

Filosofía de la tecnología: una introducción conceptual

Alvaro David Monterroza Rios

Juan Carlos Moreno Ortiz

Jorge Francisco Maldonado Serrano

Editores académicos



**Filosofía de la
tecnología:
una introducción
conceptual**

Filosofía de la tecnología: una introducción conceptual

Álvaro David Monterroza Ríos
Juan Carlos Moreno Ortiz
Jorge Francisco Maldonado Serrano

EDITORES ACADÉMICOS

John Jairo Cardozo Cardona
Daian Tatiana Flórez Quintero
Jorge Francisco Maldonado Serrano
Álvaro David Monterroza Ríos
Juan Carlos Moreno Ortiz
Diana María Muñoz González
Dairon Alfonso Rodríguez Ramírez

AUTORES

Red Colombiana de Filosofía de la Tecnología –
PHITEC



Monterroza Ríos, Álvaro David

Filosofía de la tecnología: una introducción conceptual / Álvaro David Monterroza Ríos, Juan Carlos Moreno Ortiz y Jorge Francisco Maldonado Serrano (Editores académicos) [y otros cinco autores], Primera edición, Bogotá: Universidad Santo Tomás, 2026.

250 páginas; figuras y tablas.

Incluye referencias bibliográficas e índice temático y onomástico.

ISBN: 978-958-782-732-3

E-ISBN: 978-958-782-733-0

1. Filosofía de la tecnología 2. Conocimiento tecnológico 3. Ética de la tecnología
4. Cambio tecnológico 5. Inteligencia artificial - Aspectos morales y éticos I.
Universidad Santo Tomás (Colombia).

CDD 601 CO-BoUST



LISTA DE AUTORES

Editores Académicos:

Alvaro David Monterroza Ríos

Juan Carlos Moreno Ortiz

Jorge Francisco Maldonado Serrano

Autores:

John Jairo Cardozo Cardona

Daian Tatiana Flórez Quintero

Jorge Francisco Maldonado Serrano

Álvaro David Monterroza Ríos

Juan Carlos Moreno Ortiz

Diana María Muñoz González

Dairon Alfonso Rodríguez Ramírez

© Universidad Santo Tomás, 2026

Ediciones USTA

Bogotá, D.C., Colombia

Carrera 9 n.º 51-11

Teléfono: (+571) 587 8797, ext. 2991

editorial@usta.edu.co

<http://ediciones.usta.edu.co>

Corrección de estilo: Ludwing Cepeda

Aparicio

Diagramación y montaje de cubierta: A Tres

Tintas SAS

Impresión: Imageprinting

Hecho el depósito que establece la ley

Red Colombiana de Filosofía de la Tecnología

– PHITEC

ISBN: 978-958-782-732-3

E-ISBN: 978-958-782-733-0

Primera edición, abril de 2026

Este libro hace parte de la Investigación Fodein MultiCampus “Mejoramiento de la conectividad digital y aprovechamiento de recursos tecnológicos mediante la implementación de una plataforma de aprendizaje asincrónica mediada por agentes conversacionales inteligentes” aprobada y ejecutada en la convocatoria n° 2021 financiada por la Universidad Santo Tomás y desarrollada por un equipo de profesores-investigadores adscritos a la Decanatura de Educación Abierta y a Distancia de la Universidad Santo Tomás (Bogotá, Colombia) en colaboración con la Universidad de Sevilla y la Universidad de Córdoba (España).

Universidad Santo Tomás

Vigilada MinEducación

Reconocimiento personería jurídica: Resolución 3645 del 6 de agosto de 1965, MinJusticia
Acreditación Institucional de Alta Calidad Multicampus: Resolución 014525 del 28 de julio de 2022, 8 años, MinEducación.

Se prohíbe la reproducción total o parcial de esta obra, por cualquier medio, sin la autorización expresa del titular de los derechos.

Impreso en Colombia • Printed in Colombia

Contenido

INTRODUCCIÓN	15
CAPÍTULO 1	
¿QUÉ ES TECNOLOGÍA? ASPECTOS ONTOLÓGICOS	19
INTRODUCCIÓN	19
LA TECNOLOGÍA ES UNA FORMA PARTICULAR DE TÉCNICA	22
CONCEPCIONES COMUNES DE TECNOLOGÍA	24
CONCEPCIÓN INSTRUMENTAL DE TECNOLOGÍA	25
CONCEPCIÓN INTELECTUALISTA DE LA TECNOLOGÍA	27
CONCEPCIÓN SISTÉMICA DE TECNOLOGÍA	29
POSIBILISMO TECNOLÓGICO, UNA RESPUESTA HISpanOHABLANTE AL DETERMINISMO	32
TECNOLOGÍA A LA LUZ DEL CONTEXTO DE DISEÑO Y CONTEXTO DE USO	41
¿QUÉ ES DISEÑAR?	41
EL CONTEXTO DE DISEÑO	44
EL CONTEXTO DE USO	47
¿CÓMO Y POR QUÉ CAMBIAN LAS TECNOLOGÍAS?	50
LOS MODELOS EVOLUTIVOS DEL CAMBIO TÉCNICO	56

¿MEJORES TECNOLOGÍAS? EL CONCEPTO “TECNOLOGÍAS ENTRAÑABLES”	62
PREGUNTAS	67
PARA SABER MÁS	68
CAPÍTULO 2	
CLASIFICAR EL CONOCIMIENTO TECNOLÓGICO: LA TECNOLOGÍA COMO CONOCIMIENTO	71
INTRODUCCIÓN	71
EL CONCEPTO TECHNÉ	73
CIENCIA Y TECNOLOGÍA	76
LA TECNOLOGÍA COMO INDEPENDIENTE DE LA CIENCIA	77
LA TECNOLOGÍA COMO CIENCIA APLICADA Y LA EVIDENCIA HISTÓRICA EN SU CONTRA: EL CASO DE LA MÁQUINA DE VAPOR	82
CONOCIMIENTO TECNOLÓGICO Y TÉCNICO	89
ESPECIFICIDAD DEL CONOCIMIENTO TECNOLÓGICO	90
CLASIFICACIÓN EPISTEMOLÓGICA DE DOS FORMAS DE CONOCIMIENTO TECNOLÓGICO	105
SIMONDON	106
CONOCIMIENTO Y USO DE HERRAMIENTAS	115
PREGUNTAS	123
PARA SABER MÁS	124
CAPÍTULO 3	
ÉTICA DE LAS TECNOLOGÍAS	127
INTRODUCCIÓN	127
CONTEXTO HISTÓRICO Y DESARROLLO DE LA ÉTICA DE LAS TECNOLOGÍAS	128
LAS CONDICIONES Y CARACTERÍSTICAS QUE DEFINEN LAS ÉTICAS DE LAS TECNOLOGÍAS	130

PROPÓSITOS Y FUNCIONES DE LAS ÉTICAS DE LAS TECNOLOGÍAS	130
LOS PUNTOS CARDINALES DE LAS ÉTICAS DE LAS TECNOLOGÍAS	132
LOS TIPOS DE INSTRUMENTALIZACIÓN TECNOLÓGICA	135
LOS SENTIDOS DE AGENCIA DE LAS TECNOLOGÍAS Y SUS PRESUPUESTOS ONTOLÓGICOS	137
PERSPECTIVAS ÉTICAS TRADICIONALES	140
ÉTICAS DE LAS TECNOLOGÍAS HUMANISTAS	140
ÉTICAS DE LA TECNOLOGÍA NORMATIVAS	143
ÉTICAS DE LAS TECNOLOGÍAS PRAGMÁTICAS Y PROCEDIMENTALES	145
ANÁLISIS COMPARATIVO DE LAS PERSPECTIVAS	146
NUEVAS ORIENTACIONES EN LAS ÉTICAS DE LAS TECNOLOGÍAS	147
EL GIRO EMPÍRICO EN LA FILOSOFÍA DE LA TECNOLOGÍA	149
EL GIRO SOCIAL DE LAS ÉTICAS DE LAS TECNOLOGÍAS	150
EL GIRO ONTOLÓGICO RELACIONAL EN LA FILOSOFÍA DE LA TECNOLOGÍA	152
EL GIRO MATERIAL EN LAS ÉTICAS DE LAS TECNOLOGÍAS. LA RELEVANCIA MORAL DE LOS ARTEFACTOS TECNOLÓGICOS	154
EL GIRO NORMATIVO DE LAS ÉTICAS DE LAS TECNOLOGÍAS	158
ANÁLISIS DEL CASO DE LA GEOINGENIERÍA Y EL TRANSHUMANISMO: DOS CARAS DEL SOLUCIONISMO TECNOLÓGICO CONTEMPORÁNEO	161
HACIA UNA ÉTICA CONSTRUCTIVA DE LAS TECNOLOGÍAS	168
PREGUNTAS	170
PARA SABER MÁS	171
 CAPÍTULO 4	
FILOSOFÍA DE LA INTELIGENCIA ARTIFICIAL	173
INTRODUCCIÓN	173
¿QUÉ ES INTELIGENCIA ARTIFICIAL?	175
HISTORIA DE LAS INTELIGENCIAS ARTIFICIALES	178
INTELIGENCIA ARTIFICIAL “DÉBIL” VS. “FUERTE”	184

LAS INTELIGENCIAS ARTIFICIALES GENERATIVAS	186
LOS MODELOS DE PROCESAMIENTO DE LENGUAJE NATURAL	187
CHATGPT NO PIENSA, CALCULA Y TAMBIÉN SE EQUIVOCA	192
PRINCIPIOS ÉTICOS PARA LA REGULACIÓN DE LA IA	194
LA ÉTICA DE LA IA DESDE UN PUNTO DE VISTA MÁS AMPLIO QUE EL NORMATIVO	204
ESTATUTO MORAL DE LA IA Y PERSPECTIVAS DE ANÁLISIS ÉTICO	207
DESAFÍOS ÉTICOS DE LA IA	215
PREGUNTAS	218
PARA SABER MÁS	218
CONSIDERACIONES FINALES	221
REFERENCIAS	225
SOBRE LOS AUTORES	243
ÍNDICE TEMÁTICO Y ONOMÁSTICO	247

Lista de figuras

FIGURA 1. LA TECNOLOGÍA ES UNA FORMA PARTICULAR DE TÉCNICA ORGANIZACIONAL BASADA EN PROCESOS DE DISEÑO	23
FIGURA 2. LA TRAYECTORIA HISTÓRICA DE UN GRUPO HUMANO CAMBIA CUANDO SE APROPIA DE UN NUEVO DISEÑO TECNOLÓGICO	37
FIGURA 3. TECNOLOGÍA COMO INTERSECCIÓN DE POSIBILIDADES	38
FIGURA 4. DESDE EL CONTEXTO DE DISEÑO, TECNOLOGÍA ES UN SISTEMA DE PRÁCTICAS PARA LA TRANSFORMACIÓN DE LA REALIDAD CONSTITUIDO POR AGENTES, ARTEFACTOS, CONOCIMIENTOS, NORMAS Y VALORES	47
FIGURA 5. TECNOLOGÍA DESDE EL CONTEXTO DE USO. LOS PRODUCTOS DE LAS PRÁCTICAS TECNOLÓGICAS OFRECEN NUEVAS POSIBILIDADES PRAGMÁTICAS QUE CAMBIAN EXPECTATIVAS	50
FIGURA 6. INTERRELACIONES DE CAUSAS Y EFECTO ENTRE EL CAMBIO TECNOLÓGICO CON EL DESARROLLO CIENTÍFICO, EL DESARROLLO ECONÓMICO Y EL CAMBIO SOCIAL	52
FIGURA 7. ESQUEMA REPRESENTATIVO DE UN PROCESO DE VARIACIÓN Y SELECCIÓN EN LA EVOLUCIÓN DE UNA CLASE DE ARTEFACTO (A) EN UN NICHOS DE CULTURA MATERIAL	60
FIGURA 8. CONTENIDO DEL CONOCIMIENTO TECNOLÓGICO	94
FIGURA 9. PUNTOS CARDINALES DE LA ÉTICA DE LAS TECNOLOGÍAS	133
FIGURA 10. FLUJO FUNCIONAL DE UN LLM	188

Lista de tablas

TABLA 1. TIPOS DE POSIBILIDADES QUE SE TIENE EN CUENTA EN UN DISEÑO TECNOLÓGICO	39
TABLA 2. DIVERSOS ENFOQUES EXPLICATIVOS DEL CAMBIO TÉCNICO	53
TABLA 3. COMPILACIÓN DE UNIDADES DE ANÁLISIS DE EVOLUCIÓN BASADA EN REVISIÓN BIBLIOGRÁFICA	58
TABLA 4. DECÁLOGO DE CARACTERÍSTICAS DE TECNOLOGÍAS ENTRAÑABLES	65
TABLA 5. CARACTERÍSTICAS DE LAS ÉTICAS DE LAS TECNOLOGÍAS	130
TABLA 6. TIPOS DE INSTRUMENTALIZACIÓN TECNOLÓGICA	136
TABLA 7. TIPOS DE AGENCIA DE LAS TECNOLOGÍAS	139
TABLA 8. SÍNTESIS ENFOQUES ÉTICOS TRADICIONALES	146
TABLA 9. POSIBLES OBJETIVOS DE LA IA	176
TABLA 10. HITOS DE LA IA	183
TABLA 11. EXPLICACIÓN DE LAS ETAPAS DE FLUJO FUNCIONAL DE UN LLM	189
TABLA 12. PRINCIPIOS ÉTICOS APLICABLES A LA IA	195
TABLA 13. ROLES MORALES QUE PUEDE ASUMIR UNA IA	208
TABLA 14. RIESGOS Y DESAFÍOS QUE SE PUEDEN DERIVAR DEL USO DE IA	209

Introducción

En las últimas décadas, la filosofía de la tecnología se ha consolidado como un campo de reflexión que permite comprender, con mayor precisión conceptual, cómo se configuran las prácticas humanas en entornos crecientemente mediados por artefactos, infraestructuras y sistemas técnicos. Este creciente interés no se explica únicamente por la expansión cuantitativa de dispositivos o por la aceleración de la innovación, sino por un hecho más fundamental: la tecnología reorganiza las condiciones de posibilidad para actuar, conocer, coordinarse y habitar el mundo. Por ello, las discusiones contemporáneas sobre educación, comunicación, producción, gobierno, cultura digital o transformación social remiten, de forma directa o indirecta, a problemas propiamente filosóficos: qué es la tecnología, qué tipo de racionalidad expresa, qué clase de conocimiento moviliza, y cómo se evalúa normativamente su presencia en la vida colectiva.

A pesar de esta centralidad, gran parte del debate público e incluso académico sobre tecnología continúa operando con nociones imprecisas o reduccionistas. Con frecuencia, se la identifica con un conjunto de objetos discretos, con “aplicaciones” de la ciencia, o con un proceso autónomo que avanza al margen de decisiones humanas. Estas aproximaciones suelen oscilar entre posiciones celebratorias y posiciones alarmistas, y tienden a suprimir matices decisivos: la historicidad de las formas técnicas, la mediación material de la acción, la dimensión organizacional de los sistemas tecnológicos, la distribución de agencia entre humanos y artefactos, así como la inscripción de valores en

arreglos de diseño, métricas, estándares y procedimientos. En este sentido, la filosofía de la tecnología ofrece un conjunto de herramientas conceptuales para examinar críticamente esas simplificaciones y reconstruir con claridad los supuestos que orientan la interpretación de lo tecnológico y sus efectos.

Este libro se concibe como una introducción a los temas más discutidos en el campo, organizada de manera sistemática para servir tanto a lectores que se aproximan por primera vez a estas cuestiones como a quienes buscan un marco ordenado para la docencia o la investigación interdisciplinar. Su propósito es presentar un mapa conceptual que permita comprender la tecnología como un fenómeno que no se agota en la materialidad de los artefactos ni en la intencionalidad de sus diseñadores, sino que involucra prácticas, instituciones, conocimientos, regímenes de evaluación y formas de coordinación social. En esa medida, el texto se estructura alrededor de tres familias de preguntas que atraviesan de modo recurrente la literatura especializada: preguntas ontológicas sobre la naturaleza e identidad de lo tecnológico, preguntas epistemológicas sobre el tipo de conocimiento implicado en su producción y operación, y preguntas éticas y políticas sobre su evaluación y orientación.

Un punto de partida del libro, desarrollado en el primer capítulo, es la necesidad de precisar qué debe entenderse por “tecnología” cuando se habla de sociedades contemporáneas. En ese capítulo se propone una definición que busca evitar equívocos frecuentes: por un lado, la identificación de la tecnología con cualquier forma de técnica en sentido amplio; por otro, la reducción de la tecnología a la mera presencia de artefactos o dispositivos. La definición presentada sostiene que la tecnología es una forma particular de acción técnica contemporánea mediada por procesos de diseño, caracterizada por su carácter colectivo y organizacional, y consolidada históricamente tras la primera Revolución Industrial. Además, se subraya que esta acción se apoya en modelos y representaciones, en procedimientos metódicamente programados y en mecanismos de evaluación de resultados, orientados a producir cambios causales que transforman el horizonte de posibilidades de grupos humanos. Esta definición no pretende clausurar el

debate, sino ofrecer un marco operativo para articular los problemas que el libro desarrolla.

El libro empieza con los aspectos ontológicos de la tecnología. El primer capítulo presenta concepciones comunes que han orientado tanto el pensamiento ordinario como diversas teorías filosóficas: concepciones instrumentales centradas en el artefacto como medio para fines, concepciones intelectualistas que comprenden la tecnología como derivación o aplicación del conocimiento científico, y concepciones sistémicas que la entienden como entramado sociotécnico compuesto por artefactos, infraestructuras, organizaciones, normas y prácticas. A partir de esa discusión, el capítulo sitúa la tecnología contemporánea como una modalidad de acción técnica que requiere ser pensada en términos de diseño, organización, coordinación y estabilización de prácticas, y prepara las preguntas sobre cambio técnico y sobre la relación entre artefactos, sistemas y cultura material.

El segundo capítulo se ocupa de la dimensión epistemológica, esto es, del estatuto del conocimiento tecnológico. En lugar de asumir que la tecnología es simplemente ciencia aplicada, el capítulo reconstruye los argumentos que muestran su especificidad: la orientación hacia la acción y la resolución de problemas, el papel de modelos, representaciones y prototipos, la relevancia de habilidades y procedimientos, y la presencia de criterios normativos internos que no se reducen a la verdad o la explicación, sino que incluyen funcionalidad, seguridad, confiabilidad, eficiencia y adecuación a contextos de uso. Con ello se busca clarificar cómo se produce y valida el conocimiento tecnológico, y por qué su racionalidad no coincide plenamente con la racionalidad científica, aun cuando ambas mantengan relaciones estrechas en múltiples ámbitos.

El tercer capítulo introduce la pregunta ética y política. Partiendo de la idea de que las tecnologías median la experiencia y reorganizan prácticas, el capítulo examina por qué la evaluación moral no puede limitarse a un juicio posterior sobre consecuencias, sino que debe considerar también los modos en que los artefactos y sistemas configuran posibilidades de acción, distribuyen responsabilidades y estabilizan formas de vida. Se presenta el tránsito desde enfoques centrados en riesgos, impactos y regulación, hacia perspectivas que insisten en la necesidad

de pensar la orientación de los desarrollos tecnológicos, lo que implica atender a decisiones de diseño, gobernanza, participación, responsabilidad institucional y justicia distributiva. El objetivo no es ofrecer una lista cerrada de criterios, sino mostrar los problemas que emergen cuando se toma en serio el carácter mediador y organizador de la tecnología.

El cuarto capítulo aplica el marco ontológico, epistemológico y ético a un caso de alta relevancia contemporánea: la inteligencia artificial, con énfasis en la IA generativa. Este capítulo aborda cómo reaparecen, en un dominio específico, preguntas tradicionales del campo: qué tipo de agencia se atribuye a sistemas inteligentes, qué significa conocimiento y decisión en contextos algorítmicos, cómo se distribuye la responsabilidad entre desarrolladores, organizaciones, usuarios y marcos regulatorios, y qué desafíos se plantean para la evaluación y orientación de tecnologías que intervienen en procesos de información, comunicación y producción simbólica. Con esta aplicación se busca mostrar la continuidad entre debates clásicos de la filosofía de la tecnología y transformaciones técnicas recientes, evitando tratar la IA como un tema aislado o puramente técnico.

Por estas razones, el libro sostiene que la comprensión de la tecnología requiere el esfuerzo de articular tres planos: definir su naturaleza como modalidad de acción mediada por diseño y organización, examinar el tipo de conocimiento que la hace posible, y discutir los criterios con los cuales se justifica, evalúa y orienta en la vida colectiva. El resultado es un itinerario introductorio que, sin renunciar a la complejidad del campo, ofrece al lector un conjunto de distinciones y problemas para pensar de manera más precisa qué es la tecnología, cómo opera y por qué su análisis filosófico es inseparable de la comprensión de las sociedades contemporáneas.

LOS EDITORES

Capítulo 1

¿Qué es tecnología? Aspectos ontológicos

ÁLVARO DAVID MONTERROZA RÍOS

Introducción

Comprender los aspectos ontológicos de las entidades de la realidad humana es fundamental porque abre el espacio de comprensión para plantear, y resolver provisionalmente, nuevas preguntas filosóficas. Atender a los aspectos ontológicos es abrir la caja negra que permite una comprensión en perspectiva de fenómenos que afectan y constituyen la realidad humana. Ese es el propósito de este capítulo, plantear preguntas sobre ¿por qué la naturaleza técnica del ser humano?, ¿cómo distinguir las distintas formas de acción técnica?, ¿qué características tiene la tecnología contemporánea?, ¿cómo se ha comprendido comúnmente?, ¿por qué y cómo sucede el cambio técnico? y ¿podrían existir tecnologías con cualidades más deseables? Estas son las preguntas sobre las que se desarrolla el capítulo, intentando plantear la diversidad de enfoques, pero también proponiendo posiciones sugerentes de la rica tradición hispanohablante de la filosofía de la técnica y la tecnología.

