⁴ y liberamos a la lógica de la semántica y la hicimos sintáctica como los programas de computación.

Los nuevos medios han potenciado algo que el hombre hizo siempre: imaginar realidades distintas y contarlas como un relato. Tal vez en torno a los fuegos de las cavernas, seguramente en el mito de la caverna de Platón, y se han multiplicado en modo vertiginoso.

Hoy los sistemas inteligentes están en todas partes desde el celular al dron y están transformando el modo de tratar con el banco, con la Administración pública, la Justicia, los transporte y nuestras relaciones con los otros y nosotros mismos.

No hay tiempo para reflexionar porque como bien había visto Italo Calvino, nuestro siglo esta signado por la velocidad ⁶⁵pero hay que actuar y ver si todo lo que se está haciendo tiene respeto por el otro, por la sociedad, por la humanidad, por el clima de la tierra.

La pandemia aceleró los procesos sociales, mostró las debilidades de los sistemas económicos, sociales y políticos; encerrando a la gente las hizo enfrentar con sus peores miedos: hambre, soledad, limitaciones, robots, carencias materiales y afectivas, pestes. Y nos privó de la esencialidad social que hablaba Aristóteles.

Se nos han venido encima los automóviles autónomos, las armas cibernéticas, *las fake news*, los reconocimientos faciales y la posibilidad de ubicar a cualquiera en cualquier lugar, la escasa defensa de nuestra privacidad de datos, las manipulaciones que se pueden hacer desde una red social o desde un sistema de voto, las manipulaciones con libros, películas y noticias, la facilidad con la cual se pueden limitar libertades con las nuevas tecnologías y con los encierros. A los viejos populismos del siglo XX se agregan los nuevos con rostro no tan duro, pero tanto o más violentos que aquellos, se deteriora la democracia en una época de crisis democrática.

Pero no todas son malas noticias: las nuevas metodologías permiten controlar mucho más lo que hace una Administración, favorece la participación, extiende la dimensión horizontal de la burocracia que fue vertical desde Luis XIV, crea chatbots de información para el ciudadano, alerta sobre la controlabilidad de lo que se decide y se hace. Nacen ciudades inteligentes

¿El gran público queda indiferente? Siempre el gran público fue indiferente. Inclusive la rebelión de las masas de orteguiana memoria fue simplemente para que los muchos obtuvieran los objetos exteriores de los pocos no para que cogobernaran. Pero los medios están, se trata de ponerlos en movimiento.

⁶³Carlos E. Alchourron, , Antonio A. Martino, *Logic Without Truth*, Ratio Juris, March 1990
<https://doi.org/10.1111/j.1467-9337.1990.tb00050.x>

⁶⁴Aristoteles.. *De Interpretations*. In *The Works of Aristotle*. Trans. and ed. W. D. Ross, vol. 1. Oxford: Oxford University Press, 1923

⁶⁵Italo Calvino, *Lezioni americane: sei proposte per il prossimo milenio*, Oscar Mondadori, 1985, Milano. Las seis propuestas son: levedad, rapidez, exactitud, visibilidad, multiplicidad, consistencia

Además, grandes instituciones supranacionales, estados nacionales, grandes corporaciones, iglesias, hasta la Unesco, se han percatado de los riesgos de una sociedad con sistemas inteligentes incontrolados y han comenzado a formular acuerdos, a redactar principios o pactar soluciones y vigilancias, a redactar un documento que sirva como guía mundial para vigilar el uso moral que se hace de los mismos.

Es de estos días la noticia que La Comisaria Europea de Asuntos Digitales y Antimonopolio, Margrethe Vestager, dijo que los gigantes de la tecnología tendrán que abrir sus archivos de anuncios a los reguladores e investigadores. Es probable que la medida impacte en empresas como Alphabet, propietaria de Google, Amazon, Apple y Facebook.

Sabemos que el medio no es el mensaje, pero lo limita, lo condiciona, lo coloca en otro contexto, lo puede manipular, tergiversar. Y entonces no basta con saberlo, no basta con recurrir a las soluciones éticas del pasado, no porque no sirvan sino porque hay que contextualizarlas y ello requiere que se afronte la nueva sociedad que se ha implantado.

Nace una ética global, aparecen morales ontocéntricas, biocéntricas y antropocéntricas y es necesario informarse, reflexionar, tomar partido. Se tiene clara conciencia que nuestras decisiones no solo afectan nuestra vida y nuestro futuro inmediato sino también el largo plazo y las próximas generaciones y es por eso que a las tres reglas clásicas de Asimov, se agrega una cuarta en la cual el robot debe cuidar de no afectar la perdurabilidad de la especie humana, como el mal tratamiento del ambiente o la afectación de los recursos naturales.

Se evidencian los sesgos con los cuales obramos y con los cuales creamos los sistemas inteligentes. Es fácil descubrir que cualquier regla de aprendizaje para un sistema inteligente (robótico o no) contiene las bases para hacer filtrar decisiones de mayorías prejuiciosas por las cuales es necesario hacer intervenir personas de culturas diferentes para descubrir los prejuicios anudados en la mayoría local.