Partimos de la posición de que la acción técnica es la manera en que todos los grupos humanos han posibilitado sus formas de vida, es la manera más común en que el ser humano se ha forjado a sí mismo extendiendo los límites de sus propias capacidades biológicas. Como lo decía José Ortega y Gasset (1982), sin técnica simplemente no hay

humano. Por supuesto, la transformación técnica del mundo nos ha traído problemas de gran alcance que han sido resultados no esperados de la misma acción técnica colectiva, en especial, en las sociedades globales en que la tecnología ocupa cada rincón de nuestra realidad. Es que, en la actualidad, la tecnología tiene incidencia en múltiples esferas humanas, tales como la social, política, económica, científica, cultural, moral y, por supuesto, técnica. Es tanta la incidencia de nuestra acción técnica que no nos reconocemos a nosotros mismos como una especie más del planeta, pero somos el resultado de las contingencias y casos fortuitos que nos trajeron hasta dónde estamos ahora. Nuestra condición de seres técnicos está en nuestra propia constitución como humanos, así esta afirmación parezca escandalosa para quien ve la técnica, la tecnología, las máquinas y los artefactos como frías herramientas útiles o como capuchas que esconden una supuesta esencia humana primigenia.

La tecnología es un concepto que solemos usar de múltiples formas en nuestro lenguaje cotidiano; normalmente lo asociamos a la electrónica de consumo, a la informática, como una forma de conocimiento científico aplicado, como una infraestructura o como una forma de pensamiento. No obstante, consideramos desde la filosofía que se le puede dar un significado más amplio al concepto para poder comprender mejor por qué ha sido uno de los fenómenos que más ha tenido impacto en la realidad humana. La tecnología contemporánea es capaz de cambiar los distintos ámbitos, tales como las relaciones sociales, la economía, la organización militar y política, la relación con los ecosistemas naturales, el gusto estético, los valores epistémicos y morales, y la misma identidad social e individual. Por supuesto, estos cambios se han incrementado de manera sustancial tras las revoluciones industriales y han traído múltiples beneficios evidentes frente a las condiciones colectivas del pasado, pero también nuevos riesgos y problemas que no existirían sin las trayectorias tecnológicas actuales.

¿Y qué podemos definir como tecnología? La tecnología es una forma particular de acción técnica contemporánea mediada por procesos de diseño, que se caracteriza por ser colectiva y organizacional, y que se consolidó después de la primera Revolución Industrial. Los procesos de diseño tecnológico son guiados teóricamente mediante modelos y representaciones que requieren acciones metódicamente programadas y la evaluación

de sus resultados, esto es, la tecnología es una actividad creadora, basada en métodos sistemáticos que usa sus propios medios representacionales, fundamentada en conocimientos fiables (sobre todo científicos), para llevar a cabo un cambio causal en una parcela de realidad que transforma el horizonte de posibilidades de un grupo humano. Para ilustrarlo, consideremos el diseño y la implementación de una nueva infraestructura tecnológica como un sistema de acueducto, una red de carreteras, nuevos equipos hospitalarios, la adquisición de dispositivos de electrónica de consumo o sistemas de telecomunicaciones. Este tipo de intervenciones transforma de manera efectiva las condiciones de posibilidad de una comunidad determinada, al introducir cambios significativos en sus prácticas cotidianas: resuelve ciertos problemas, pero también ocasiona otros imprevistos, además de modificar las expectativas y producir nuevas necesidades percibidas para tal grupo humano.

En el caso de un sistema de acueducto, por ejemplo, se soluciona el suministro de agua potable en una comunidad rural, lo que mejora las condiciones de consumo, higiene y salubridad. Sin embargo, emergen nuevos problemas y necesidades, como el mantenimiento y la reparación de tuberías, el pago periódico del servicio o la gestión de fallas en el suministro, etc. De este modo, el acueducto pasa de ser visto como un lujo para la comunidad a ser una necesidad básica, alterando de forma estructural la trayectoria histórica de la comunidad.

No obstante, una definición completa de un fenómeno complejo como la tecnología nunca se logra, por lo cual durante el capítulo se presentarán apartes para explicar los diversos aspectos de lo que entendemos por tecnología como, por ejemplo, destacar el carácter técnico que ha sido constitutivo de la identidad humana y su orden social, pero que con las tecnologías contemporáneas surgen cambios significativos y problemáticos en esta relación. Por lo tanto, en primer lugar, intentaremos describir qué rasgos en particular, desde el contexto de diseño y del contexto de uso, puede comprender la tecnología. También el carácter contingente, cambiante y parcialmente acumulativo de los logros humanos de la acción técnica humana, y, por último, una reflexión sobre el carácter abierto y normativo de la tecnología y las posibles vías para una actitud crítica, pero propositiva, frente a los cambios tecnológicos.

La tecnología es una forma particular de técnica

Nuestra historia evolutiva y cultural ha sido moldeada por los resultados de nuestras propias acciones técnicas al punto de abrir nuevas trayectorias evolutivas propias en la que coevolucionamos con nuestro entorno de artefactos (Monterroza-Ríos, 2019). Existe mucha evidencia para afirmar que los humanos somos esencialmente seres técnicos, ya que, sin técnica no habría humano tal como lo conocemos (Ortega y Gasset, 1982; Mithen, 1998; Byrne, 2004). Según José Ortega y Gasset (1982), nuestras capacidades práctico-mentales de memoria individual, y las habilidades para poder plasmar la memoria en elementos materiales externos (Stiegler, 1998; Simondon, 2007), nos da esa capacidad de separarnos parcialmente de la realidad inmediata. De esta manera, podemos reconstruir continuamente el pasado abriendo el horizonte de posibilidades a lo que “podría ser”, lo que nos da una particular manera de existir diferente a otros animales, pues no estamos completamente atados a las necesidades biológicas, sino que vivimos en el reino de las posibilidades (Broncano, 2012).

El mismo Charles Darwin (2013) en *The Descent of Man*, escrito en 1871, había afirmado ya que la técnica humana es un rasgo distintivo de nuestra especie, ya que es una forma de acción (en principio azarosa y casi accidental) que produjo nuevas habilidades, un cambio en los microentornos, un cambio de dieta, etc., que dejó como resultado un entorno de objetos funcionales con los cuales establecimos una relación (cercana a un ecosistema) que sustentan y posibilitan la mayoría de las prácticas en cualquier cultura (Flórez Vega, 2024). Esto es, hay un efecto retroalimentador que producen las acciones técnicas, ya que cambiaron las mismas trayectorias evolutivas de nuestra especie (Monterroza, 2018). Pero específicamente ¿qué es técnica?

Las técnicas humanas son formas de acción prácticas guiadas por fines y propósitos, en principio, funcionales. Es una manifestación de la agencia humana que sería algo así como la capacidad de transformar una parcela de la realidad inmediata mediante acciones intencionadas reflexivas. La acción técnica humana tiene la particularidad de ser colectiva, por lo tanto, aprendida de otros pares humanos, y apoyada en andamios culturales como herramientas, materiales, conocimientos y

normas de acción (Monterroza-Ríos y Novikova, 2020). No solo los humanos tienen técnica, muchos animales tienen estas acciones y sus artefactos (nidos, telarañas, hormigueros, etc.); pero las técnicas humanas tienen la particularidad de ser aprendidas formal y explícitamente (no solo instintivas o imitadas), además de ser desarrolladas con apoyo en entornos enriquecidos que llamamos *cultura material*.

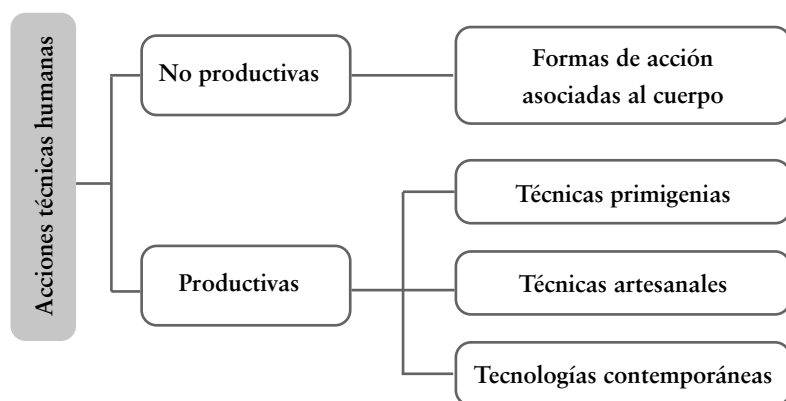


Figura 1. La tecnología es una forma particular de técnica organizacional basada en procesos de diseño

Fuente: elaboración propia.

De acuerdo con lo anterior, las técnicas humanas son formas de acción que pueden ser o no productivas: las acciones no-productivas están relacionadas con el manejo del cuerpo que deja resultados intangibles, mientras que las acciones productivas también están relacionadas con el manejo del cuerpo, pero dejan un producto concreto (figura 1). Cualitativamente han existido diversos tipos de técnicas productivas, que van desde las presentes en cualquier cultura primigenia o ancestral con la que se construyen herramientas, habitáculos, modificaciones corporales o del paisaje, hasta las formas productivas contemporáneas de producción industrial.

¿Qué diferencia hay entre las técnicas tradicionales y la tecnología industrializada? La respuesta es que a la tecnología la describimos

como la forma de producción que predomina en las sociedades contemporáneas y que tuvo su origen durante las revoluciones industriales de Europa y Norteamérica, sucedidas entre finales del siglo XVIII y mediados del XX. En esa época, se produjeron cambios decisivos, como la sustitución generalizada de herramientas por máquinas, la introducción de nuevas fuentes de energía para el trabajo mecánico (carbón, electricidad, petróleo), la organización especializada de la producción en fábricas y, especialmente, la separación de la “concepción” de la “fabricación” de productos con el uso de representaciones para la fabricación de piezas siguiendo planes de diseño. Estos cambios desencadenaron la división institucional del trabajo, en los cuales la planificación, la ejecución, el control y la innovación de la producción de bienes ya no sucede a escala individual, sino a escala organizacional (Quintanilla, 2017).

En términos generales, podemos decir que la tecnología contemporánea es una forma de acción técnica colectiva y organizacional mediada por un conjunto de métodos, máquinas, herramientas, conocimientos representacionales y operacionales (principalmente científicos), habilidades y valores plasmados en procesos colectivos de diseño representacional; sujeta a marcos normativos, éticos y pragmáticos que predominan las formas de producción técnica en el siglo XXI. No obstante, para llegar a esta afirmación hay que examinar primero algunas concepciones comunes del concepto “tecnología” que predominan en la divulgación, los sistemas educativos, las discusiones académicas y, en general, en el lenguaje cotidiano.

Concepciones comunes de tecnología

Como la tecnología es un fenómeno técnico que cambia las condiciones materiales de las sociedades en general, es difícil clasificarla de una sola manera. En ese sentido, se suele tener concepciones comunes sobre lo que es, tanto en el ámbito cotidiano como en los ámbitos más académicos. En principio, (1) a la tecnología se le suele asociar cotidianamente con el conjunto de artefactos o máquinas, pero, si observamos con cuidado, una máquina o una herramienta no hace nada

por sí sola, requiere de la interacción práctica de agentes humanos en marcos sociales más amplios para que la tecnología sea capaz de cambiar el estado de cosas.

También (2) se le suele asociar con conocimientos científicos aplicados, lo cual no es algo equívoco, pues una de las características de la tecnología contemporánea es que los ingenieros, tecnólogos, diseñadores y técnicos han aprendido de las metodologías y parten de los conocimientos de las ciencias físicas, matemáticas, biológicas, químicas, etc., para solucionar problemas prácticos. Sin embargo, reducir tecnología a “ciencia aplicada” puede ser incompleto, ya que en la práctica tecnológica hay también elementos cognitivos que no son formalizables en enunciados, principios, leyes o modelos matemáticos tales como los conocimientos operacionales, las habilidades, los valores, ciertas aptitudes, etc.; además, se compone también de elementos materiales esenciales como máquinas, herramientas e infraestructuras que no se pueden reducir a conocimiento o información.

Por otro lado (3), hay concepciones de tecnología que la conciben como un sistema sociotécnico, la cual incluye los elementos contextuales del entorno social, político y económico. Esta noción sociotécnica de la tecnología es mucho más completa que las anteriores, pero, en su afán de señalar los aspectos sociales de la tecnología, no se suelen dar criterios de demarcación claros con otras esferas sociales como la económica, política, moral, etc. Veamos con más detalle las concepciones más comunes de tecnología y qué ventajas y desventajas tienen a la hora de explicar el fenómeno tecnológico.

Concepción instrumental de tecnología

La concepción instrumental o artefactual de la tecnología es una perspectiva que entiende a la tecnología como el conjunto de herramientas, máquinas, sistemas o infraestructuras creados y utilizados por los seres humanos para alcanzar objetivos específicos y satisfacer necesidades. Desde esta concepción, la tecnología son sus aspectos materiales y físicos, y es vista como un medio para un fin. Su valor se deriva de su capacidad para facilitar la realización de tareas y resolver problemas. Es la concepción más cercana al sentido común y se encuentra en el

lenguaje cotidiano. Desde esta perspectiva, una máquina es tecnología, una red de comunicaciones es tecnología, un celular es tecnología, etc. Es el concepto de tecnología que encontramos con más frecuencia en los medios de comunicación y divulgación.

Esta concepción permite identificar los artefactos concretos y divide tajantemente el ámbito humano del ámbito técnico. Aunque es la visión más cercana al sentido común, también es adoptada tácitamente por algunos críticos del impacto negativo de la tecnología (Marcuse, 1968), así como de divulgadores acríticos que celebran cada avance tecnológico (Kurzweil, 2005). Consideramos que es un problema considerar instrumentalmente la tecnología de esta manera por las siguientes razones:

En primer lugar, un artefacto solo, incluso los artefactos más automatizados, no son capaces de hacer nada por sí solos; requieren del acople estructural con otras redes de artefactos y con los humanos para instaurar una práctica que sí transforme causalmente el mundo (que se supone que es lo que hacen las tecnologías) (Broncano, 2006). Pongamos un ejemplo práctico: una lavadora automática (un artefacto) tiene la función propia de “lavar ropa”, pues una lavadora sola no hace nada, requiere tener conexión acoplada con los sistemas de acueducto, alcantarillado y eléctrico, requiere el uso de materiales y la entrada y selección de ropa sucia, pero además requiere de un conjunto de intenciones, habilidades y conocimientos de un agente humano en un entorno doméstico. El agente humano, en este conjunto de operaciones al que llamamos “práctica”, debe saber normas tácitas de operación, pero también valores para seguir sus fines (la higiene es un valor) y también para operar la lavadora (el cuidado es otro valor). Esta serie de elementos deben ser tenidos en cuenta a la hora de concebir la tecnología, pues, como observamos en el ejemplo, las tecnologías no pueden ser solo los artefactos.

Además, en la concepción instrumental, la tecnología es considerada neutral en términos valorativos, ya que su función es simplemente servir como un medio para alcanzar un fin determinado por el usuario, una forma de racionalidad instrumental (Adorno y Horkheimer, 1947/2003). La responsabilidad ética recae en las personas que utilizan los artefactos, no en los artefactos en sí mismos. Desde esta perspectiva,

los artefactos no tienen valores inherentes, y cualquier impacto ético, social o ambiental que puedan tener es resultado de las acciones y decisiones humanas (García Palacios et al., 2001).

La concepción instrumental de la tecnología ha sido criticada por diversos autores y filósofos de la tecnología (Winner, 1980; Feenberg, 1991; Quintanilla, 2017). Estos críticos argumentan que los artefactos tecnológicos no son neutrales y que, de hecho, están moldeados por valores y objetivos humanos. Además, como estos sistemas de artefactos tienen impactos significativos en la sociedad, la cultura y el medio ambiente, es importante reconocer y abordar estos impactos en lugar de simplemente considerarlos como un medio neutral para lograr fines humanos.

Concepción intelectualista de la tecnología

Esta concepción entiende la tecnología como un conjunto de métodos que facilitan la aplicación de conocimientos derivados de las ciencias naturales a fines prácticos. En términos sintéticos, la tecnología es una forma de ciencia aplicada. Uno de los defensores implícitos de esta idea es la que propuso, ya hace un tiempo, el filósofo argentino Mario Bunge (Bunge, 1963) para quien la tecnología es un conjunto de prácticas y procesos basados en la ciencia que se utilizan para resolver problemas prácticos y alcanzar objetivos específicos. En efecto, según esta idea, la tecnología se trata de la aplicación del método científico y las teorías científicas para alcanzar objetivos prácticos. Para el autor, esto plantea problemas filosóficos interesantes como la naturaleza del conocimiento tecnológico, el poder validador de la acción, la relación entre las reglas tecnológicas y las leyes científicas, así como los efectos de los pronósticos tecnológicos en el comportamiento humano (Bunge, 1963).

Por supuesto, el estudio de la relación entre la ciencia y la tecnología es una cuestión que ha planteado muchas líneas desde la filosofía. Uno de los efectos de considerar la tecnología de esta manera es que conlleva implícitamente una concepción del modelo lineal de desarrollo, como el de Vannevar Bush (1945), esto es, el modelo de innovación en el cual se sugiere que hay un desarrollo consecutivo entre ciencia básica, ciencia aplicada, tecnología (artefactos) e innovación.

Autores como los de los *Cuadernos de CTS* de la OEI (García Palacios et al., 2001) señalan que una imagen intelectualista de la tecnología lleva a decir equivocadamente que: 1) la tecnología es principalmente un conjunto de reglas tecnológicas; 2) las reglas tecnológicas son consecuencias deducibles de las leyes científicas; y 3) el desarrollo tecnológico depende de la investigación científica (García Palacios et al., 2001, p. 38), lo que muestra la relación de subordinación de la tecnología a la ciencia.

Existen varias razones por las cuales la tecnología no puede entenderse como subordinada al conocimiento científico. John M. Staudenmaier presenta algunas de estas. En primer lugar, la tecnología no es simplemente una aplicación de la ciencia, ya que puede modificar y evolucionar los conceptos científicos existentes. En segundo lugar, el conocimiento tecnológico tiene su propia especificidad y no puede ser reducido a la ciencia, pensemos en las ciencias exclusivas de entidades artificiales como la ciencia de los materiales o la cibernética. En tercer lugar, las ingenierías, que son disciplinas netamente tecnológicas, hacen contribuciones importantes a los problemas que la ciencia no ha abordado. Finalmente, la tecnología depende de las habilidades técnicas, que son diferentes de las habilidades científicas, lo que subraya aún más la autonomía de la tecnología como práctica en sí misma (Staudenmaier, 1985).

Ilkka Niiniluoto ha explorado la relación entre la ciencia y la tecnología; según el autor finlandés, no existe una relación lineal entre ciencia y tecnología, sino una continuidad entre ellas, pues ambas comparten características similares en cuanto a estructura, lógica y metodología. La ciencia busca adquirir conocimiento sobre el mundo natural y sus leyes, mientras que la tecnología se enfoca en desarrollar conocimiento para resolver problemas prácticos de la vida humana, además, destaca la importancia de la innovación, el diseño y la evaluación en el desarrollo científico (Niiniluoto, 1997) que es un aspecto que casi que ha olvidado la filosofía de la ciencia. Para un estudio más detallado sobre los problemas de comprender la tecnología como ciencia aplicada, se recomienda el artículo de Flórez et al. (2019). En el capítulo 2 de este libro se profundizará sobre los aspectos epistemológicos de la tecnología, retomando la discusión sobre las formas

de conocimiento no representacionales y las habilidades necesarias en las prácticas tecnológicas.

Concepción sistémica de tecnología

La concepción sistémica de tecnología se refiere a la comprensión de la tecnología como un sistema complejo y dinámico compuesto por diferentes elementos interdependientes. En esta concepción, la tecnología no se ve simplemente como un conjunto de objetos o herramientas, sino como un conjunto de procesos, relaciones y sistemas que interactúan entre sí para cumplir una función específica. Desde una perspectiva sistémica, la tecnología se considera como una combinación de componentes físicos, como máquinas, herramientas y equipos, junto con procesos y sistemas organizacionales, como procedimientos, reglas y políticas. Además, la tecnología también puede incluir componentes sociales y culturales, como las actitudes y valores de los usuarios, así como las normas y estándares de la sociedad en la que se utiliza.

La concepción sistémica de la tecnología implica que cualquier cambio o mejora en uno de los componentes del sistema puede tener efectos en cascada en otros componentes del sistema. Por lo tanto, para comprender el impacto de la tecnología, es necesario considerar todos los componentes y su interrelación. De esta manera, se pueden identificar posibles problemas y soluciones, y se puede desarrollar una estrategia integral para la implementación y el uso de la tecnología.

Estas ideas han sido plasmadas por diferentes autores que vienen de la historia, la sociología y la filosofía de la tecnología. Entre ellos, Bijker et al. (1987), quienes afirman que la tecnología es un conjunto de componentes interconectados que funcionan juntos para lograr un propósito específico. Estos sistemas incluyen tanto aspectos técnicos (máquinas, dispositivos, infraestructura) como no técnicos (actores sociales, instituciones, políticas, regulaciones). Además, estos sistemas tecnológicos evolucionan a través del tiempo y se ven afectados por factores como la innovación, la competencia, la política y la economía. Por otro lado, Wiebe E. Bijker y Trevor Pinch son los denominados fundadores del enfoque de la construcción social de la tecnología (SCOT), pues en su obra *The Social Construction of Technological*

Systems (Bijker et al., 1987) destacan la importancia de analizar las tecnologías como sistemas sociales y técnicos interconectados. Esta obra reúne una serie de estudios de caso y ensayos teóricos que analizan cómo las tecnologías son construidas socialmente y cómo los sistemas tecnológicos son moldeados por factores sociales, culturales, económicos y políticos.

Otros autores cercanos a las ideas de tecnología como un sistema sociotécnico son los teóricos de la conocida teoría del actor red (ANT) que es un enfoque teórico y metodológico desarrollado en la década de 1980 por sociólogos como Latour (2008), Law (2009) y Callon (1986). La teoría del actor-red se ha utilizado ampliamente en los estudios de ciencia, tecnología y sociedad (STS) para analizar cómo los actores sociales y los objetos técnicos se relacionan e interactúan dentro de redes complejas (Monterroza-Ríos, 2017). La ANT sostiene que tanto los actores humanos como los no humanos (objetos técnicos, artefactos, sistemas, etc.) pueden desempeñar roles importantes en la formación y el mantenimiento de redes sociales y tecnológicas. En lugar de considerar a los objetos técnicos como meros productos de la actividad humana, la ANT reconoce que estos artefactos también pueden influir activamente en las acciones y relaciones de los actores humanos. La ANT se enfoca en el estudio de redes heterogéneas, que incluyen tanto actores humanos como no humanos y que pueden abarcar una amplia variedad de escalas y contextos. Estas redes son complejas, dinámicas y pueden cambiar con el tiempo a medida que los actores y las relaciones dentro de la red se transforman. La ANT enfatiza que las redes no son estáticas, sino que están sujetas a procesos de estabilización y cambio. Los actores y las relaciones dentro de una red pueden cambiar con el tiempo, lo que puede llevar a la formación de nuevas redes, la disolución de redes existentes o la reconfiguración de las relaciones y roles de los actores (Tirado y Domenech, 2005).

Esta forma de ver la tecnología tiene muchas ventajas y ha sido central en las reflexiones transdisciplinarias de los estudios de ciencia y tecnología; no obstante, su principal ventaja, que es su carácter constructivista, es también su propia debilidad, ya que hereda las críticas generales a los constructivismos. Entre las principales críticas tenemos que el constructivismo tiende a enfocarse en las dimensiones sociales y

culturales de la tecnología, lo que, salvo la teoría del actor-red, puede llevar a desatender los aspectos materiales y técnicos de la tecnología. Además, se señala que al centrarse en cómo la sociedad construye y moldea la tecnología, los constructivistas pueden subestimar la importancia de los aspectos técnicos en sí mismos. Por otro lado, las teorías constructivistas pueden caer en un determinismo social, atribuyendo demasiada influencia a las estructuras y actores sociales en la conformación de la tecnología. Al hacerlo, podrían subestimar el papel de la creatividad individual, la innovación y las decisiones técnicas en el desarrollo tecnológico. Por otro lado, algunos autores sostienen que las teorías constructivistas no abordan de manera suficientemente crítica las relaciones de poder y las desigualdades sociales en la tecnología, ya que argumentan que el constructivismo podría hacer más para examinar cómo las tecnologías pueden contribuir a reproducir y perpetuar dichas asimetrías (Haraway, 1995). Por otro lado, Ana Fernández Zubieta señala la ambivalencia epistemológica de los constructivismos al no proporcionar criterios claros para determinar qué evidencia es más sólida o qué métodos y teorías son más eficaces a la hora de explicar la ciencia y la tecnología, pues son muy buenos para deconstruir los fenómenos, pero no nos dan criterios para decidir cómo una teoría o desarrollo tecnológico puede estar mejor construido (Fernández Zubieta, 2009).

A pesar de las críticas, es la concepción de tecnología que mejor ha explicado hasta el momento las relaciones bidireccionales entre los diseños tecnológicos con los aspectos contextuales, ha mostrado su carácter complejo y evolutivo (Bijker et al., 1987), así como su importancia en mantener el orden social estable en el tiempo (Latour, 2008). También ha dado pie para que surja el llamado “giro empírico” en la filosofía de la tecnología y los estudios sociales de la ciencia y la tecnología, y que se refiere a un cambio en el enfoque desde la teoría abstracta y especulativa hacia investigaciones empíricas y basadas en la evidencia. Este cambio fue impulsado por la necesidad de conectar las reflexiones filosóficas sobre la tecnología con situaciones y problemas concretos del mundo real (Ihde, *Technology and the Lifeworld: From Garden to Earth*, 1990).

Posibilismo tecnológico, una respuesta hispanohablante al determinismo

En esta sección vamos a proponer una etiqueta inspirada en diversas reflexiones de filósofos hispanohablantes de la tecnología que, en términos generales, consideran la tecnología como una práctica que reúne un conjunto de condiciones pragmáticas para cambiar las condiciones reales de posibilidad de un grupo humano, con sus consecuencias abiertas a futuros posibles, no como un destino inevitable seguido por un solo propósito en la historia. Por lo tanto, “posibilismo” es un término de contrapeso para el término “determinismo”, que es una etiqueta común en los estudios filosóficos de la tecnología. El “posibilismo” considera la filosofía de la tecnología como una rama de la filosofía de la acción (no de la filosofía de la ciencia), esto es, porque la tecnología es un tipo especial de acción que tiene una estructura intencional compleja, realizada por un sujeto colectivo y heterogéneo, y que da como resultado entidades (artefactos) que también tienen niveles de realización ontológica heterogéneas (Broncano, 2006; Monterroza, 2018). Por lo tanto, en esta sección vamos a describir, en primer lugar, cuáles han sido las ideas generales que se asocian al determinismo tecnológico, tanto sus versiones pesimistas como sus versiones optimistas. Una vez hecha esta caracterización, presentaremos lo que consideramos una concepción posibilista de la tecnología.

El determinismo tecnológico sostiene que el desarrollo tecnológico es el principal motor del cambio social, económico y cultural en una sociedad (Mumford, 1934; Ellul, 1954). Según esta perspectiva, la implementación de nuevas tecnologías y sus aplicaciones en diferentes ámbitos de la vida determina en gran medida la evolución de las sociedades humanas, influyendo en aspectos como la economía, la política, las relaciones sociales y la cultura (Marx, 1867). Esta manera de concebir la tecnología supone que su dinámica es autónoma, pues su desarrollo sigue una lógica interna, llevando a la sociedad a adaptarse a ella (Ellul, 1954). Bajo esta idea, el cambio tecnológico es inevitable porque sigue sus fines y valores (normalmente de eficacia y eficiencia) independientes de la mayoría de las dinámicas sociales. Existen formas de determinismo tecnológico tanto optimistas (Kurzweil, 2005; Kelly,

2016; Toffler, 1970) como pesimistas (Mumford, 1934; Ellul, 1954). Los optimistas, a veces llamados “futuristas”, a menudo ven la tecnología como una solución a los problemas sociales y económicos. Argumentan que la innovación tecnológica puede mejorar la calidad de vida y resolver problemas como la pobreza, la enfermedad y la desigualdad, que son deseables por supuesto; sin embargo, olvidan que la historia de la tecnología nos muestra que a cada problema resuelto por medios técnicos surgen nuevos problemas inesperados. Por otro lado, el determinismo pesimista, muy popular en las humanidades, la literatura y el arte, ven la tecnología como una fuerza que puede tener consecuencias negativas para la sociedad, como la pérdida de empleo, la erosión de la privacidad, la desigualdad y la alienación. Son como los viejos profetas que advierten que la tecnología se ha vuelto una *hubris* sobre la que ya nada o poco se puede hacer frente al destino. Por supuesto, el determinismo tecnológico es una etiqueta externa; ningún pesimista u optimista tecnológico va a considerarse a sí mismo como determinista.

Las concepciones deterministas de la tecnología han sido fuente de críticas en varios sus aspectos, pues tienden a restar importancia a la agencia humana en el proceso de cambio tecnológico. Al concebir la tecnología como una fuerza autónoma, los deterministas tienden a minimizar el papel de la agencia humana en la orientación del cambio tecnológico, pues suponen que la aparición y adopción de nuevas tecnologías son inevitables y que la sociedad no tiene más opción que adaptarse a ellas. Sin embargo, los críticos sostienen que esta visión de la inevitabilidad puede conducir a una actitud pasiva y fatalista, en lugar de fomentar la participación activa y el debate sobre cómo y cuándo se deben desarrollar y adoptar nuevas derivas técnicas (Feenberg, 1991; Broncano, 2005; García Palacios et al., 2001).

La etiqueta de “posibilismo” que proponemos no es solo la contraparte del determinismo, sino que puede pensar la tecnología como una intersección de posibilidades que abren nuevos horizontes de acción e imaginación que no estaban allí sin los productos de la acción tecnológica. Así, un camino abre la posibilidad de ser transitado y abrir la comunicación y el comercio entre grupos distintos; un conjunto de habitáculos abre las posibilidades de ser un nodo de vivienda, de exploración o de explotación de los recursos

naturales alrededor del mismo poblado, etc. El planteamiento posibilista es el siguiente: el determinismo tecnológico supone la existencia de una “esencia” humana ahistórica que puede ser ocultada (pesimismo determinista) o potenciada (optimismo determinista) por la tecnología que avanza de manera irrestricta impulsada por la competencia económica o militar. No obstante, no hay tal esencia primigenia humana, hay efectos de las trayectorias culturales contingentes que han hecho del ser humano, “humano”; es el producto híbrido entre naturaleza y artificio que se ha hecho a sí mismo por sus propias acciones técnicas. Es una alusión a la ya clásica interpretación orteguiana de que “el hombre no tiene esencia sino historia” (Ortega y Gasset, 1914).

Afirmamos que esta puede ser una contribución de la filosofía de la técnica hispanohablante porque estas ideas se encuentran en la tradición orteguiana cuando señala en las *Meditaciones de la técnica* que con la técnica creamos posibilidades completamente nuevas al instaurar objetos que no hay en la naturaleza del hombre (Ortega y Gasset, 1982, sección III). El considerar la técnica como una forma de acción para abrir nuevas posibilidades reales es una idea implícita que está presente en autores diversos de habla hispana como Zubiri (1962; 2010), Quintanilla (2017), Broncano (2001), Liz (2002), Olivé (2000), Echeverría (1999) y Lawler y Vega (2009). Cada autor y obra enfatiza en diversos aspectos ontológicos, epistemológicos y axiológicos de la técnica pero que, en nuestro concepto, comparten cierta concepción de técnica como constitutiva de la realidad humana abierta, en que el ser humano se hace a sí mismo a través de la técnica. No es que otros autores de otras tradiciones no lo hayan considerado de esta manera (Simondon, 2007; Stiegler, 1998; Leroi-Gourhan, 1971), pero no se cierran solo a las reflexiones críticas como algunos autores cercanos a la idea de tecnología autónoma (Mumford, 1934; Ellul, 1954).

De todos estos autores, es Fernando Broncano quien de manera explícita desarrolla la idea posibilista de tecnología, que se observa en sus diferentes obras, tales como *Mundos artificiales* (2001), *Entre ingenieros y ciudadanos* (2006), *La melancolía del ciborg* (2009) y *La estrategia del simbiote* (2012), además de múltiples artículos y colaboraciones:

Mi propuesta es que desarrollemos todo lo que hay de posibilista en el dominio de la técnica, alejándonos de las formas ocultas de determinismo cultural que subyacen a muchas de estas discusiones. En esta concepción modalmente posibilista, la tecnología es un modo de transformar colectivamente el presente, sujetándose a constricciones normativas que nacen ya desde dentro de la propia naturaleza de las acciones tecnológicas. En esta perspectiva, la imaginación de lo posible no es un pensamiento ajeno al que la acción técnica se sometería como una racionalidad que se atiene a fines, sino que formaría parte de la trama misma del pensar lo técnico como pensar posibilidades genuinas. (Broncano, 2006, p. 58)

El posibilismo planteado por Broncano se basa en una premisa, en principio trivial: el futuro se modifica con cada ejercicio de acción libre, pues al realizarse una acción suceden cambios infinitesimales que hacen bifurcaciones de las trayectorias futuras; es posibilista porque solo el pasado está escrito, no hay rumbos determinados en la historia más que en las consecuencias de las acciones del presente. En este sentido, la tecnología es una subclase de acción intencional colectiva o, más precisamente, una forma de agencia. Los diseños tecnológicos participan en la creación de futuros posibles (p. 69).

La concepción posibilista supone que la acción tecnológica se describe por ciertas cualidades que están presentes en gran parte de los procesos de diseño tecnológico. Así, un diseño tecnológico es una respuesta “novedosa” y “eficiente” a un “problema” práctico; no obstante, hay que examinar con cuidado los conceptos de problema, novedad y eficiencia, pues no tienen el carácter común presentes en las críticas culturales a la tecnología. Veamos cómo estos tres conceptos se entrelazan entre sí para describir una buena parte de los procesos de diseño tecnológico contemporáneo (Broncano, 2006, pp. 88-91):

1. Problema: la categoría problema supone que existen alternativas a las situaciones actuales, que el curso de acción de los acontecimientos puede ser de otra manera. Los animales no reflexivos no tienen problemas, tienen reacciones, pues están sujetos al flujo causal de su existencia. Justamente gran parte

de la investigación tecnológica es encontrar problemas, es decir, situaciones mal planteadas que pueden tener cursos de acción alternativos.

2. Novedad: la novedad es fundamental en la concepción posibilista de tecnología, pues sugiere explorar posibilidades dentro de recursos limitados (tiempo, materiales, de conocimiento, recursos económicos, etc.). A diferencia de los determinismos e instrumentalismos que asocian las tecnologías contemporáneas a la repetición y al tedio, la concepción posibilista de tecnología considera que las actividades de diseño son justamente la exploración de posibilidades novedosas a la luz de los problemas identificados previamente.
3. Eficiencia: este concepto no tiene que ver con el valor taylorista de productividad. Para el posibilismo, la eficiencia es que al inducir una trayectoria histórica nueva no se modifiquen todas las demás adyacentes. En términos más sencillos, el ideal de una eficiencia total de un diseño tecnológico es que pueda solucionar un problema sin generar otros. Por supuesto esto no se alcanza en una situación real, no obstante, es un parámetro y cualidad deseable en cualquier práctica tecnológica.

Observamos el esquema (figura 2) de una posible trayectoria histórica de un grupo humano. Un nuevo diseño tecnológico puede cambiar las condiciones de posibilidad que cambien dicha trayectoria. Como lo mencionamos desde un principio, este diseño puede solucionar un problema a la luz de un momento histórico, pero es posible que abra otros. En todo caso, la tendencia de posibilidades de acción para el grupo humano no será la misma si no hubiera adoptado un nuevo diseño tecnológico que cambie sus posibilidades de acción, así como sus expectativas y sus nuevos problemas. En ese sentido, se crean nuevos valores y fines como componente esencial del modo en que los seres humanos hacen su historia, muchas veces bajo las condiciones que ellos mismos crean y a veces por las consecuencias indeseadas de esas acciones (Broncano, 2006, p. 69). Así, concebir la tecnología de esta

manera es una forma de acción social en la que, al elaborar planes o proyectos, esas posibilidades se ven más cercanas y, por lo mismo, surge un momento reflexivo sobre la acción que harían reales esas posibilidades una vez que sabemos que existen las capacidades necesarias para llevarlas a cabo.

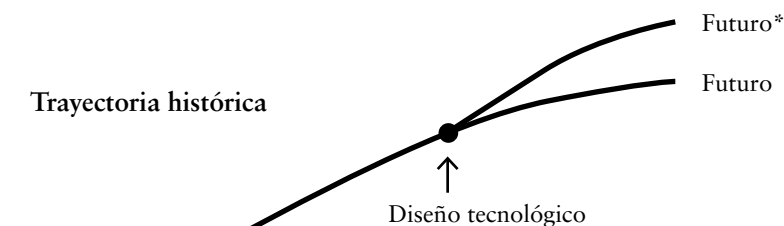


Figura 2. La trayectoria histórica de un grupo humano cambia cuando se apropia de un nuevo diseño tecnológico

Fuente: elaboración propia.

Para llevar a cabo estos fines y propósitos se debe tener una capacidad agente lo suficientemente efectiva para hacer confluir diversos espacios de posibilidad que transcurren en las vidas humanas. Broncano propone en *Mundos artificiales* (2001) estos espacios de posibilidad (físico, conceptual, pragmático y legítimo) que en su intersección conforman las posibilidades reales y genuinas que crean un diseño tecnológico. En la figura 3 se muestra una manera gráfica de entender esta confluencia de espacios para generar nuevas posibilidades; en principio, lo que se propone puede parecer abstracto, pero al examinar la figura 3 y la tabla 1 se concretizan ejemplos particulares de lo que se quiere decir conforme a la visión posibilista de la tecnología.

La tecnología ocupa el espacio de intersección de todas estas posibilidades, es decir, lo que hace que lo (1) imaginado y conceptualmente

posible se intercepte con (2) lo físicamente posible, en un marco de (3) lo pragmático y (4) legítimamente posible para crear nuevas posibilidades reales a ciertos grupos humanos. Los procesos de diseño tecnológico hacen real lo que es posiblemente factible a la luz de ciertos propósitos prácticos y fines humanos. La creatividad técnica es la capacidad de encontrar una intersección de todas las posibilidades mencionadas (que a veces no es posible) bajo una cantidad considerable de restricciones físicas, imaginativas, conceptuales, económicas, jurídicas y morales.

Además, la intersección de las posibilidades (las posibilidades genuinas que crean los diseños tecnológicos) se expande y reconfigura con cada nuevo desarrollo tecnológico, en el sentido de que una tecnología abre el horizonte de posibilidades para nuevas tecnologías, y así continuamente. Por ejemplo, la electrónica de transistores posibilita los circuitos integrados, que posibilitan a la vez los microprocesadores; en esa misma cadena, los microprocesadores posibilitan los circuitos programables, que posibilitan a la vez el *software* y así hasta que se agota o se reemplaza una trayectoria tecnológica.

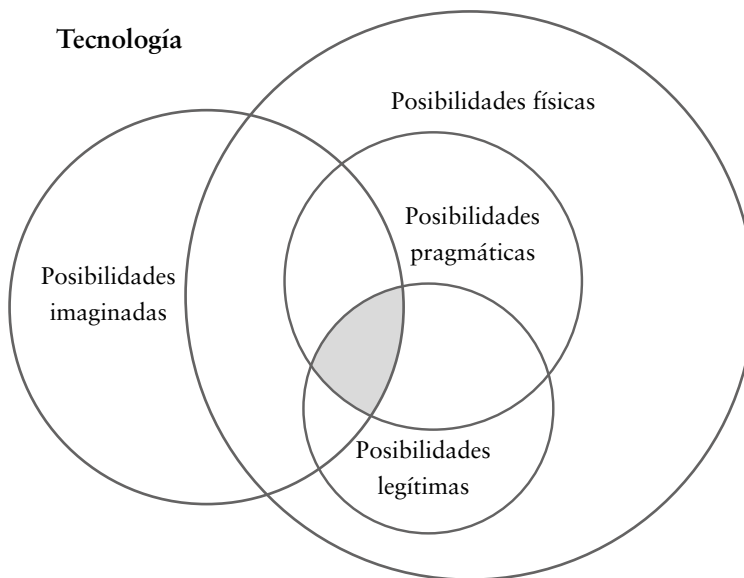


Figura 3. Tecnología como intersección de posibilidades

Fuente: basada en Broncano (2001, p. 122).

En la tabla 1 se describe a qué se refieren cada uno de los espacios de la figura 3, por qué los diseños tecnológicos suceden en esa intercepción de posibilidades y por qué ese espacio (gris) no es estático, sino que cambia tanto espacial como temporalmente.

Tabla 1. Tipos de posibilidades que se tiene en cuenta en un diseño tecnológico

Tipo de posibilidad	¿Qué es posible?	¿Cómo se expanden estas posibilidades?
Posibilidades físicas	Es todo lo físicamente posible, la naturaleza tiene una estructura causal que responde a ciertas leyes que hacen lo físicamente posible. Una intervención técnica al mundo (un artefacto, una máquina, una infraestructura) debe responder a las leyes físicas, químicas, biológicas, etc. Un desarrollo tecnológico debe estar necesariamente en lo físicamente posible, si se diseña un sistema de acueducto, se debe tener en cuenta el peso del agua y las alturas de las tuberías sobre la superficie de la tierra debido al efecto gravitatorio del planeta. Hay restricciones físicas que no se pueden obviar.	Las ciencias naturales, las matemáticas y afines crean nuevo conocimiento sobre el mundo físico, químico, biológico, etc., y sus propiedades que hace que el universo de lo físicamente posible se expanda, pues se conocen más propiedades y relaciones de las entidades físicas que sirven de caldo de cultivo para la innovación tecnológica.
Posibilidades imaginadas	Aunque nuestra imaginación está alimentada por el mundo físico donde vivimos, la literatura nos muestra que las posibilidades imaginadas no están necesariamente atadas a las posibilidades físicas, pero sí a lo conceptualmente posible, ya que con los conceptos colectivos del lenguaje hacemos nuevos mundos posibles. El sueño de volar, la inmortalidad, los viajes en el tiempo, etc., nos muestran que la imaginación estimula los límites de la creación técnica para hacer posible lo imaginado.	El espacio de posibilidades imaginadas y conceptuales se expande con los mismos resultados de las acciones tecnológicas, así como con el desarrollo de las ciencias y las matemáticas. Además, las artes, la literatura y el cine también contribuyen a reconfigurar los futuros posibles y sus consecuencias.

Tipo de posibilidad	¿Qué es posible?	¿Cómo se expanden estas posibilidades?
Posibilidades pragmáticas	Constituyen un subconjunto de las posibilidades físicas que, además de respetar las leyes y restricciones materiales, están condicionadas por marcos pragmáticos propios de contextos sociales e históricos específicos. Estas posibilidades dependen de variables como la organización social, las estructuras económicas y los modos de producción disponibles en un determinado estadio cultural. En consecuencia, no todo lo físicamente posible resulta pragmáticamente realizable. Por ejemplo, aunque desde el punto de vista físico sería viable construir un edificio íntegramente en oro, dicha opción no es pragmáticamente posible debido a limitaciones de recursos, costos de extracción, viabilidad técnica y condiciones constructivas. De este modo, la noción de posibilidad pragmática introduce un criterio de factibilidad situado que restringe el dominio de lo físicamente posible.	El espacio de posibilidades pragmático crece con el desarrollo de nuevas tecnologías, pues amplía el horizonte de posibilidades y recursos técnicos, económicos y habilidades prácticas que crean estas tecnologías.
Posibilidades legítimas	Existe otro subconjunto de las posibilidades físicas que corresponde de lo legítimamente posible, esto es, que depende de nuestras convicciones morales y las convenciones legales. Son fundamentales y tienen que ver con lo que la ética de la tecnología nos sugiere, no porque podamos hacerlo, debemos hacerlo. Hay muchos casos como por ejemplo la clonación de humanos, el diseño genético de bebés, etc., que son casos de cosas que física y técnicamente se puede hacer, pero no se debe hacer por los riesgos y consecuencias imprevistas.	La forma del espacio de posibilidades legítimas está relacionada con los cambios en los valores de un grupo humano a través del tiempo. Lo que nos parece correcto y legítimo no es igual en cada época. El desarrollo de ciertas tecnologías como las TIC ha abierto la comunicación global reconfigurando los valores de cierto grupo humano debido a las influencias culturales externas.