Nace la identidad digital y es necesario velar sobre ella.⁶⁶

A las viejas tradiciones filosóficas se agregan las nuevas que acentúan la necesidad del multiculturalismo y la imperiosa necesidad de enfoques holísticos sin necesidad de perder la especificidad y la profundidad de los conocimientos puntuales. Recordando a Calvino: multiplicidad, visibilidad, exactitud y consistencia. Occidente va quedando chico y hay que hacer las cuentas también con la cultura oriental que nuestra cultura tomo siempre como “excentricidad” tal vez provechosa, como demuestra Jung, pero excéntrica.

La activación de las neurociencias demuestra sin lugar a dudas que el que medita tiene mayor conectividad de sinapsis que quien no lo hace y entonces la cultura tiene que ver directamente con la salud y con la inteligencia no solo de los sistemas sino del hombre mismo.

Hasta los artistas reclaman un ámbito ético a su labor sobre todo si se sirven de sistemas inteligentes y reivindican algo que siempre estuvo en la cultura flotando con dificultades de expresión: la ética está indisolublemente ligada a la estética.

Un joven acude a una clínica para un chequeo rutinario; en esa clínica hay cinco pacientes esperando trasplantes de órganos. Para poder vivir, dos de ellos necesitan un

⁶⁶ El decreto del DL 16 luglio 2020, n. 76 de simplificación digital italiano que comprende la carta d'identita digitale.

pulmón, otros dos necesitan sendos riñones y el quinto requiere un corazón. Curiosamente el joven que ha ido a hacerse el chequeo tiene el mismo grupo sanguíneo que ellos, lo que le convierte en el donante idóneo. ¿Es lícito matar a una persona para salvar a otras cinco?⁶⁷

En 2009, durante un experimento en el Laboratorio de Sistemas Inteligentes en la École Polytechnique Fédérale de Lausanne en Suiza, robots programados para cooperar entre sí (en la búsqueda de un recurso beneficioso y evitar uno venenoso) finalmente aprendieron a mentir a cada uno otro en un intento de acumular el recurso beneficioso. Un problema en este caso puede haber sido que los objetivos eran "terminales" (es decir, en contraste, los motivos humanos últimos típicamente tienen la cualidad de requerir un aprendizaje interminable).

Entre asustados, ignorantes, exaltados, angustiados, delirantes, habrá que concluir que “Algunas personas temen que la inteligencia artificial nos haga sentir inferiores, pero entonces cualquiera de nosotros tendría que sentir complejo de inferioridad al mirar una flor”.⁶⁸

La economía tiene que rendirle cuentas a la moral, no es posible que solo las leyes del mercado decidan la falta de trabajo de millones de personas, la hambruna y la angustia de muchos.

Con cada situación, con cada mundo que se va creando con los sistemas inteligentes es necesario plantearse hasta donde las acciones que se generan y las consecuencias que derivan pueden traer felicidad o dolor a los seres humanos que serán implicados por las mismas porque, en definitiva, como decía Kant: La moral no es realmente la doctrina de cómo hacernos felices, sino de cómo debemos hacernos dignos de la felicidad.

Se trata siempre de un tema de fines y medios y la frase “el fin justifica los medios” históricamente carece de sentido. Sucede que si se quiere un fin fuertemente se termina aceptando los medios, pero el medio puede polucionar, envenenar de tal manera el fin de hacerlo inaceptable. Por eso se debe tener en cuenta desde el punto de vista moral que cuando se acepta un fin los medios para llegar a él no deben tornarlo indeseable. Torturar a un maestro para aprender todo lo que él sabe es un ejemplo muy claro y burdo, pero la vida real está llena de claro oscuros, es allí donde se debe indagar sobre los medios y no basta con decir “no lo sabía”.

Cualquier persona razonable puede averiguar los peligros que implica cualquier uso de cualquier cosa. Es tan obvio esto que se empiezan a distinguir entre riesgos pequeños, medianos y grandes. Se puede admitir la discusión sobre el uso de un medio que tiene poco riesgo pero que podría éste aumentar, si los beneficios son enormes, pero no se trata solo de un tema de cantidad⁶⁹

Para terminar nada mejor que la paradoja de Fermi: él se preguntaba cómo era posible que cuando se buscaban vidas inteligentes estos siempre aparecían en la Tierra cuando era estadísticamente improbable que no hubiese vida inteligente en el vasto universo. Su hipótesis

⁶⁷Judith Jarvis Johnson, The Trolley problem The Yale Law Journal, 1985

⁶⁸Alan Kay, *The Computer "Revolution" Hasn't Happened Yet!* Educom Conferences, 1998. Alan Kay es también el autor de la famosa frase “la mejor manera de predecir el futuro es inventándolo”

⁶⁹ Si descalificar el magnífico trabajo del Mit para hacer el juego de la Moral Machine y sin desprestigiar la intuición de la creadora del dilema del tranvía, a veces, el tema de las cantidades es humanamente intolerable

era que cuando la tecnología va más rápido de la razón, la vida inteligente se destruye a sí misma. Tomémoslo como una advertencia.