Fuente: elaboración propia.

Esta manera abierta de concebir la tecnología puede estar en mejor sintonía con muchas de las propuestas de control social de la tecnología que han venido de múltiples tradiciones como los estudios CTS (García Palacios et al., 2001; Moreno-Ortiz y Vinck, 2021), la teoría crítica de la tecnología de Feenberg (Feenberg, 1991), el diseño socialmente responsable (Papanek, 1977) o incluso las iniciativas multilaterales como los Objetivos de Desarrollo Sostenible (ODS) que fueron adoptados por la Asamblea General de las Naciones Unidas en 2015 como parte de la Agenda 2030 para el Desarrollo Sostenible (ONU, 2015).

Tecnología a la luz del contexto de diseño y contexto de uso

¿Qué es diseñar?

La capacidad de diseño técnico ha sido uno de los rasgos humanos más característicos desde sus orígenes. En efecto, al ser humano se le conoce tanto como *Homo sapiens* como con *Homo faber* (Ortega y Gasset, 1982) por esa facultad de intervenir materialmente su realidad inmediata gracias a sus acciones técnicas mediadas por su plasticidad mental (Zawidzki, 2018) ya que permite imaginar futuros posibles en los cuales muchas veces sus deseos sobrepasan sus posibilidades prácticas (Ortega y Gasset, 1982). Por lo tanto, las capacidades de diseñar, esto es de proyectar sus acciones y evaluar sus resultados, han estado presentes en todos los grupos humanos manifestados en sus diversas hazañas técnicas. En ese sentido, todas las culturas, civilizaciones y grupos humanos han tenido capacidades de diseño, que se manifiestan en las habilidades y conocimientos técnicos de personas especializadas en un determinado oficio; estas personas, que llamamos artesanos, aprendieron sus oficios de manera práctica a través del entrenamiento y la repetición en talleres con otros artesanos. Por decirlo de una manera sintética: las capacidades de diseño en la producción artesanal se encontraban de manera implícita en la memoria mental y corporal de los artesanos.

El artesano tradicional ha sido el principal reproductor de formas y funciones de los artefactos en la mayor parte de la historia de la humanidad después de Neolítico. Por su propia constitución, en los procesos artesanales, las innovaciones han sido pausadas y resultado de un trabajo paulatino de desarrollo de habilidades por entrenamiento de maestro a aprendiz. El artesano es habilidoso en la reproducción de formas que han tenido éxito previamente, y ha aprendido a hacerlo al observar la producción de otros ejemplares similares. Las variaciones en los diseños artesanales son mínimas y son resultado del conocimiento operacional adquirido por el artesano o de las demandas de los usuarios de los objetos que produce. La mente del artesano funciona principalmente como un reproductor de formas, no como un creador, a excepción de los elementos creativos que surgen ocasionalmente (Broncano, 2006, p. 71).

El separar la concepción de los artefactos (con el uso de medios representacionales) de la fabricación (industrializada) va a ser un hito fundamental que cambia la forma de hacer acciones técnicas e inaugurando lo que ahora se conoce como tecnología. El separar la proyección de la producción (Gay y Samar, 2007) trae enormes ventajas en la manera de optimizar los procesos productivos, en especial, cuando los artefactos por construir son grandes y costosos, o que se requiere una optimización de los materiales o controlar mejor la calidad. Por esta razón, la tecnología industrializada tiene en su constitución la ejecución de planes de diseño llevados a cabo con el apoyo de procesos de diseño representacional. Lo que consideramos de manera disciplinar como diseño ingenieril en la actualidad está relacionado con las técnicas de representación apoyadas en medios materiales (bocetos, planos, prototipos, etc.) para proyectar las acciones técnicas. El diseño implica un lenguaje simbólico abstracto de dibujos y representaciones gráficas que permiten varias transformaciones distintas a las técnicas artesanales tradicionales. Esto trajo múltiples cambios que menciona Fernando Broncano (2001, p. 97) que acá se explican y amplían de manera sintética:

- En la producción artesanal tradicional se trabaja directamente con los materiales, mientras que en los procesos de diseño se hace con representaciones (bocetos, planos, prototipos, etc.).

- En la producción artesanal se hace un proceso de ajuste y acoplamiento pieza por pieza mientras que en los procesos mediados por el diseño es posible la preconfiguración de las piezas exactas.
- En los procesos mediados por el diseño se puede modificar una y otra vez las representaciones, mientras que en la producción artesanal se puede hacer solo en el proceso de ensamble del artefacto.
- En los procesos de diseño exigen otro tipo de conocimientos más allá de las habilidades propias de los procesos artesanales.
- En los procesos de diseño, al usar representaciones normalizadas cuantitativamente, se pueden usar operaciones matemáticas y conocimiento científico para predecir o evaluar propiedades deseables, que es algo que no se puede lograr con la producción artesanal.
- En los procesos de diseño se fomenta las novedades al poder trabajar sobre representaciones con las que es más fácil de hacer modificaciones, aprovechando los conocimientos y materiales de los procesos de diseño anteriores.
- Adicionalmente, los procesos de diseño basados en medios representacionales permiten que dicho proceso sea colectivo, esto es, que no dependan de una sola persona y, por lo tanto, posibilitan la inteligencia colectiva que multiplica la creatividad técnica para resolver los problemas del proceso.

Por esta razón, existe una relación tan estrecha entre la industrialización y los procesos de diseño en planos, pues son corresponsables. Dicho con una analogía: *los procesos de diseño son a la acción técnica lo que la escritura fue para el lenguaje hablado*. El poder apoyar las acciones sobre medios representacionales multiplicó las capacidades efectivas de la acción técnica y sus alcances. El gran desarrollo tecnológico que tuvo lugar después de las revoluciones industriales (y

sus impactos sociales, culturales, económicos, morales, etc.) ha sido posible gracias a la implementación metódica de procesos de diseño.

Los procesos de diseño han establecido dos ámbitos de comprensión del desarrollo tecnológico: el contexto de diseño y el contexto de uso. Son dos unidades de estudio para la comprensión del desarrollo y la implementación de tecnologías, ya que ambos influyen en cómo se conciben, diseñan, implementan y utilizan las soluciones tecnológicas.

La tecnología misma, como un sistema de acciones y prácticas, se puede mirar desde el contexto de diseño, así como desde el contexto de uso. En el contexto de diseño, se establecen los propósitos, fines, restricciones, normas tácitas y valores de lo que se quiere reconfigurar en un paisaje de cultura material. En el contexto de uso, en el que se reconfiguran las condiciones de posibilidad a un grupo de usuarios, se puede comprender cómo sus expectativas y posibilidades de acción pueden abrir nuevas trayectorias de diseño. Nuestro concepto de tecnología puede explorarse desde el proceso de diseño para convertir las posibilidades imaginadas en posibilidades reales, hasta la adopción y apropiación de artefactos y sistemas tecnológicos que cambian elementos concretos de las posibilidades pragmáticas para una comunidad o grupo de usuarios.

El contexto de diseño

El contexto de diseño se refiere al entorno y las condiciones en las que los diseñadores y desarrolladores trabajan para crear una solución tecnológica. Este contexto puede incluir factores tanto internos como externos que influyen en la concepción, el desarrollo y la implementación de un sistema tecnológico.

En la actualidad, la concepción clásica de diseño ha cambiado significativamente: ha pasado de entenderse, desde sus orígenes, como herramienta de proyección basada en técnicas de dibujo técnico, a concebirse como una disciplina que trasciende la mera producción y se configura como una forma de pensamiento orientada al diseño de condiciones de posibilidad para diferentes grupos de usuarios. En este sentido, organizaciones internacionales sobre diseño, como la *World Design Organization* (WDO), definen el diseño

como “un proceso estratégico de resolución de problemas que conduce a una mejor calidad de vida [...] a través de productos, sistemas, servicios y experiencias” (WDO, 2022). Esta definición muestra varias cosas: la primera es que el diseño es un “proceso estratégico”, lo que significa que se parte de una situación en la que existe un problema y a través de un proceso iterativo se van planteando y perfeccionando posibles soluciones, proceso el cual es continuamente retroalimentado por los resultados parciales evaluados por el mismo equipo de diseño o los posibles usuarios. El siguiente elemento clave es que se pueden diseñar varias cosas y no solo objetos “a secas”, ya sean productos, sistemas, estrategias, servicios o experiencias, pero estos resultados no se plasman en el aire, sino con elementos materiales concretos que son proyectados, modelados, fabricados y puestos a prueba en el contexto de uso adecuado. Otro elemento importante es la novedad, esto es, las soluciones propuestas (aunque no sean completamente originales) deben mostrar derivas creativas a lo que ya existe, ya sea con otras propiedades o en otros contextos, susceptibles siempre de ser mejorados, de modo que la innovación es un valor fundamental en los procesos de diseño contemporáneo. Un elemento adicional es que los procesos de diseño llevan, en la mayoría de los casos, procesos de investigación, esto es, un proceso que requiere los estadios propios de una investigación aplicada, tales como la revisión de antecedentes, el análisis del contexto y los actores involucrados en el fenómeno, la descripción de una metodología clara de trabajo, el uso de instrumentos y técnicas adecuados, el análisis y la evaluación de los resultados. Esto también supone un trabajo colectivo entre varios profesionales del diseño involucrando otras disciplinas dependiendo de la naturaleza del fenómeno (Domínguez-Rendón y Monterroza-Ríos, 2019, p. 78).

Algunos elementos clave del contexto de diseño incluyen factores de distintos tipos, tales como sociales, culturales, económicos, tecnológicos, legales, entre otros. Acá se presenta la lista de algunos de ellos que son tenidos en cuenta a la hora de iniciar un proceso de diseño de un nuevo desarrollo tecnológico desde el contexto de diseño (Goodwin, 2009):

- Necesidades y requisitos del usuario: comprender las necesidades, deseos y limitaciones de los usuarios potenciales es fundamental para diseñar una solución.
- Competencia y mercado: analizar el entorno competitivo y las tendencias del mercado ayuda a identificar oportunidades, amenazas y posibles puntos de diferenciación.
- Restricciones técnicas y recursos: evaluar las limitaciones técnicas, así como los recursos disponibles (humanos, financieros, materiales) para desarrollar el producto.
- Normativa y regulaciones: entender las leyes, regulaciones y estándares aplicables al producto o servicio, ya que pueden afectar su diseño, implementación y comercialización.
- Costumbres y valores: considerar las diferencias culturales y los valores de los usuarios para asegurar que el diseño sea apropiado y respetuoso con las diversas sensibilidades.

Ahora, desde el contexto de diseño, la tecnología es un sistema heterogéneo de acciones para transformar causalmente la realidad. Quintanilla (2017) es uno de los filósofos que más ha insistido en el carácter sistémico y heterogéneo de la tecnología. Para el autor, es la forma de acción técnica contemporánea por excelencia, es decir, como un sistema de acciones colectivas apoyado en (y guiado por) andamios culturales tangibles (máquinas, herramientas, imágenes, planos, etc.) e intangibles (conocimientos, habilidades, normas, valores, etc.). Los fines de estas intervenciones están sujetos a marcos sociales tales como políticas públicas, planes de desarrollos, normas jurídicas, intereses privados corporativos, necesidades sociales (reales o percibidas) de acuerdo con los valores y formas de gobernanza de una sociedad (Monterroza-Ríos y Escobar-Gómez, 2021).

En la figura 4 observamos cómo, desde el contexto de diseño, la acción humana cooperativa y organizacional de equipos de diseño (ingenieros, técnicos, operarios, administradores, etc.) dispone de un conjunto de artefactos, conocimientos, habilidades y aptitudes para llevar a cabo una solución (en forma de producto, servicio, proceso, sistema, etc.) para cambiar el espacio de posibilidades del grupo de usuarios objetivo bajo un conjunto de marcos sociales, normativos y valorativos.

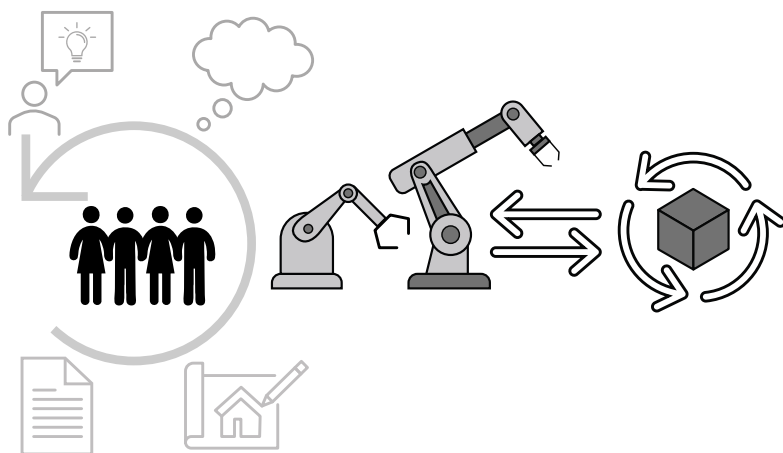


Figura 4. Desde el contexto de diseño, tecnología es un sistema de prácticas para la transformación de la realidad constituido por agentes, artefactos, conocimientos, normas y valores

Fuente: basado en Quintanilla (2017).

El contexto de uso

El contexto de uso se refiere al entorno y las condiciones en las que los usuarios finales interactúan y utilizan un artefacto o sistema tecnológico. Este contexto puede variar ampliamente según el tipo de tecnología, el propósito y los usuarios.

Desde el contexto de uso, la tecnología puede entenderse como parte de los resultados de un proceso de diseño que transforma el horizonte de posibilidades pragmáticas para un grupo humano, sea

como intervenciones físicas o como intervenciones simbólicas. En la “intervención física” de la realidad, los procesos dejan como resultado artefactos (obras públicas, objetos de uso, edificaciones, sistemas informáticos, etc.) que instauran una novedad en la cultura material y establece nuevas condiciones de posibilidad y cambios en algunas prácticas culturales. En la “intervención simbólica” se diseñan nuevos significantes, significados, estándares, procesos o modelos de gestión que, al igual que la intervención física, instauran una novedad en la cultura material que establece nuevas condiciones de posibilidad y cambios en algunas prácticas culturales. Tanto las transformaciones físicas como las simbólicas se hacen necesariamente con elementos materiales, pues solo de esta manera se cambia el horizonte real de posibilidades y se mantienen en el tiempo (Broncano, 2012; Latour, 2008). Incluso los elementos simbólicos requieren referentes materiales que evocan pensamientos, emociones y experiencias previas; pensemos en los objetos sagrados o de culto, como lo dice Lambros Malafouris, pensamos involucrados con las cosas (Malafouris, 2013).

En este sentido, los procesos mismos de diseño tienen en cuenta una serie de aspectos desde el contexto de uso que se han vuelto fundamentales en el desarrollo de nuevas tecnologías, como lo señalan algunos autores (Norman, 2005):

- Entorno físico: las condiciones ambientales, como la ubicación, la iluminación, el ruido, el clima y la ergonomía pueden afectar la experiencia del usuario y el rendimiento de un producto de diseño.
- Entorno social: las interacciones con otras personas, ya sean usuarios o no, así como las normas sociales y las expectativas de comportamiento, pueden influir en cómo se utiliza y se perciben los productos de diseño.
- Habilidades y conocimientos del usuario: las capacidades, habilidades y conocimientos previos de los usuarios pueden

afectar su capacidad para utilizar y aprovechar una intervención tecnológica.

- Tareas y objetivos: las metas y actividades específicas que los usuarios quieren lograr con el producto influyen en cómo lo utilizan y en qué funcionalidades priorizan.
- Dispositivos y sistemas complementarios: otros dispositivos y sistemas con los que la tecnología debe interactuar o integrarse para funcionar correctamente.

En la figura 5 señala cómo, desde el contexto de uso, los resultados de los procesos de diseño que son implementados o apropiados por una comunidad o grupo de usuarios cambian el paisaje de posibilidades de acción que puede transformar las prácticas cotidianas mismas.

Una de las perspectivas que más ha señalado el papel del uso en la dirección de las trayectorias tecnológicas es el programa SCOT (Bijker et al., 1987) al destacar cómo los usuarios co-construyen activamente los artefactos tecnológicos a través de sus prácticas, adaptaciones y apropiaciones. Los usuarios no son meros receptores pasivos de productos de diseño, sino que desempeñan un papel activo en su configuración y uso. Los teóricos de programa SCOT sostienen que las tecnologías pueden tener múltiples significados e interpretaciones en función de las experiencias, expectativas y necesidades de los diferentes grupos de usuarios (flexibilidad interpretativa). Estas interpretaciones flexibles pueden llevar a diferentes patrones de uso y a la adaptación de la tecnología en función de las necesidades de cada grupo. Además, identifican a los “grupos relevantes” como aquellos actores sociales que tienen un interés en una tecnología en particular y que influyen en su desarrollo, implementación y uso. Estos grupos pueden incluir usuarios, fabricantes, reguladores, entre otros. SCOT sugiere que el análisis de las interacciones y negociaciones entre estos grupos relevantes es fundamental para comprender cómo se construye y utilizan los sistemas sociotécnicos.

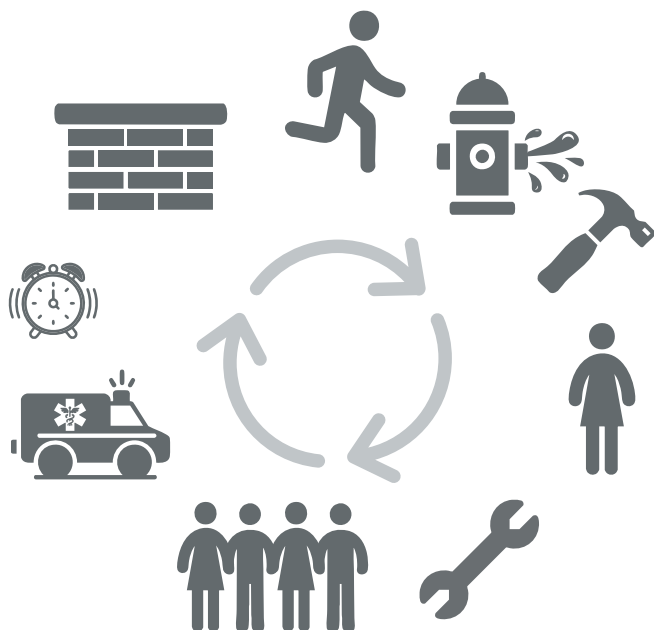


Figura 5. Tecnología desde el contexto de uso. Los productos de las prácticas tecnológicas ofrecen nuevas posibilidades pragmáticas que cambian expectativas

Fuente: elaboración propia.

Desde los estudios de diseño contemporáneo, se tiene claro que existe una conexión fundamental entre el contexto de diseño con el contexto de uso; no tiene sentido concebirllos de manera aislada, pues, en efecto, las nuevas metodologías y principios de diseño (Krippendorff, 2006) lo que buscan es que haya un solapamiento entre estos.

¿Cómo y por qué cambian las tecnologías?

Una de las características de la acción técnica y las modificaciones de la cultura material es que, hasta cierto punto, tienen un carácter acumulativo en el sentido de que los nuevos objetos, registros y modificaciones materiales sirven de repositorio objetivo de memoria colectiva para un grupo humano (Stiegler, 1998). Es por esta razón que un arqueólogo puede reconstruir una buena parte de las prácticas cotidianas de un

colectivo con la interpretación de las relaciones entre los objetos que sobreviven al paso del tiempo. Como lo hemos mencionado, el carácter parcialmente acumulativo de los elementos de una cultura material permite al grupo que habita en ella abrir horizontes de posibilidades de acción que no serían posibles si no estuvieran esos artefactos allí. La cultura material, al ser materialmente objetiva, puede perdurar en el tiempo para que las próximas generaciones no empiecen de cero, pues su aprendizaje de prácticas estará apoyado no solo en las generaciones anteriores de personas, sino también en el conjunto acumulado de útiles, herramientas, artefactos cognitivos y símbolos materiales que están disponibles en dicho entorno cultural.

Durante casi toda la historia de la humanidad, ha existido una evolución relativamente parsimoniosa de las diversas técnicas y sus artefactos posibilitadores; no obstante, como ya lo mencionamos, hubo una disrupción en la manera de producir bienes apoyadas en nuevas fuentes de energía, en la división del trabajo, en la transformación de talleres a fábricas y, en especial, en la separación de la concepción de los objetos de su fabricación (Gay y Samar, 2007), es decir, empezó la industrialización de los sistemas de producción. Desde entonces, el cambio técnico ha sido mucho más vertiginoso, pues se mezclan en una relación multidireccional entre el desarrollo tecnológico con el cambio social, con el desarrollo económico y con la misma investigación científica. Esto es, sabemos que el desarrollo económico influye el cambio técnico, pero también lo contrario; sabemos que el desarrollo de la ciencia impulsa el desarrollo tecnológico, pero también lo contrario; además, sabemos que el cambio técnico también influye el cambio social, pero las sociedades pueden impulsar o no un tipo determinado de tecnología. Además, los factores económicos, sociales y políticos están interrelacionados entre sí. Esto nos deja una malla compleja de relaciones en donde no sabemos de manera clara cómo repartir los efectos y las causas (figura 6).

La tecnología tiene un carácter contingente, esto es, “contingente” porque refiere a algo que no es necesario, que no es esencial o que no está determinado por la naturaleza de las cosas, sino que depende de circunstancias particulares. En el contexto de la tecnología, la contingencia se refiere al hecho de que las decisiones y las acciones relacionadas

con la tecnología no están determinadas únicamente por la naturaleza de la tecnología en sí misma, sino que dependen de las circunstancias sociales, políticas, culturales y económicas en las que se desarrollan. Por este motivo, incluso pequeños cambios en las condiciones iniciales de una sociedad pueden cambiar fuertemente una trayectoria tecnológica. Así, decisiones políticas, cambios económicos, nuevos descubrimientos científicos, epidemias, desastres naturales o cambios en los valores de una sociedad pueden favorecer el florecimiento de una tecnología y el desuso o la obsolescencia de otra. Esto hace que las proyecciones tecnológicas a largo plazo resulten equivocadas y hasta ridículas en ciertos casos. En la literatura y el cine encontramos muchos ejemplos: se proyectaba un inicio del siglo XXI con autos voladores, la colonización de la luna, los viajes espaciales y las guerras nucleares; pero no se pronosticaron las revoluciones digitales que generaron internet y las tecnologías de las comunicaciones, y posteriormente la popularización de las inteligencias artificiales.

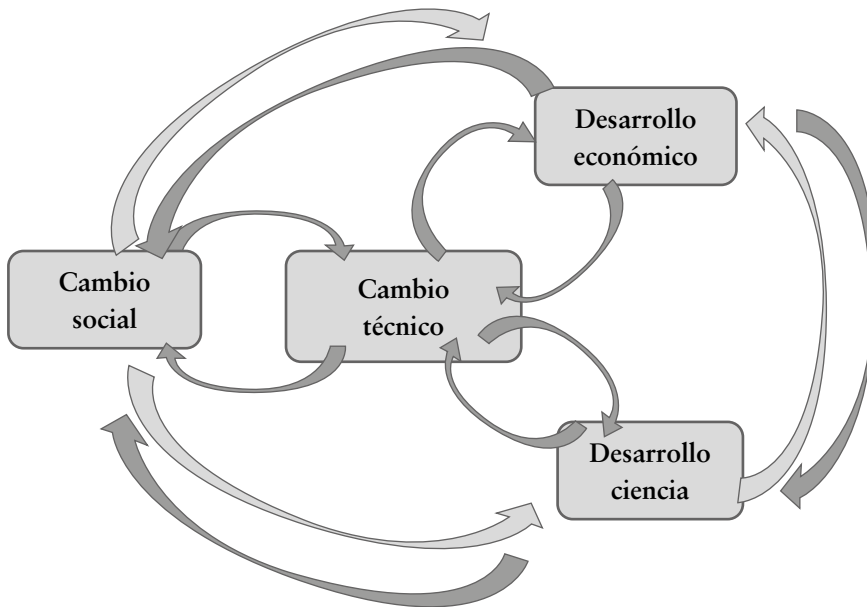


Figura 6. Interrelaciones de causas y efecto entre el cambio tecnológico con el desarrollo científico, el desarrollo económico y el cambio social

Fuente: elaboración propia.

No obstante, la explicación del cambio técnico ha dado lugar a diversas iniciativas y aproximaciones teóricas que aparecieron con las primeras teorías económicas del siglo XIX hasta los modelos epistemológicos y constructivistas de nuestros tiempos. Entre algunos de los múltiples referentes encontramos los autores referenciados en la tabla 2.

Tabla 2. Diversos enfoques explicativos del cambio técnico

	Modelos de cambio técnico	Referentes bibliográficos
1	Explicaciones deterministas	Moore (1963), Kurzweil (2005)
2	Explicaciones economicistas	Marx (1867), Schumpeter (1942), Dosi (1982), Mokyr (1990)
3	Explicaciones culturalistas e informacionales	Dawkins (2000), Mosterín (1993)
4	Explicaciones evolucionistas darwinistas	Basalla (2011), Ziman (2000) Nelson y Winter (1982)
5	Explicaciones evolucionistas lamarquianas	Elster (1983)
6	Explicaciones desde sistemas complejos	Weisbuch (1991), Kirman (1993)
7	Explicaciones constructivistas	Bijker et al. (1987)
8	Explicaciones a través de mecanismos de individuación-concretización	Simondon (2007)

Fuente: elaboración propia.

Explicaciones deterministas: las explicaciones deterministas del cambio técnico (1) suponen intrínsecamente un “modelo lineal de desarrollo” en el que existe una lógica autónoma del desarrollo tecnológico casi independiente del contexto social y económico; en ese sentido, la tecnología tiene una lógica autónoma que sigue valores generales intrínsecos como la eficiencia, la eficacia o la productividad. Como se mencionó en secciones anteriores, es la concepción de cambio técnico predominante en la divulgación tecnocientífica, de los divulgadores de la innovación y, por lo tanto, la más común en el ámbito cotidiano. No

existe una posición académica propia, pero conceptos como “la ley de Moore” (Moore, 1963) o la “singularidad” (Kurzweil, 2005) denotan este tipo de aproximaciones.

Explicaciones economicistas: por otro lado, existen explicaciones de corte económico (2) que señalan las causas del desarrollo tecnológico a factores económicos. En ellas se toman como unidad de cambio el crecimiento económico (Schumpeter, 1942), el crecimiento empresarial (Dosi, 1982; Mokyr, 1990) y, en algunos casos, el cambio social (Marx, 1867).

Explicaciones informacionales: las explicaciones informacionales (3) consideran que todos los elementos culturales (desde los conocimientos y valores hasta los artefactos y las tecnologías) pueden plantearse como unidades de información. Esta aproximación informacional del cambio técnico (Dawkins, 2000; Mosterín, 1993) ha propuesto explicar el cambio cultural (por lo tanto, también técnico) a través de unidades de información cultural (llamadas “memes”) que se transmiten de generación a generación y, hasta cierta manera, están sometidas a procesos evolutivos.

Explicaciones evolucionistas: una de las familias más populares de modelos de cambio técnico tiene inspiración en los mecanismos evolutivos, tanto en las aproximaciones darwinistas (4) (Basalla, 2011; Ziman, 2000; Nelson y Winter, 1982) como las no darwinistas lamaquianas (5) (Elster, 1983). En estas aproximaciones, las unidades de estudio evolutivo (conocimientos, artefactos, empresas, etc.) están sometidas a procesos de variación y selección a través de mecanismos adaptativos que explican el carácter histórico y contingente del cambio técnico.

Explicación desde sistemas complejos: por otra parte, existe un enfoque de estudiar el cambio técnico desde la perspectiva de los sistemas complejos (6) (Weisbuch, 1991; Kirman, 1993), el cual se centra en el estudio de las propiedades emergentes y la dinámica de sistemas compuestos por numerosas partes que interactúan entre sí. Estos estudios utilizan un enfoque matemático y computacional para modelar y analizar sistemas complejos.

Explicaciones constructivistas: en el estudio del cambio técnico también se pueden mencionar las aproximaciones constructivistas sociales (7) como el modelo SCOT (Bijker et al., 1987) en el que existe un

proceso de cambio técnico a partir de la negociación de grupos sociales relevantes a las invenciones técnicas propuestas por inventores y compañías. El artefacto o sistema sociotécnico va adoptando cambios hasta que se estabilizan los diseños por la negociación de grupos sociales relevantes que serían tanto los diseñadores, las compañías, los usuarios, los grupos de resistencia a una tecnología y los agentes gubernamentales.

Explicaciones como mecanismos de individuación: por último, también podemos referir la explicación de Simondon (2007), escrita en 1958, sobre el proceso de *concretización* (8) a la que están sometidas los objetos técnicos (en especial las máquinas). Según Simondon, los objetos iniciales (abstractos) pueden ser bastante rígidos y separados de su contexto, lo que limita su adaptabilidad y eficiencia. A medida que los objetos técnicos se utilizan y se someten a iteraciones, comienzan a adaptarse a su entorno y a integrar mejor sus funciones. Las limitaciones y las redundancias se eliminan gradualmente, y los objetos técnicos se vuelven más eficientes y versátiles. Al final, los objetos técnicos alcanzan un estado de (cuasi)equilibrio en el que sus funciones, estructuras y relaciones están altamente integradas y optimizadas (objetos concretizados).

Como podemos observar, el tipo de explicaciones del cambio técnico va a depender de lo que se entiende por tecnología y cuál es el interés del investigador sobre lo que quiere describir, por ejemplo, un enfoque economicista puede ser utilizado para una descripción general de tecnologías de largo aliento, atendiendo principalmente a las relaciones sociales y económicas, mientras que un enfoque evolucionista o constructivista puede ser más útil para estudiar un artefacto en particular y las relaciones que estableció con una práctica o el impacto que generó en un determinado caso histórico.

Para tener un estado del arte y la discusión de algunas de estas familias de teorías, recomendamos revisar el documento *Una introducción a la historia de las teorías de la evolución de la tecnología*, de Luciano Gallón (Gallón, 2007), que muestra un seguimiento sistemático y sintético de autores en mapas conceptuales desde el siglo XIX hasta el cambio de milenio.

A continuación, nos centraremos en las explicaciones de tipo evolutivo, ya que en nuestro concepto estas aproximaciones revelan bien el aspecto contingente de la tecnología, pues son modelos: históricos (que explican, no predicen); tienen un carácter antiesencialista; no son teleológicos y no utilizan jerarquías explicativas. Para más detalles del porqué de la elección de los modelos evolutivos, recomendamos el texto de Domínguez y Monterroza (2015).

Los modelos evolutivos del cambio técnico

Nuestros mundos artificiales tienen una altísima diversidad, relationalidad y complejidades que los asemejan a los ecosistemas naturales. Existen tantos y tan diversos objetos, partes de ellos, infraestructuras, sistemas, procesadores de información, imágenes, textos, etc., que en vez de considerarlos individualmente merecen ser estudiados en conjunto como un nicho ecológico artificial en permanente cambio y relación con los grupos humanos que los habitan. George Basalla (2011) señalaba que cuantificando solo las patentes (como registros formales de invenciones nuevas en la década de 1990) su cantidad ya triplicaba al número de especies de seres vivos descubiertos hasta ese momento. Durante las últimas dos décadas, podemos afirmar que la cantidad de invenciones técnicas ha seguido creciendo impulsada principalmente por la revolución de las tecnologías de la información y la comunicación. La reflexión que hace Basalla es que, si el modelo darwiniano pudo explicar el origen y la diversificación de especies vivas, también podría explicar la diversidad del mundo artificial. Por supuesto, los mecanismos de reproducción entre los seres vivos y los objetos artificiales tienen que ser distintos, pues, mientras los organismos tienen capacidades de autorreproducción, los artefactos requieren del esfuerzo técnico humano para la reproducción de sus formas y diseños.

El concepto de evolución está vinculado a un cambio de estado o a un despliegue o desenvolvimiento, y su resultado es un nuevo aspecto o forma del elemento en cuestión. Por lo tanto, evolución es netamente un cambio y ese cambio no necesariamente tiene una dirección o un propósito, por ende, es equivocado asociar este término al progreso o

aumento de una cualidad particular. La evolución, en términos darwinianos, no puede valorar si una trayectoria evolutiva es mejor que otra, ya que no existen patrones universales para compararlas, de hecho, solo se refiere a la eficacia biológica de una especie en términos de adaptación a un determinado ambiente, por lo que son valoraciones relativas y no absolutas. La evolución natural por selección planteada por Darwin se basa en tres mecanismos (Freeman y Herron, 2002):

- Debe haber una variación previa, o anterior a que suceda el proceso evolutivo.
- Debe existir un método para seleccionar esa variación, logrando pasar a la generación siguiente.
- Debe existir un mecanismo para que retenga esa variación.

Para los seres vivos, se recurre a los genes como “unidad de análisis”, como lo dice la teoría sintética de la evolución (Gould, 2006), en que existen variaciones por mutaciones o cambios en la secuencia genética que es al azar. La selección de esa variación se debe a que esa mutación fue provechosa o desventajosa para esa especie en el hábitat en que vive; si la mutación fue provechosa (que sucede en la minoría de los casos), el individuo o grupos podrán tener más descendencia y asegurar la retención de esa variación. Por lo tanto, este sencillo mecanismo podría explicar la enorme diversidad de especies vivas en el planeta, así como la historia de las especies pasadas y los ancestros comunes. El “árbol filogenético” de las especies no sería tanto un árbol sino un coral en que los individuos actuales que viven en la superficie del coral están soportados sobre los restos de los individuos que vivieron en el pasado (Gould, 2006).

Darwin y Wallace no solo dieron un modelo que explicaba la evolución de las especies vivas, sino que ofrecieron un modelo general que puede ser usado para modelar cualquier proceso desarrollo o proceso histórico que implique variaciones que puedan ser después escogidas o seleccionadas para ser reproducidas (Aranguren, 2010, p. 317).

Donald Campbell propuso un modelo evolutivo general que intentó aplicarlo a la generación de conocimiento en su propuesta de “epistemología evolucionista” que denominó “variación ciega y retención selectiva” (Campbell, 1960) aplicable a cualquier unidad de análisis que siga un proceso de desarrollo. En el caso del cambio técnico, los diferentes enfoques evolutivos han utilizado diferentes tipos de “unidades de análisis” como se muestran en la tabla 3.

Tabla 3. Compilación de unidades de análisis de evolución basada en revisión bibliográfica

Unidad de análisis evolutivo	Algunos exponentes
Empresas	Nelson y Winter (1982), Dosi (1982)
Conocimientos – Teorías	Popper (1979), Campbell (1960), Simonton (1999)
Información cultural	Dawkins (2000), Mosterín (1993)
Funciones técnicas	Lewens (2015), Preston (2009)
Usos tecnológicos	Bijker et al. (1987)
Artefactos	Basalla (2011), Ziman (2000), Broncano (2001)

Fuente: basado en Domínguez y Monterroza (2015).

Creemos que la unidad de análisis más interesante y con la cual se obtiene más información de las novedades por el registro de patentes son los artefactos. Además, en otros escenarios hemos defendido que los artefactos son los nodos objetivos de relaciones tales como la forma, la función, el material, la historia individual del artefacto, la historia filogenética del artefacto, las relaciones del artefacto con otros objetos y con sus usuarios. Por este motivo, posibilitan las prácticas y mantienen la cohesión social, lo que los convierte en las unidades fundamentales de los ecosistemas artificiales; son dispositivos objetivos que redireccionan flujos causales de materia, energía e información (Monterroza, 2018).

De todas maneras, hay que dejar claro que al no tener la capacidad de reproducirse por sí mismos, los replicadores, esto es, la entidad que transmite su estructura durante la replicación, serían los diseños

plasmados en medios representacionales (planos, prototipos, patentes, etc.) o en la memoria intelectual y corporizada de los artesanos. Las variaciones de los artefactos son propuestas por personas, grupos, compañías, etc., y son seleccionadas por un entorno de humanos y objetos (sociedad, mercado, comunidad, público objetivo, etc.). A diferencia de Campbell, las variaciones de los objetos no son necesariamente ciegas y pueden surgir del estudio intencionado de las condiciones iniciales. En efecto, Dean Simonton afirma que en el modelo de Campbell la variación no debe considerarse necesariamente como algo discreto (cien por ciento ciego o cien por ciento intencional), sino como un continuo entre los extremos entre la variación completamente ciega o la variación completamente intencionada (Simonton, 1999).

En la figura 7 mostramos un ejemplo esquemático desde que aparece la primera versión de una clase de artefacto (forma A) que puede surgir de un uso creativo de otro artefacto, de la inspiración en la naturaleza o de una hibridación de otros artefactos; la forma A inaugura una nueva clase artefactual. Esta forma A fue seleccionada dentro de un nicho de cultura material, pero con el tiempo empieza a ser sometida a variaciones por los agentes técnicos de dicho nicho. Así aparecen nuevas formas inspiradas en A como lo son la forma B, B*, C o D, que también son sometidas a procesos de selección de su entorno (el nicho de cultura material). Como se muestra en la figura 7, la coordenada horizontal señala el paso del tiempo (gen 1 a gen 7) en que se van diversificando las formas; algunas son seleccionadas por un tiempo, pero llega un momento que deja de reproducirse (forma A), otras inspiran nuevos linajes que van a mantenerse durante varias generaciones (formas B, C o D). Observamos que sigue el proceso de variación con formas como E, F, G, I o H aumentando la diversidad si el entorno es favorable a las variaciones. También es importante anotar que en las invenciones técnicas es posible que existan hibridaciones (forma H) que no suceden con los seres vivos, pues las hibridaciones de los organismos no tienen descendencia. En todo caso, toda la diversidad de formas comparte algo: un ancestro común; por lo que es una característica importante en los modelos evolutivos, esto es, las invenciones enriquecen un horizonte de posibilidades para otras invenciones.

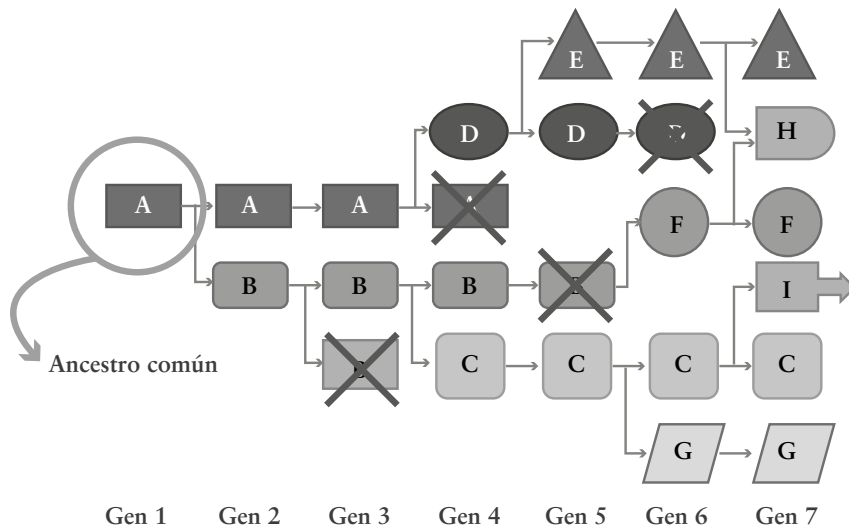


Figura 7. Esquema representativo de un proceso de variación y selección en la evolución de una clase de artefacto (A) en un nicho de cultura material

Fuente: elaboración propia.

Las explicaciones evolutivas, como la acá presentada, tienen varias cualidades por mencionar: la primera es que todas las invenciones siguen cierta continuidad, es decir, que ninguna invención es completamente novedosa en sentido estricto, pues recurre a formas, materiales, funciones, conocimientos y habilidades que existen previamente en una cultura material. El modelo no se contradice con las invenciones disruptivas, ya que pueden ser explicadas de la misma manera que las invenciones incrementales. También ofrece un marco de explicación general para modelos de tecnologías particulares como el ciclo de vida de las tecnologías repartidas en etapas de introducción, crecimiento, madurez y declive inspirados en la idea de la curva en “S” de Richard Foster (1986). Los modelos evolutivos son los suficientemente flexibles, pero a la vez lo suficientemente demarcados, como para describir una trayectoria histórica de una tecnología en particular.

Ya mencionamos que las ventajas de los modelos evolutivos del cambio técnico están en que, al ser modelos históricos, no teleológicos

y adaptativos, explican bien las trayectorias contingentes tecnológicas; sin embargo, esto hace que no sean capaces de predecir las tecnologías a mediano y largo plazo. En principio, esto no parece un gran problema si lo que se quiere explicar es la diversidad y el carácter contingente de la tecnología y su relación con los entornos de cultura material, aunque esos “ecosistemas” tengan procesos de diseño intencionales. También, al ser sistemas dinámicos son adaptativos, pues cambian en función del entorno y ese entorno es transformado permanentemente por sus mismos resultados. Esto nos da una idea: no controlamos los sistemas adaptativos, pero sí podríamos controlar una o varias propiedades de ese sistema, esto es, preservar lo que nos parezca valioso.

A pesar de estas ventajas, aún se presentan limitaciones que deben ser mencionadas para el caso del cambio técnico:

1. En el cambio técnico, la variación no es azarosa sino intencional, esto es, hay procesos dirigidos de diseño que no suceden con los seres vivos y sus ecosistemas. Existen localmente procesos de cambio técnico que son intencionales y no azarosos.
2. Debido a que hay procesos intencionales en el mundo técnico, las explicaciones evolutivas no aclaran de dónde surgen las innovaciones (mutaciones) técnicas. En el caso de los seres vivos, las variaciones se explican por las modificaciones en la recombinación de los replicantes (genes), pero en la técnica se necesita una teoría convincente de la creatividad (Simonton, 1999).
3. En los procesos técnicos existe hibridación, entendida como la combinación y recombinación de formas, diseños y funciones provenientes de distintos linajes tecnológicos. A diferencia de los seres vivos, donde la hibridación entre especies suele ser limitada y en muchos casos no produce descendencia fértil, en el ámbito técnico constituye un mecanismo fundamental de innovación. La historia de la tecnología ofrece numerosos ejemplos en los que la integración de principios, mecanismos y arquitecturas diversas ha dado lugar a nuevos artefactos y trayectorias

técnicas. Los motores de combustión interna, los turborreactores, los circuitos integrados, las computadoras personales y los teléfonos inteligentes ilustran cómo estos procesos de combinación generan nuevos linajes tecnológicos.

Estos problemas dejan abierta una serie de preguntas: ¿cómo explicar los procesos intencionales locales?, ¿cómo explicar las innovaciones o mutaciones?, ¿cómo explicar las hibridaciones y los usos creativos?, ¿cómo dar detalles de los métodos de creación técnica? Son preguntas que abren una veta de oportunidades para la reflexión filosófica.

¿Mejores tecnologías? El concepto “tecnologías entrañables”

Una de las paradojas del desarrollo tecnológico son las tendencias de diseño centradas en las tecnologías “amigables”. Esto es, una mirada del diseño tecnológico que aboga por hacer diseños amigables con los usuarios. Donald Norman, un psicólogo cognitivo experto en diseño centrado en el usuario, es uno de los autores que propone una serie de principios de diseño para garantizar una experiencia de usuario efectiva y agradable en la interacción con productos y sistemas. Uno de los principios fundamentales es la visibilidad, que se refiere a la importancia de que los elementos y opciones del diseño sean fácilmente perceptibles para el usuario. Esto facilita la navegación y el uso del producto o sistema, mejorando la experiencia general.

Otro aspecto es la retroalimentación, la cual establece que los usuarios deben recibir información clara y oportuna sobre los resultados de sus acciones. Esto les permite confirmar si han realizado correctamente una acción o si necesitan corregir errores. El mapeo es otro principio de diseño que se centra en la relación entre los controles y sus efectos. Un buen mapeo garantiza que esta relación sea fácil de comprender y utilizar, permitiendo a los usuarios predecir los resultados de sus acciones y operar el sistema de manera intuitiva. La consistencia en el diseño es esencial para facilitar el aprendizaje y el uso del sistema. Esto

implica mantener una presentación, terminología y funcionamiento coherentes a lo largo de la experiencia del usuario.

Norman también sugiere aprovechar las *affordances*, o invitaciones al uso, como características que sugieren cómo se puede interactuar con un objeto de manera intuitiva, es decir, un buen diseño debe comunicar claramente cómo se deben utilizar los elementos o controles, y qué acciones son posibles. La flexibilidad es otro aspecto importante del diseño centrado en el usuario, pues los productos deben adaptarse a las necesidades y habilidades de diferentes usuarios, permitiendo personalizar la experiencia de uso según las preferencias individuales. Por último, la tolerancia a errores es un principio que busca minimizar la posibilidad de errores y facilitar su corrección. Esto incluye prevenir errores antes de que ocurran y ofrecer opciones para deshacer acciones, asegurando una experiencia de usuario más satisfactoria (Norman, 2005, 2010). A estas cualidades de un diseño, Norman les llama “transparencia” y sucede cuando no se tienen barreras intuitivas para la interacción con la multitud de artefactos con los que interactuamos a diario. En principio, este tipo de iniciativas puede facilitar el desarrollo de prácticas cotidianas en que las dificultades cognitivas y prácticas que tienen los objetos cotidianos pueden reducirse.

No obstante, aunque estas ideas de convertir la interacción con los artefactos cotidianos en un envoltorio amigable y significativo pueden ser deseables, podría tener unas consecuencias negativas a largo plazo. Nos convierten en analfabetas tecnológicos porque desconocemos la lógica y la forma como están diseñados los mecanismos internos de una herramienta, máquina, *software* o sistema. Nos aumenta la comodidad y el bienestar frente a la interacción con los objetos cotidianos, pero nos puede llevar a una alienación total del desarrollo tecnológico al volvernos “niños” que desconocemos las lógicas técnicas que terminan condicionando nuestras propias prácticas, el orden social y nuestras identidades. Es el lado oscuro del diseño amigable, los diseños “transparentes” que promueve Norman y otros, en realidad son “opacos” porque terminamos desconociendo los mecanismos por debajo de la capa amigable de las tecnologías, lo que evita una mirada crítica y un control social de estas. No se puede pedir prudencia o tener una actitud crítica de lo que no se entiende. En cuanto a las

tecnologías digitales, este desconocimiento es aún más pronunciado, como el caso de los algoritmos de las redes sociales que seleccionan de una manera opaca nuestro contenido. Si bien las redes sociales están conformadas por una interfaz atractiva, amigable y llena de estímulos para la interacción, termina envolviendo a sus usuarios en contenidos que desincentivan sus facultades cognitivas de atención o incluso a la sumisión en agujeros ideológicos en que todas sus creencias son reforzadas continuamente.

Frente a este desconocimiento generalizado de los mecanismos internos de las tecnologías y sus consecuencias alienadoras, filósofos como Miguel Ángel Quintanilla han propuesto el concepto de tecnología “entrañable”. Es una propuesta normativa, también de origen hispano (Quintanilla, 2002, 2012; Quintanilla et al., 2017), que busca identificar criterios y cualidades a las tecnologías de amplio impacto social para superar la alienación de sus usuarios, para que así puedan comprender mejor y controlar sus impactos y posibilidades de acción. Miguel Ángel Quintanilla ha reutilizado la idea hegeliana de “alienación”, que utiliza y desarrolla Marx con respecto a la sociedad y producción capitalista, para describir lo que está sucediendo con las tecnologías contemporáneas. Marx, en los *Manuscritos de París de 1844*, le dio un sentido más práctico al término “alienación” como el proceso mediante el cual los individuos se sienten separados, desconectados y deshumanizados en el sistema y sociedad capitalista. Marx identificó cuatro dimensiones principales de la alienación que genera el modo de producción capitalista: 1) la alienación del trabajo; 2) alienación del producto del trabajo; 3) la alienación de otros seres humanos y; 4) la alienación de la propia naturaleza humana (Marx).

Este concepto de alienación ha sido mencionado por una buena parte de la tradición filosófica y está relacionado con lo “extraño” y distante al sujeto, que no le pertenece. La alienación se refiere a la sensación de extrañamiento que experimentan los usuarios debido a la desvinculación técnica, cultural y/o representacional con las tecnologías. Esto puede ocurrir cuando los usuarios no entienden cómo funcionan las tecnologías o cómo se relacionan con su vida cotidiana. Como vimos, la tecnología aliena al trabajador al transformarlo en un subordinado a la máquina y al sistema de producción (Marx),

aliena al ciudadano al convertirlo en consumidor y al situarlo en un mundo virtual mediante los influyentes medios de comunicación masiva (Adorno y Horkheimer). Además, aliena al pensamiento crítico al subyugarlo a una episteme de producción (Heidegger) y a la sociedad en general al trasladar las categorías del pensamiento técnico a la esfera política y la organización social (Mumford).

En ese sentido, el término “entrañable” sería lo contrario a lo “extraño”, pues es un adjetivo que se utiliza para superar el extrañamiento en ciertos modos de expresión de las tecnologías y no solo para describir algo o alguien que provoca sentimientos de afecto y conexión. El adjetivo entrañable no se refiere solo al aspecto emocional de los diseños, sino que hay cualidades morales y políticas que deben ser tenidas en cuenta para considerar una tecnología como abierta, cercana y comprensible. Según los promotores del concepto, es un enfoque que promueve el diseño y consumo de tecnologías abiertas en lugar de aquellas propietarias y cerradas, busca fortalecer el mercado de segunda mano en vez de aceptar la obsolescencia programada, abandona la mentalidad de usar y desechar, y adopta el principio de mantenimiento y reparación constante de nuestros artefactos. Además, enfatiza en mantener el control sobre nuestras máquinas en lugar de depender de ellas, y se prefiere soluciones tecnológicas compatibles con el desarrollo sostenible, con los valores de la solidaridad y la cooperación en vez de la competencia y el éxito individual (Quintanilla et al., 2017). A continuación, en la tabla 4 presentamos el decálogo de características y rasgos que debería tener, en cierto grado, una tecnología entrañable que propone Miguel Ángel Quintanilla.

Tabla 4. Decálogo de características
de tecnologías entrañables

Número	Característica de la tecnología	Descripción
1.	Abiertas	Accesibles y apropiables
2.	Polivalentes	Susceptibles de usos alternativos
3.	Dóciles	Controlables por el usuario
4.	Limitadas	Consecuencias previsibles, aplicación del principio de precaución si no son previsibles

Número	Característica de la tecnología	Descripción
5.	Eventualmente reversibles	Posibilidad de revertir cambios en caso de falla; evitar proyectos tecnológicos que cambien el mundo de forma irreversible y con riesgo de destrucción
6.	Recuperables	Susceptibles de mantenimiento activo y recuperación de residuos; tecnologías accesibles al ciudadano
7.	Comprensibles	Diseño manifiesto, transparente y no opaco; fácil de usar sin instrucciones
8.	Participativas	Facilitan la cooperación humana
9.	Sostenibles	Permiten el ahorro y reciclado de energías y recursos
10.	Socialmente responsables	La implantación de una nueva tecnología no contribuye a empeorar la situación de los colectivos más desfavorecidos

Fuente: basado en Quintanilla M. Á. (2012).

La tabla 4 describe diez características esenciales de la tecnología que promueven un enfoque sostenible y responsable en su diseño y aplicación. Estas características incluyen la accesibilidad, la versatilidad en sus usos, el control por parte del usuario y la limitación de sus consecuencias a través de la precaución. Además, se enfatiza en la importancia de la reversibilidad, la recuperación, la comprensión y la participación en el diseño tecnológico. Por último, se aboga por la sostenibilidad y la responsabilidad social en la implantación de nuevas tecnologías, garantizando que no empeoren la situación de los colectivos más desfavorecidos.

Por supuesto, los parámetros de la tabla 4 son cualitativos y bastante generales, y no dictámenes rígidos normativos. No chocan ni con la lógica de la economía de mercado ni con las economías centralizadas por el Estado. Son guías para la acción para tener en cuenta en el diseño, la apropiación y la evaluación de tecnologías. Este es un ejemplo de que la filosofía de la tecnología puede formular iniciativas prácticas para proponer cambios en la manera en que transformamos los entornos artificiales y, por lo tanto, nuestras formas de vida.

Como lo mencionamos en la sección de cómo cambian las tecnologías, los ecosistemas artificiales son análogos en complejidad de relaciones, diversidad y evolución a los ecosistemas naturales, por

consiguiente, no pueden controlarse de manera general al ser sistemas adaptativos, pero sí podemos escoger qué tipo de valor y propiedad de esos ecosistemas son más valiosos. Este conjunto de valores que nos propone Quintanilla (tabla 4) es una muestra práctica de ellos.

Preguntas

1. ¿De las concepciones comunes de tecnología (instrumental, intelectualista o sistémica) cuál encuentras más convincente y por qué? ¿Puedes identificar alguna limitación o aspecto problemático en cada una de ellas?
2. Considerando el concepto de “posibilismo” en tecnología, ¿cómo podrías argumentar en contra del determinismo tecnológico? ¿Qué implicaciones tendría adoptar una postura posibilista para nuestra relación con las tecnologías y su desarrollo? ¿Qué problemas tiene abordar una postura posibilista?
3. Al analizar el contexto de diseño y el contexto de uso, ¿cómo influyen estos aspectos en la forma en que entendemos y valoramos las tecnologías? ¿Puedes pensar en ejemplos de tecnologías cuya comprensión y valoración cambian significativamente según el contexto de diseño o el de uso?
4. ¿Cuáles son las cualidades de una tecnología entrañable? ¿Cuáles son algunas de las ventajas y desventajas de enfocarse en aspectos entrañables al evaluar y seleccionar tecnologías? ¿Puedes pensar en ejemplos de tecnologías entrañables y cómo afectan nuestra vida cotidiana?
5. A partir del caso de estudio de los modelos IA de procesamiento de lenguaje natural como ChatGPT, ¿cómo podemos abordar las cuestiones ontológicas y éticas relacionadas con estas tecnologías? ¿De qué manera el impacto de ChatGPT

en las esferas de la realidad humana plantea preguntas filosóficas sobre la naturaleza de la tecnología y nuestra relación con ella?

Para saber más

- Para una aproximación de una antropología de la técnica se recomienda el libro clásico de José Ortega y Gasset, *Meditaciones de la técnica*, que nos muestra una serie de reflexiones sobre el papel de la técnica en la constitución de lo humano. Es un texto accesible y escrito con un lenguaje sencillo y muy elegante (Ortega y Gasset, 1982).
- Para explorar sobre las concepciones comunes y una descripción general sobre qué es tecnología desde diferentes aproximaciones y temas correlativos, se recomienda el segundo capítulo del libro *Ciencia, tecnología y sociedad: una aproximación conceptual* escrito en conjunto por diversos autores iberoamericanos y editado por la Organización de Estados Iberoamericanos (OEI) (García Palacios et al., 2001).
- Para una filosofía de la ciencia de corte posibilista, se recomienda el libro *Entre ingenieros y ciudadanos. Filosofía de la tecnología para días de democracia*, de Fernando Broncano. Es una pequeña obra maestra, con un enfoque bastante novedoso, para comprender aspectos políticos y morales de la tecnología, sin recurrir al fatalismo típico de algunos críticos culturales, y dando criterios desde la filosofía para una actitud crítica y propositiva sobre la relación entre tecnología y sociedad (Broncano, 2006).
- Para la ampliación del concepto “tecnología entrañable”, se recomienda el libro colaborativo de Quintanilla et al. (2017) que propone una relación política y moral sobre de las tecnologías contemporáneas para superar la sensación de extrañamiento

que se ha vuelto común en los últimos tiempos (Quintanilla et al., 2017).

- Para comprender el funcionamiento de las inteligencias artificiales, en especial, los modelos enormes de procesamiento de lenguaje natural, recomendamos el video de Nate Gentile *¿Cómo funciona ChatGPT? La revolución de la inteligencia artificial* (Gentile, 2023), quien describe de una manera absolutamente clara y sencilla cómo funcionan por dentro estos modelos explicando los conceptos y tecnicismos necesarios para una comprensión global. <https://youtu.be/FdZ8LKijBhQ>

Capítulo 2

Clasificar el conocimiento tecnológico: la tecnología como conocimiento

DAIAN TATIANA FLÓREZ QUINTERO

JORGE FRANCISCO MALDONADO SERRANO

DAIRON ALFONSO RODRÍGUEZ RAMÍREZ

Introducción

Presentar un capítulo sobre epistemología de la tecnología implica considerar la necesidad del estudio del conocimiento tecnológico en la variedad de sus aspectos. Después de hablar de ontología pasamos a la revisión de los estudios epistemológicos de la tecnología porque corresponden a una etapa que desarrolla y amplía los estudios ontológicos de la tecnología. En efecto, dan paso a preguntar por las implicaciones político-sociales, éticas y educativas de la tecnología.

La pregunta epistemológica central siempre se refiere al conocimiento; en este caso, puede formularse así: ¿qué tipo de conocimiento es el conocimiento tecnológico? Este interrogante es el punto de partida

para abordar la epistemología de la tecnología. Para comprender el planteamiento y la posible respuesta a esta cuestión, es necesario revisar de manera introductoria el problema de la técnica y establecer la relación y la diferencia entre la ciencia y la tecnología.

La pregunta epistemológica implica saber algo sobre el significado del conocimiento tecnológico y, por derecho, de lo que significa *tecnología*, por lo menos en parte. En este sentido, la problemática de la especificidad, naturaleza o significado la comparte la epistemología con la ontología, respecto de la tecnología. En la primera sección del capítulo se revisará este asunto en clave histórica, porque dicho enfoque permite entender lo que se ha concebido por *técnica* y *tecnología*, esto es, ubica de manera crítica la idea que actualmente se puede tener de tecnología.

En la misma dirección, conviene señalar la pertinencia de otra pregunta clave para la epistemología, dado que esta se ocupa del estudio del conocimiento científico: ¿qué es la tecnología? Cualquier estudio que intente definir o delimitar —así sea de forma relativa— un objeto de análisis debe primero cuestionar sus propios supuestos. En este caso, debemos empezar por preguntarnos ¿qué tan científica es la tecnología? O, inversamente, ¿qué tan tecnológica es la ciencia? ¿Basta usar el concepto de *tecnociencia*?

A este respecto, se puede sostener que la tecnología es ciencia aplicada o, por el contrario, que constituye algo completamente distinto. En ambas perspectivas, se presupone que la especificidad de la tecnología se define en relación con la ciencia, motivo por el cual la segunda sección ofrece un recorrido sobre este problema.

Comprender el conocimiento tecnológico implica, por tanto, establecer una concepción clara de la tecnología y de su vínculo, de una u otra manera, con la ciencia. Elaborar una epistemología de la tecnología da por sentada su especificidad y requiere una revisión de la forma concreta en que puede estudiarse la tecnología misma. La tercera sección se divide en un estudio sobre la especificidad del conocimiento tecnológico, un análisis sobre la operación de diseño propia de la tecnología y, finalmente, un examen sobre las operaciones corporales que toda la tecnología implica. Estos tres aspectos se entrelazan en un equilibrio pragmático que alterna entre creatividad e invención, saber-que y saber-cómo.

El concepto *techné*

La sola idea de tecnología parece implicar dos dimensiones diferentes: *techné* y *logos*. Esto querría decir que la tecnología es el *logos* de una técnica; dicho en otras palabras, el discurso, el estudio o ciencia de una técnica. Habría una especie de contradicción. En parte porque el término griego *techné* suele traducirse por expresiones como “habilidad”, “oficio”, “experticia”, “profesión” o “arte”, y vincularlo al plano gnoseológico, al *logos*, puede resultar extraño. ¿Por qué habría un *logos* de la técnica o una ciencia de la tecnología?

Sin embargo, para los griegos en tiempos de Platón, y para él mismo, hablar de *techné* significaba no solo la *praxis* efectiva de un oficio o profesión, sino también el conocimiento en el cual tal *praxis* se encuentra fundamentada. En los diálogos, Platón se refiere a esta clase de conocimiento con el nombre de *logos*, término que suele traducirse, en este contexto, como “explicación”. En ese sentido, la expresión *techné* comportaba una clara dimensión epistemológica.

En el diálogo *Timeo* (Platón, 2012), el papel del demiurgo es el de ser legislador y el de ser artesano¹. El demiurgo es artesano porque fabrica las imágenes a partir de las formas, es decir, en la medida en que contempla las formas puede hacer las imágenes. Dicho en otras palabras, porque conoce (contempla) las formas, puede fabricar copias de ellas. De ahí que, para Platón, cualquier actividad técnica, cualquier saber hacer técnico estuviera ligado a un saber teórico o contemplativo.

Así, *techné* era una habilidad que implicaba conocimiento, por lo menos de cierto tipo; algo que Aristóteles explicita aún más. De los

1 Conviene recordar que, aunque hoy nos referimos a Platón y Aristóteles con cierta confianza editorial, sus obras, como las conocemos hoy, han sido recuperadas y, en cierto sentido, restauradas. El diálogo *Timeo* fue el único que circuló por Europa después del siglo IV a. C. hasta que en el siglo XII, la renaciente civilización europea se encontró con la avanzada civilización musulmana en Toledo y volvió a tener acceso a varios manuscritos de los antiguos filósofos griegos (Rubenstein, 2003, cap. I). En el caso de Aristóteles, solo su *Organon* se mantuvo al lado del *Timeo*; su *Metafísica*, *Ética a Nicomaco*, *De anima* y otras obras entraron propiamente en el radar de la civilización europea en el XII.

manuscritos aristotélicos, para efectos de esta problemática de la técnica, interesan la *Metafísica* y la *Ética a Nicómaco* porque en estos presenta de forma sintética su manera de entender los niveles de conocimiento (Arroyave Montoya, 2013, p. 45).

El estagirita, como se le llama, explicitó en el primer capítulo del libro I de la *Metafísica* que los niveles de conocimiento propios del ser humano son la experiencia (*empeiría*), la técnica (*techné*) y la ciencia (*episteme*). En términos generales, por una parte, tener experiencia o saber una técnica se ubica en un extremo del conocimiento, y quiere decir que la persona sabe qué hacer para darle solución a algo; la experiencia es un conocimiento que tiene una utilidad. En el otro extremo, la ciencia se caracteriza por ser un conocimiento que no tiene utilidad. En medio, encontramos la técnica, que para Aristóteles tiene en común con la experiencia que es un conocimiento que sirve para dar solución a situaciones, esto es, un conocimiento que también tiene utilidad. No perdamos de vista, entonces, que la ciencia, a juicio de Aristóteles, es un conocimiento que tiene valor por sí mismo y no en razón de su utilidad.

Por otra parte, ciencia y técnica tienen en común, por oposición a la experiencia, el hecho de que implican saber teoría, conocer las causas o, como también las llama, *principios*; lo que se señalaba con la idea platónica de demiurgo. En efecto, lo específico de quien posee un arte o conocimiento técnico, frente al que solo posee experiencia radica en que este no conoce las causas que explican o dan cuenta de lo que hace, mientras que aquel sabe por qué hace algo y puede dar cuenta y razón de su hacer en una situación. En otras palabras, el que tiene experiencia sabe qué debe hacer o sabe efectuar la acción X para darle solución o respuesta a un caso específico, pero no puede dar cuenta de la razón por la cual se debe hacer o efectuar dicha acción X. La experiencia es un conocimiento limitado al caso, mientras que la técnica no está limitada al caso, sino que, por conocer la causa o el principio, sabe por qué aplica la acción X.

La diferencia entre ciencia y técnica Aristóteles la detalla mejor en el libro VI de la *Ética a Nicómaco*. Como se mencionó, en general la diferencia estriba en que toda técnica siempre tiene alguna utilidad, sea para satisfacer necesidades (agricultura, medicina) o para dar placer o

entretenen (música, teatro). Es decir, es un conocimiento que establece reglas de acción para obtener un fin externo a ese conocimiento: la salud se obtiene por la aplicación del arte médico. Esto quiere decir que el arte médico, en sí mismo, no es la salud, sino que es el medio gracias al cual se busca la salud.

La ciencia, en cambio, no satisface nada externo a sí misma, sino que solo tiene su necesidad en ella misma. Esto quiere decir que ciencia es el conocimiento de lo que existe por necesidad y no puede ser de otra manera, y realmente no tiene aplicación y no puede tenerla. En otras palabras, Aristóteles entendía que la ciencia era el conocimiento de lo que es eterno e inmutable, de lo que no está sujeto al cambio o al azar.

Se debe tener cuidado y no asumir que lo que se acepta como ciencia en la actualidad equivale a lo que Aristóteles aceptaría como ciencia. Por ejemplo, la astronomía es una ciencia en la actualidad y lo era para Aristóteles, porque el movimiento de los astros era constante. Sin embargo, la química, que se acepta sin lugar a dudas como ciencia en la actualidad, no puede serlo para el estagirita, precisamente porque tiene que ver con asuntos contingentes o azarosos.

Ahora bien, de esta distinción aristotélica interesan dos aspectos: la técnica debe tener un producto, trátase de un objeto físico existente o de la transformación de una cosa física existente; en este sentido, la técnica está relacionada con el azar, porque, aunque haya causas que se puedan tener en cuenta, el cambio a nivel físico no es necesario, es decir, tiene un grado azaroso que permea cualquier realidad física. Lo importante es que ambos aspectos aseguran la dimensión pragmática del conocimiento técnico, o el hecho de que se trate de un conocimiento útil. Y ambos asuntos nos explican por qué la ciencia aristotélica no tendría aplicación.

La razón de la inutilidad de la ciencia para Aristóteles y la utilidad del saber técnico estriba en que lo que las técnicas producen o fabrican no es algo que tenga que existir necesariamente o que se tenga que fabricar de modo obligatorio; y esto es así porque lo que resuelven o mejoran tampoco tiene que existir necesariamente: caer en enfermedad no es algo que tenga que suceder necesariamente, como tampoco lo es curarse; puede que alguien tenga un accidente y necesite del arte de la medicina, o puede que no. Hay cosas que se dan por azar y

no por necesidad, razón por la cual, para Aristóteles, todo lo que no exista o no se dé por necesidad no es objeto de ciencia, mientras que lo que existe por azar y no por necesidad es objeto de conocimiento técnico. Conviene leerlo directamente del filósofo:

Azar y técnica se relacionan con las mismas cosas tal como afirma Agatón: la Técnica ama al azar y el azar a la Técnica. Por tanto, tal como se ha dicho, la Técnica es una disposición acompañada de razón verdadera relativa a la fabricación, y la carencia de Técnica, por el contrario, es una disposición acompañada de razón errónea relativa a la fabricación —en lo tocante a aquello que puede ser de otra manera—. (*Ética a Nicómaco*, 1140a)

Así, los objetos producto de las técnicas son contingentes, cambiantes, azarosos o no obligatorios, mientras que los objetos de la ciencia no lo son, ni producidos ni sujetos al azar. Al contrario, la ciencia versa sobre objetos que no se producen, sino que están ahí desde siempre —Aristóteles dirá que por naturaleza—, y estos objetos son, de suyo, inmutables en alto grado, no sujetos al azar.

Con todo, se debe tener claro que la ciencia con la que contamos en el siglo XXI desarrollada como ciencia moderna en dos fases (mecanicista o clásica y relativista y cuántica) (Kuhn, 2004; Prigogine y Stengers, 1990)) tiene otro sistema paradigmático tan contrastante con la concepción griega de ciencia y de técnica que se debe volver a estudiar la diferencia entre tecnología y ciencia. En efecto, desde las reflexiones filosóficas de los siglos XX y XXI, encontramos maneras más adecuadas para entender la relación entre ciencia y tecnología. Es lo que se examinará en la siguiente sección.

Ciencia y tecnología

Se podría pensar, *prima facie*, que la defensa de la fórmula que identifica la tecnología con la ciencia aplicada, y que ha sido ampliamente rechazada en la tradición filosófica sobre la tecnología, es un verdadero despropósito de cara a la evidencia histórica disponible. Sin embargo,

precisamente al prestar una escrupulosa atención de la evidencia, causa estupor que filósofos e historiadores de la tecnología hayan renunciado tan fácilmente a la amplia evidencia que sí valida dicha fórmula. La extrañeza aumenta cuando advertimos que algunos de ellos son físicos.

La tecnología como independiente de la ciencia

Los casos que apoyan la fórmula mencionada son espectaculares, ya que muestran cómo la ciencia ha conducido a aplicaciones prácticas que han cambiado a las sociedades para siempre. En defensa de la fórmula, el primer caso al que se recurrirá es *al descubrimiento de la inducción electromagnética por Faraday*. Este caso es sorprendente porque incluso detractores de la fórmula que identifica a la tecnología con la ciencia aplicada reconocen que el desarrollo de la industria eléctrica fue posible gracias a la ciencia:

Como ya hemos visto, la mayor parte de los adelantos tecnológicos fueron consecuencia de descubrimientos empíricos llevados a cabo por hombres eminentemente prácticos: [...] ya dijimos que hasta hace relativamente poco tiempo la tecnología ha aportado más a la ciencia que la ciencia a la tecnología. La industria eléctrica es una excepción en el sentido de que su nacimiento y desarrollo fueron consecuencia directa de investigaciones científicas. (Derry y Williams, 1960, p. 608)

Faraday se inspiró en el descubrimiento de Oersted de que una corriente eléctrica producía magnetismo, por lo que se preguntó si lo contrario era posible. Para probarlo construyó un electroimán enrollando una bobina (un cable eléctrico) y haciendo pasar una corriente. Después colocó otra bobina en las cercanías y la conectó a un amperímetro para medir el paso de la corriente eléctrica. Faraday observó que cuando pasaba una corriente eléctrica por la primera bobina no sucedía nada en la segunda, no había corriente eléctrica. Sin embargo, se dio cuenta de que cuando apagaba la corriente en la primera bobina durante unos instantes se generaba corriente eléctrica en la segunda bobina [...] con ello comprendió

que era el magnetismo cambiante en el tiempo el que generaba la corriente eléctrica. (Farías y Cuevas, 2018, p. 45)

Gracias a este descubrimiento, se diseñó el primer generador mecánico, el cual fue expuesto en París, menos de un año después de que Faraday leyese en 1831 su ya clásica comunicación ante la Royal Society, por un fabricante de instrumentos, Hyppolyte Pixii, cuyo generador movido a mano tenía las bobinas fijas, mientras que el imán, en forma de herradura, giraba en torno a ellas :

Los primeros generados producían corriente alterna: esto es, la dirección de flujo eléctrico cambiaba constantemente, con una frecuencia que dependía de la velocidad a que se hacía girar la máquina [...] hacia finales del siglo se verificó que para su uso a gran escala, la corriente alterna tenía grandes ventajas sobre la continua². (Derry y Williams, 1960, p. 612)

El descubrimiento de la inducción electromagnética fue esencial para el comienzo de la producción de corriente eléctrica alterna y el transporte de electricidad hacia finales del siglo XIX y, por tanto, para la electrificación de la economía y de la sociedad. Este impresionante descubrimiento motivó la siguiente leyenda —cuya verdad, como muchas otras está en cuestión—:

Una vez el primer ministro británico visitó el laboratorio de Faraday y le preguntó para qué serviría la investigación que estaba adelantando, a lo que Faraday respondió: No lo sé, pero apuesto a que algún día usted le pondrá un impuesto. (Farías y Cuevas, 2018, p. 46)

Como puede advertirse, el descubrimiento de Faraday, el cual tuvo lugar en la ciencia, fue la base de la generación de electricidad en nuestras

2 Este conocimiento también sirvió de base para el diseño del motor eléctrico. El primer motor para corriente alterna fue inventado en 1888 por Nikola Tesla.

centrales eléctricas. Este caso constituye, sin lugar a dudas, una prueba de que los descubrimientos en la ciencia no solo tienen aplicaciones espectaculares en el dominio de la tecnología, sino que, además, tienen efectos decisivos en el desarrollo de las sociedades modernas.

Otro de los casos que sirven de apoyo a la ecuación que identifica la tecnología con la ciencia aplicada es la invención de la radio gracias a los estudios de Maxwell (1865): la predicción de Maxwell de ondas electromagnéticas fue comprobada por primera vez por el físico alemán Heinrich Hertz (1888). Años después, en 1895, el físico italiano Guglielmo Marconi usó las ondas electromagnéticas para establecer la primera telegrafía sin cables como medio de comunicación a larga distancia.

Ahora ocurre con frecuencia que las aplicaciones tecnológicas del descubrimiento no sean ni siquiera percibidas por el científico. Como bien lo señala Raynaud:

El físico alemán Heinrich Hertz dio una ronda de conferencias para presentar su descubrimiento de las ondas. En una de ellas, en 1892, algunos estudiantes le preguntaron sobre el uso práctico de las ondas electromagnéticas. Esta fue su respuesta: “No tienen ninguna utilidad. El experimento demuestra solamente que Maxwell tenía razón en su teoría. Tenemos esas ondas electromagnéticas que no se pueden ver a simple vista. ¡Pero están ahí! —¿Qué se puede esperar de esto?—. Nada, creo yo”. Hertz ignoraba que apenas tres años después, en 1895, Guglielmo Marconi, inspirándose en sus investigaciones, llevaría a cabo la primera conexión por radio abriendo así camino a fructíferas aplicaciones. (Raynaud, 2018, p. 76)

Como bien se sabe, el conocimiento sobre las ondas electromagnéticas es la base para el diseño de innumerables dispositivos y artefactos que usamos a diario, tales como la televisión, la radio, los teléfonos móviles, los microondas, los radares, el *bluetooth*.

El último caso que se ha de mencionar es sorprendente, ya que prueba cómo la ciencia puede transformar el mobiliario de los objetos que nos rodean (por ejemplo, computadores, teléfonos móviles,

satélites). Se trata del desarrollo de la mecánica cuántica, gracias a la cual fue posible la invención de tecnologías increíblemente versátiles, como el láser y los microchips.

Para mostrar cómo la teoría cuántica conduce a la invención del láser, Farías y Cuevas cuentan que:

En un trabajo publicado en 1917, Einstein se ocupó de la teoría cuántica de la radiación. La mecánica cuántica aún no estaba lo suficientemente desarrollada, pero Einstein estaba convencido de la validez de la ley de cuantización de la energía³ y decidió aplicarla a la interacción de la luz con átomos y moléculas. Einstein formuló una ecuación (otra de las que lleva su nombre) con la que predice el fenómeno de emisión estimulada, que es el mecanismo empleado en el láser (el primer láser no se inventaría hasta 1960). (Farías y Cuevas, 2018, p. 278)⁴

Otra de las aplicaciones esplendorosas de la mecánica cuántica es el diseño del transistor. Este diminuto dispositivo, inventado en los laboratorios Bell por los físicos John Bardeen, Walter Brattain y William Shockley, es el resultado del surgimiento en los años treinta —del siglo pasado— de la física del estado sólido.

Lo que resulta inquietante es que la invención del transistor no se vea como un caso a favor de la fórmula que aquí se defiende, sino que, por el contrario, se tome como evidencia histórica para refutarla. Así lo vio Skolimowski:

3 Fue Einstein quien propuso la cuantización de la energía, por ejemplo, sostuvo que los fotones estaban empaquetados en unos *quantum*. En 1900, Planck descubrió la constante fundamental que lleva su nombre, que es utilizada para calcular la energía de un fotón. El físico descubrió que la radiación no es emitida ni absorbida en forma continua, sino en pequeñas cantidades a las que denominó “cuantos”.

4 La teoría de la relatividad es esencial para el diseño de los sincrotrones: aceleradores de partículas que generan fuentes de luz muy intensa de energía variable, y que son herramientas fundamentales para estudiar la física del estado sólido (hay unos cincuenta sincrotrones en el mundo). Cfr. Farías y Cuevas (2018, p. 278).

Si no fuera por la necesidad de resolver problemas tecnológicos específicos, muchas propiedades de los cuerpos físicos nunca habrían sido examinadas y muchas de las teorías incorporadas posteriormente al cuerpo de la ciencia pura nunca se habrían formulado. Quizás los ejemplos más obvios se encuentren en las ciencias de la electrónica y la física espacial. El desarrollo de los ordenadores resultó de la sustitución de los tubos por los transistores. Al desarrollar los transistores se formularon muchas propiedades y leyes que rigen el comportamiento de los semiconductores, que de otra manera nunca se habrían formulado. (Skolimowski, 1966, p. 373)

El lector advertirá que, en opinión de Skolimowski, fue el diseño del transistor el que favoreció el descubrimiento de las leyes que gobiernan los semiconductores. Ahora bien, como se puede leer en la refutación de este caso histórico (Flórez, 2019), lo que aquí importa es mostrar, como bien lo indican Farías y Cuevas, que “la física del estado sólido es, sin dudas, la reina de las aplicaciones de la física” (p 272).

El término “física del estado sólido” se refiere a un área de la física que se desarrolló entre 1930 y 1960 (recordemos que el transistor se inventó en 1948), y que tiene su fundamento en la física cuántica. Como bien lo señalan Farías y Cuevas:

Podemos afirmar, sin exagerar, que la física del estado sólido hizo posible toda la tecnología más reciente, desde los materiales usados en electrónica, las células solares, luces LED, el láser de estado sólido, hasta los materiales superconductores que se emplean en la resonancia magnética nuclear. (Farías y Cuevas, 2018, p. 272)

Estos no son los únicos casos que sirven como evidencia a favor de la definición de la tecnología como ciencia aplicada. En el dominio de la nanotecnología y la biotecnología, podríamos mostrar otros casos.

Como es bien sabido, la nanotecnología es el área de investigación que estudia, diseña y fabrica materiales o sistemas a escalas nanoscópicas (Takeuchi, 2016, p. 15). También se define como la habilidad de controlar la materia átomo por átomo. Entre los artefactos que se han diseñado gracias a este dominio —en el que incluso detractores de la

fórmula como Vermaas (2011) encuentran evidencia a favor de la ecuación— están los *microscopios electrónicos*, los cuales usan electrones —en lugar de fotones— para formar imágenes de objetos diminutos. El diseño del microscopio electrónico se basó en la física cuántica y su historia se remonta, por lo menos a 1925, cuando Louis de Broglie sugirió que el electrón tiene una naturaleza ondulatoria.

Lo cierto es que la evidencia a favor no parece impresionar a los detractores de la fórmula, para quienes *solo un caso* en contra les disuade de abandonarla: se trata del célebre caso de la máquina de vapor. En lo que resta de este acápite, voy a adelantar un examen crítico de este caso paradigmático.

La tecnología como ciencia aplicada y la evidencia histórica en su contra: el caso de la máquina de vapor

Niiniluoto (1997) denomina *perspectiva estándar* a la tesis que identifica la tecnología con la ciencia aplicada, ya que, en su opinión, es el punto de vista más ampliamente defendido por quienes se han dado a la tarea de examinar la naturaleza de la tecnología. Sin embargo, dicha tesis no parece tan universal. De hecho, a finales de los sesenta, uno de los pocos filósofos de la tecnología que la defendió fue Mario Bunge⁵ (1966), y si se quiere hacer justicia a la evolución de sus ideas, Bunge fue también uno de los primeros filósofos de la tecnología en reconocer que *la tecnología es más que ciencia aplicada*, ya que “no es ajena a la teoría y no es la mera aplicación de la ciencia pura” (Bunge, 1960).

Al examinar el alcance de la ecuación [tecnología= ciencia aplicada] se tiene que, con base en ella se plantea que una *conditio sine qua non* para diseñar o inventar artefactos —que es una de las actividades propias de la tecnología—, la cual consiste en disponer de teorías y leyes científicas sistemáticamente formuladas, de las cuales, los diseños y artefactos son su aplicación. Dicho en otras palabras, sin la formulación explícita de principios científicos no sería posible diseñar

5 “The terms 'technology' and 'applied science' will be taken here as synonymous” (Bunge, 1966, p. 19).

artefactos. Ahora bien, en la historia de la tecnología parece que abundan los casos que muestran que ha sido posible desarrollar desde técnicas —por ensayo-error— hasta el diseño de artefactos sofisticados, *sin el beneficio de la ciencia*. Feibleman, por ejemplo, argumenta⁶ que en la práctica médica se conocían las propiedades curativas de la efedrina⁷, de la coca y de la quinina⁸ mucho antes de que se desarrollara la bioquímica como una disciplina científica consolidada (Feibleman, 1961, p. 310). Este caso histórico muestra que la técnica —por ejemplo, la medicina empírica— va un paso adelante de la ciencia. Esta conclusión también se ve respaldada por el caso de la fermentación. Bien se sabe que la técnica de la fermentación tiene miles de años de antigüedad. Según los registros paleo antropológicos, desde hace aproximadamente 9000 años el ser humano dominó la técnica de la fermentación para producir alcohol —según el antropólogo Patrick E. McGovern—. Sin embargo, los principios científicos que explican la naturaleza misma del proceso de fermentación no se descubrieron hasta Pasteur durante el siglo XVIII.

Si la evidencia histórica anterior no luce convincente, podríamos ofrecer dos casos históricos adicionales: el diseño de los molinos de viento y la técnica de los pigmentos. Los molinos de viento se extendieron por toda Europa alrededor del siglo XII, sin embargo, algunos de los principios científicos que explican su funcionamiento, entre ellos

6 Para un análisis crítico de este argumento y otros, cfr. Flórez y García (2017).

7 Que se conocía en extremo oriente como *Ma huang*, y era ampliamente utilizada en la medicina tradicional china.

8 La quinina era conocida por sus propiedades curativas por los nativos americanos, pero no se incorporó al acervo cultural europeo hasta que no fueron descubiertas sus propiedades antimaláricas. Cuando los europeos trajeron la malaria a América, los nativos descubrieron que una de sus medicinas tradicionales, la quina o corteza del quino, ofrecía alivio a los síntomas de esta enfermedad. Los incas conocían las propiedades medicinales de las plantas que crecían en los Andes y en la selva amazónica, entre ellas había un árbol que producía la amarga corteza que podía curar muchas dolencias, calambres resfríos y arritmias. La palabra quechua “quina” significa corteza, pero esta corteza de propiedades extraordinarias se conoció con el nombre de quina-quina, “corteza de cortezas” dando así origen al nombre quinina (cfr. Wikipedia).

la dinámica de fluidos, no fueron formulados sino hasta el siglo XVIII, específicamente en 1738, cuando Daniel Bernoulli publicó sus estudios sobre hidrodinámica. Por otra parte, es un hecho histórico que mientras la técnica de los pigmentos es milenaria, la teoría científica que explica la estructura química de algunos pigmentos (ya que no todos son compuestos de coordinación) es relativamente reciente. La teoría química de los compuestos de coordinación fue formulada en 1893 por Alfred Werner. Gracias a las contribuciones teóricas de Werner, la química inorgánica dispuso, a partir de entonces, de una nomenclatura sistemática de los compuestos de coordinación, entre ellos algunos de los pigmentos sintetizados en el siglo XVIII (cfr. Flórez et al., 2019)⁹.

Para ofrecer una réplica a estos contraejemplos, se podría simplemente decir —tal y como lo plantea Hans Radder— que:

La explicación de Bunge podría ser interpretada como una definición de la tecnología y, como tal, estaría inmune a los contraejemplos empíricos. Si un caso determinado no encaja en la descripción

9 De ser necesario, podríamos ofrecer casos históricos de la misma naturaleza, *ad nauseam*. Por ejemplo, la brújula fue inventada mucho antes de dispusiéramos de una teoría científica que explicara su funcionamiento. Los chinos la inventaron en el siglo IX con el propósito de ubicarse a mar abierto, pero los estudios científicos sobre el magnetismo se adelantaron siete siglos después por varios científicos destacados, entre ellos William Gilbert en su reconocida obra *De magnete*. La explicación más sofisticada fue formulada en el siglo XIX por Maxwell en su teoría electromagnética. A pesar de este tipo de evidencia, no debe sorprendernos que se hayan inventado técnicas que preceden en siglos a la ciencia que explica su éxito, dado que históricamente —y, en consecuencia, ontológicamente— la técnica es muchísimo más antigua que la ciencia. La paleoantropología muestra que los subhomínidos dominaron la técnica mucho antes de alcanzar la capacidad de pensar y razonar. De hecho, el *Homo habilis*, quien precede en el árbol filogenético al *Homo sapiens*, ya tenía la habilidad de diseñar utensilios de piedra. Por eso en los estudios de paleoantropólogos y paleogenéticos se otorga tanta importancia a la capacidad para fabricar y usar instrumentos. Cabe agregar que si bien la técnica tiene millones de años de existencia, la tecnología es, al igual que la ciencia, una recién llegada a la cultura humana. Con esta observación no pretendemos sostener que la tecnología sea un fenómeno exclusivamente moderno, ya que hay tecnología antigua (por ejemplo, la invención de la polea por Arquímedes) —tan antigua como la ciencia— y la tecnología moderna es todavía más joven, pues su origen se halla en la primera Revolución Industrial.

de la tecnología como ciencia aplicada, entonces, no sería, por definición, una tecnología. (Radder, 2009, p. 71)

Sin embargo, puesto que el caso que se considera históricamente como una prueba palmaria —e irrefutable— de que la tecnología no es ciencia aplicada es el diseño de la máquina de vapor, no sería suficiente con una contrarréplica en la línea de Radder, ya que parecería muy contraintuitivo concluir con él simplemente que una máquina tan sofisticada como el motor a vapor no es una instanciación de tecnología. Justamente por ello se propone rebatir el argumento en el propio terreno en el que ha sido formulado, a saber, el de la historia de la tecnología.

De acuerdo con el desarrollo histórico de la invención de la máquina, Feibleman recuerda que su diseño precedió en más de medio siglo a la teoría que explica su funcionamiento: “Carnot descubrió la ciencia pura de la termodinámica como resultado de sus esfuerzos por mejorar la eficiencia del vapor” (1961, p. 314). En virtud de ello, es frecuente encontrar en la literatura historiográfica y filosófica la afirmación de que “es más lo que la ciencia le debe a la máquina de vapor, que lo que la tecnología le adeuda a la ciencia” (Ihde, 2013, p. 54). Los historiadores Derry y Williams concuerdan con Ihde. Para ellos, la máquina de vapor convierte la energía térmica en energía mecánica, y su rendimiento se explica por las leyes de la termodinámica. Pero estas leyes no fueron establecidas hasta mediados del siglo XIX y, por lo tanto, en modo alguno se había entendido la naturaleza del calor antes de tal descubrimiento. Esta circunstancia recuerda una vez más que tecnología no es, de ninguna manera, sinónimo de ciencia aplicada (Derry y Williams, 1960, pp. 494-495).

Pieter Vermaas et al. también son partidarios de esta idea:

Después de 1970, la perspectiva de la ciencia aplicada fue abandonada por la mayoría de los autores que reflexionaron sobre la naturaleza y los métodos de la ingeniería. Un argumento importante en contra de esta visión consistió en que se habían producido todo tipo de desarrollos tecnológicos que sólo mucho más tarde pudieron ser respaldados científicamente. Las primeras generaciones de máquinas de vapor, por ejemplo, se construyeron sin que los científicos supieran

cómo describir el funcionamiento de esos artefactos particulares. De manera similar, los transistores y, más recientemente, los superconductores de alta temperatura se crearon antes de que los científicos pudieran comprender con precisión qué procesos físicos eran responsables del funcionamiento exitoso de tales innovaciones. En estos casos, las aplicaciones industriales revolucionarias llegaron mucho antes que la comprensión científica. (Vermaas et al., 2011, p. 56)

En efecto, mientras la máquina de vapor fue patentada por el ingeniero mecánico James Watt en 1769, tuvimos que esperar cincuenta años para que la ciencia que explica su funcionamiento, por ejemplo, que la termodinámica “viera la luz” cuando Sadi Carnot publicó, en 1824, sus *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*. También se sabe que casi 81 años después Lord Kelvin y Clausius publicaron su descubrimiento de las leyes de la termodinámica. Con base en estos acontecimientos se argumenta que, puesto que el diseño de la máquina de vapor precede en décadas a la teoría científica que explica —de manera sistemática— su funcionamiento, no es legítimo identificar la tecnología con la ciencia aplicada. No obstante, la opinión que favorecieron mayoritariamente filósofos e historiadores según la cual Watt inventó la máquina de vapor sin echar mano de ningún recurso teórico es falsa.

Los registros históricos muestran que Watt fue contratado en la Universidad de Glasgow en 1757 como ingeniero y allí conoció a quien fuera su maestro: el célebre químico Joseph Black, reconocido en la historia de la química por su descubrimiento del calor latente. Watt y Black mantuvieron una relación estrecha hasta que este último se mudó a Edinburgh en 1766. Aunque Watt insistió en que había llegado a las mismas conclusiones que Black de manera independiente, lo cierto es que Black no solo ejerció un fuerte influjo en su pensamiento, sino que además le asesoró —a través de abundante correspondencia de la cual hay evidencia— en el diseño y montaje de muchos de los experimentos que Watt realizó en el tiempo previo a la invención del condensador de la máquina de vapor¹⁰.

10 Para una ampliación de este punto, ver Flórez, D. (2023). Semantic changes in Technology. En L. Anderson y E. Lepore (Eds.), *The Oxford Handbook of Applied Philosophy of Language* (pp. 594-612). Oxford University Press.

Un examen cuidadoso de la historia de la invención de la máquina de vapor revela que Watt disponía de recursos teóricos que sirvieron de base para el diseño e invención del motor a vapor, pues estaba firmemente comprometido con la que se conocería como teoría material del calor. Aunque esta es una teoría aún más arcaica —o primitiva— que la misma teoría del calórico¹¹, la teoría material del calor planteaba que el calor era una sustancia química (Miller, 2009, p. 44). Dicha teoría también postulaba la existencia del calor latente. La terminología asociada a este concepto fue introducida por Black para explicar los cambios de estado (por ejemplo, líquido a gaseoso). Tanto Black como Watt entendían el calor latente como una propiedad química intrínseca de la materia ordinaria. De hecho, Black había interpretado “el calor latente” como la *causa* de los cambios de estado de la materia: el agua se convertía en vapor, por ejemplo, por el calor adicional que tenía que ser suministrado, y que no era “detectable por el termómetro” (razón por la cual lo llamó “calor latente” o calor oculto) y que se combinaba con la materia del agua para producir vapor (cfr. Miller, 2009, p. 45). Sin embargo, tal y como lo plantea Miller en su estupendo libro, Watt exploró los alcances de esta idea mucho más allá de lo que lo hiciera el propio Black (p. 45), tanto así que consiguió formular la bien conocida ley de Watt, la cual estipula que la suma del calor latente y sensible es una constante. “Watt no sólo usó

11 La teoría del calórico fue una teoría ampliamente defendida durante el siglo xviii que explicaba el calor como un fluido hipotético que impregnaba la materia. Entre sus principales defensores se encuentra Antoine Lavoisier, considerado uno de los padres de la química moderna. Las ideas del maestro de Watt fueron decisivas para la maduración misma de la teoría del calórico: “A partir de las ideas de Black se afianzó la teoría que defendía la existencia de un fluido invisible que entraba y salía de una sustancia aumentando o disminuyendo su temperatura. La teoría del calórico se basaba en dos premisas fundamentales: i) el fluido no se crea ni se destruye, ii) la cantidad de calórico transportado hacia o desde el objeto es directamente proporcional a la masa y a la temperatura del objeto. De esta forma, al introducir más calórico en una sustancia, esta se calentaba, hasta que finalmente el calórico se desbordaba y fluía en todas direcciones. Esta era la razón por la cual la calidez de un objeto al rojo vivo se dejaba sentir a gran distancia; el calor del sol, por ejemplo, se notaba a 150 millones de kilómetros. Al poner en contacto un objeto caliente con otro frío, el calórico fluía desde el primero al segundo. Ese flujo hacía que el objeto caliente se enfriara y que el frío se calentara” (Taton, 1972, p. 150).

esta idea en su desarrollo del ‘trabajo expansivo’ de la máquina de vapor, sino que también, y de un modo aún más importante, en sus ideas sobre la transformación del agua en aire” (Miller, 2009, p. 45).

Es por ello que resulta sorprendente que los filósofos de la tecnología (e incluso muchos historiadores) hayan ofrecido una imagen distorsionada del diseño de la máquina como resultado de la aplicación del método de ensayo-error *simpliciter*, en total ausencia de recursos teóricos, y de James Watt no como un hombre de ciencia, sino como un mecánico muy ingenioso:

Nos predisponemos a considerar que Black habita en el mundo de la “teoría” y Watt en el de la “práctica”. Se nos anima a ver a Black como una figura cosmopolita de la Ilustración escocesa y a Watt como un ingeniero inteligente colgado de sus faldas, y que su relación implica una “transferencia de intelecto a artesanía”. Esto es un error. (Miller, 2009, p. 89)

Lo cierto es que tal vez lo logros de Watt en el dominio de la ingeniería mecánica terminaron por opacar sus contribuciones en el dominio de la química que, en retrospectiva, corresponden a los primeros conatos de la química científica, lo cual nos hizo olvidar —como bien lo reconoció el ingeniero Andrew Ure— “que la máquina de vapor fue el triunfo más espléndido que se erigió sobre la ciencia del calórico” (Miller, 2009, p. 71).

Ahora bien, quienes rechazan la tesis de que la tecnología es ciencia aplicada, con base en este caso histórico, podrían conceder que Watt disponía de recursos teóricos para ser “aplicados” en su diseño del motor a vapor, sin embargo, podrían insistir en que dichos recursos teóricos eran demasiado rudimentarios como para explicar de manera adecuada el funcionamiento de la máquina de vapor. Con el propósito de hacer el argumento más contundente, podrían añadir que la termodinámica (que emergió cincuenta años después) sí explica con notable éxito el funcionamiento de la máquina de vapor, puesto que mientras la teoría material del calor es evidentemente falsa, la termodinámica es verdadera.

Dicha refutación está lejos de ser incontestable. El argumento mencionado tiene el defecto —para nada despreciable— de incurrir en un

anacronismo. Para aclarar lo anterior, es importante recordar que, en la ecuación que identifica la tecnología con la ciencia aplicada, el alcance del término “ciencia” debe interpretarse de manera restrictiva y presuponer a la ciencia disponible en su momento o sus consecuencias teóricas o prácticas. Si se concede esto, se podría imaginar una línea de tiempo en la cual el diseño de los artefactos es la aplicación de los recursos teóricos de la ciencia disponible, que en todo caso siempre son tentativos y son susceptibles de mejoramiento permanente.

La renuencia a admitir la ecuación [tecnología = ciencia aplicada] revela ciertos prejuicios sobre la naturaleza de la ciencia y sobre la manera de interpretar correctamente los hechos de la historia de la ciencia. Cuando se propone interpretar la expresión “ciencia disponible” desde una perspectiva histórica y pedimos que se tomen en cuenta las implicaciones y consecuencias teóricas y prácticas de las teorías científicas que hay tras un desarrollo tecnológico particular, se rechaza una visión parcial y exageradamente demandante sobre la ciencia desde la cual se rechazan las explicaciones incompletas y, sobre todo, las que provienen de teorías falsas o superadas por competidoras más capaces¹².

Así las cosas, queda claro que la historia nos muestra la relación de aplicabilidad entre el conocimiento científico y el tecnológico. ¿Significa esto que la tecnología es un tipo de conocimiento de inferior rango, como suponía Aristóteles (ver primera sección)? ¿O se trata más bien de una invitación a pensar la especificidad del conocimiento tecnológico? La siguiente sección busca presentar este recorrido que se encuentra en diferentes perspectivas.

Conocimiento tecnológico y técnico

Cuando se enfatiza que la tecnología es un conocimiento productivo, se recurre a las visiones sobre el diseño, de las cuales examinamos

12 Un análisis más detallado de esta discusión lo adelantamos en Flórez y García (2019).

cuatro. Cuando se centra en la capacidad de acción o uso que supone el conocimiento tecnológico, se recurre a las explicaciones sobre el conocimiento práctico (*know-how*) y sobre las habilidades técnicas y manuales.

Especificidad del conocimiento tecnológico

Supóngase el siguiente escenario imaginario: tras realizar un viaje en el tiempo, Platón visita una de las legendarias bibliotecas de la Universidad de Oxford. Durante el viaje adquiere todo el caudal del conocimiento filosófico acumulado hasta 1944. Una vez, instalado allí —y ante la mirada atónita de los estudiantes—, pide a la bibliotecaria que le conduzca a los estantes donde se hallan los libros más importantes de epistemología a partir de 1945.

La bibliotecaria le entrega algunos gruesos ejemplares. Después de hojearlos un par de horas, se sorprende al advertir que los epistemólogos —desdeñando la navaja de Ockham— han multiplicado las variedades de conocimiento, a tal punto de que en la literatura contemporánea relevante se dispone de una docena de categorías, entre ellas: el conocimiento incorporado (o corporizado, o “embodied knowledge”) (Merlau-Ponty, 1945); *el know that vs el know how* (Ryle, 1949); el conocimiento tácito vs. el conocimiento explícito (Polanyi, 1958); el conocimiento visualizado (Ferguson, 1992); o el conocimiento material (o “thing knowledge”) (Baird, 2004). Tras una larga meditación, el fundador de la Academia reflexiona: *si atribuimos la capacidad de conocer al cuerpo —o incluso a partes de este—; y a objetos; nada parece impedir que atribuyamos conocimiento a un martillo o a mi ceño fruncido*¹³.

13 No sabrá advertirle a la lectora que esta es sólo una reacción imaginaria. Realmente, tanto Platón como su discípulo aventajado Aristóteles reconocieron que la τέχνη es una forma de ἐπιστήμη (episteme) o de conocimiento, como se vio en la sección anterior. Para un análisis más detallado sobre esta concepción que, parece haberse perdido en el tiempo, recomiendo el estimulante capítulo de Mitcham y Waelbers (2009). Por otro lado, estas no son las únicas variantes de conocimiento. También se podría incluir el conocimiento teórico vs. el conocimiento práctico y el conocimiento operacional.

Tras ensayar este juego de la imaginación, se le anticipa a la lectora que en esta sección se parte del supuesto de que las variantes arriba mencionadas son instancias de conocimiento genuino, por lo que no se discutirá si el *know how* (en adelante, *saber como*) es realmente o no un tipo de conocimiento distinto del *know that* (en adelante, *saber que*). En su lugar, se argumentará que los siguientes dos planteamientos —que suscriben los epistemólogos contemporáneos de la tecnología— son falsos. A esta posición filosófica se le denominará la *perspectiva estándar*. Los planteamientos afirman lo siguiente:

P1: La característica epistémica distintiva de la tecnología es que es un *saber como*, por ejemplo, P1 identifica a la tecnología con el *saber como* —*simpliciter*—. Esta es la posición que más ampliamente se ha defendido en la literatura en las dos últimas décadas, como puede verse en De Vries y Meijers (2013), Marc de Vries (2016) y Vermaas et al. (2013).

P2: Puesto que una parte importante del conocimiento tecnológico es no-proposicional (por ejemplo, el *saber como*; el conocimiento corporizado; el conocimiento visualizado), el *saber que* no captura la naturaleza del conocimiento tecnológico.

Para mostrar por qué ambos planteamientos son falsos, primero conviene examinar en qué consisten algunas de las variantes de conocimiento que sirven —de acuerdo con la *perspectiva estándar*— al análisis de la estructura epistémica de la tecnología. Comencemos con la clásica distinción de Ryle entre *saber que* y *saber como*, ya que el *saber como* parece subsumir —*prima facie*— otras variantes de conocimiento.

La distinción de Ryle entre “saber como” y “saber que”

En su célebre *The Concept of Mind* (1949), Gilbert Ryle introduce una distinción, que es decisiva en los debates epistemológicos del siglo XXI, especialmente en el dominio de la epistemología de la tecnología. Se trata de la distinción entre *saber que* y *saber como*. El primer tipo de saber se caracteriza por comprender todo aquel conocimiento que se puede articular mediante proposiciones, *por ejemplo, sabemos que* “el

Blackbird —un avión militar diseñado en los años sesenta— triplica la velocidad del sonido”. El segundo tipo de saber comprende todas las habilidades, destrezas y disposiciones, tales como ser capaz de montar en bicicleta, atarse los cordones o nadar en estilo mariposa¹⁴.

Algunas de las discusiones en la epistemología contemporánea, que giran en torno a esta distinción, se ocupan de examinar la independencia del *saber como* respecto del *saber que*, ya que hay quienes sostienen que el *saber como* es, finalmente, una variante del *saber que*. A esta posición se le conoce como “intelectualismo”. El intelectualista defiende —*prima facie*— las siguientes dos tesis:

TI1: el *saber como* es una variante del *saber que*. O dicho de otra manera, el *saber como* es una forma de *saber que*.

TI2: el *saber como* y el *saber que* no son dos tipos de conocimiento distintos; y no son independientes.

Por contraste, el antiintelectualista mantiene que:

TA1: el *saber como* y el *saber que* son dos tipos de conocimiento distintos en naturaleza y son independientes.

Estas no son las únicas posiciones filosóficas que se podrían mantener: también se podría defender —*prima facie*— un antiintelectualismo moderado —justo como el suscrito aquí—. Este plantea que:

TAm: el *saber como* y el *saber que* son dos tipos de conocimiento distintos en naturaleza, pero en lo que concierne a la tecnología, estos dos tipos de saberes no son independientes.

14 Realmente no es evidente la relación entre *saber como* y habilidad. Para un análisis detallado acerca de los contraejemplos que se han ofrecido en la literatura filosófica contemporánea, y algunas de las teorías explicativas sobre la naturaleza del *saber como*, ver Knowledge How de Jeremy Fantl (2017).

Como se indicó arriba, la *perspectiva estándar* mantiene que la tecnología se puede identificar con el *saber como* —*simpliciter*—. Así lo sostienen, entre otros, De Vries y Meijers:

El conocimiento tecnológico se expresa en parte de manera no proposicional. Hay varias razones para esto. En primer lugar, una parte importante del conocimiento tecnológico es del tipo “saber como”. Ryle introdujo en su libro *The Concept of Mind* (1949) este término para indicar un conocimiento similar a la habilidad que no puede expresarse completamente en proposiciones (por ejemplo, saber cómo martillar un clavo en una pieza de madera). A menudo, ese tipo de conocimiento permanece tácito (Polanyi, 1967). Está incorporado en las personas y es difícil de transferir a otras personas. En segundo lugar, como argumentó Ferguson (1992), los dibujos y diagramas técnicos contienen conocimientos que son casi imposibles de expresar en términos proposicionales. (De Vries y Meijers, 2013, p. 73)

En el pasaje anterior, ambos autores vacilan sobre el carácter no-proposicional del conocimiento tecnológico, pues afirman que “el conocimiento tecnológico se expresa *en parte* de modos no-proposicionales”. Además, sostienen que: “*una parte importante* del conocimiento tecnológico es del tipo que Ryle denomina, ‘saber como’”.

Con ello, pareciera que De Vries y Meijers (2013) dejan un resquicio al conocimiento proposicional, al subrayar que una parte del conocimiento tecnológico no es proposicional¹⁵; por lo que cabría preguntarse lo siguiente: si se imagina el conocimiento tecnológico como un círculo, y una parte de su área corresponde al saber no-proposicional —o al *saber como*—, ¿cómo sería el conocimiento restante?

15 De Vries también sostiene: “Existe conocimiento tecnológico que no se puede expresar adecuadamente en proposiciones. Cuando un carpintero dice que sabe cómo golpear el clavo en el lugar exacto para que se vaya directo a la madera, probablemente no pueda expresarlo con palabras. Es más probable que diga algo como: ‘bueno, simplemente lo sé; no sé exactamente cómo, pero siento cómo hacerlo’. Gilbert Ryle, basado en este tipo de consideración, ha definido una distinción entre saber-que y saber-cómo” (De Vries, 2016, p. 25).

A este respecto, véase la figura 8, pero nótese inicialmente el lado derecho del círculo, que representa el conocimiento no proposicional.

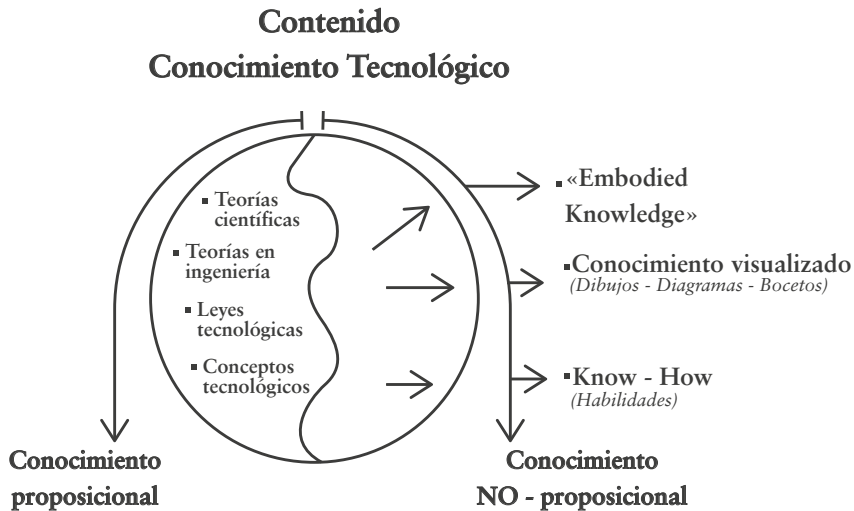


Figura 8. Contenido del conocimiento tecnológico

Fuente: elaboración propia.

Una parte del conocimiento no-proposicional corresponde al denominado “conocimiento corporizado”. Lo cierto es que ni en el pasaje citado anteriormente —ni el resto del artículo— los autores ofrecen ejemplo alguno de lo que podría ser considerado una instancia de conocimiento corporizado.

Como es bien sabido, en la literatura filosófica hay distintas explicaciones sobre la naturaleza misma del conocimiento corporizado. Este término se remonta a la fenomenología de Merleau-Ponty, particularmente a su trabajo *Phénoménologie de la perception* (1945). Allí, el filósofo francés sostiene que *el conocimiento incorporado* o corporizado se caracteriza porque es *el cuerpo el que sabe* —inter alia— *cómo actuar* (por ejemplo, saber cómo bailar o cómo andar en bicicleta). Este conocimiento no es el resultado de la deliberación racional. Tampoco es necesario verbalizarlo, ni representar en la mente

todos los procedimientos requeridos para ejecutar los movimientos que desencadenan la acción. Se trata de un conocimiento que parece estar impreso en el cuerpo. El sujeto cognoscente aquí es el cuerpo mismo, no la mente. (O más precisamente, es la mente-cuerpo, por ejemplo, el conocimiento “que tienen” las manos al teclear las letras en el teclado del computador). Este tipo de conocimiento además es inconsciente y no-articulado.

Con base en esta perspectiva, se puede afirmar que:

X es una instancia de conocimiento corporizado

1. Es el cuerpo —o alguna parte del cuerpo— el que sabe.
2. X es pre-reflexivo¹⁶ (lo que significa que, pese a ser intencional, no requiere de la deliberación).
3. X es no-proposicional (por ejemplo, no es necesario articularlo mediante proposiciones).

Teclear es una instancia de conocimiento corporizado dado que: 1) son mis manos las que saben cómo desplazarse a través del teclado —raras veces con el auxilio de la vista—; 2) solo en la etapa de aprendizaje es necesario pensar dónde está cada una de las letras, pero una vez se adquiere la destreza, se puede teclear hasta con los ojos cerrados; 3) para teclear no preciso de proposiciones.

Bien se podría conceder, lejos de la vieja tradición filosófica, que hay conocimiento corporizado. Sin embargo, resulta difícil identificar instancias de este tipo de conocimiento en el dominio de la ingeniería y que sean distintivos de ella. Piénsese, por ejemplo, en el diseño. No hay nada más deliberado que un diseño ingenieril. Ni siquiera los

16 La experiencia corporal es, según Merlau-Ponty, pre-reflexiva. Cabe aclarar que lo pre-reflexivo es una forma de conciencia porque es intencional, por ejemplo, cuando camino, el movimiento guarda una relación intencional con la superficie.

descubrimientos serendípicos satisfacen la condición de no ser el resultado de una laboriosa deliberación¹⁷.

Hay otras explicaciones disponibles acerca de la naturaleza del conocimiento incorporado, como la de Gibson (1979), quien sostiene que una instancia de conocimiento corporizado correspondería a una propiedad física del entorno que le ofrece (*affordances*) a un individuo la oportunidad de realizar una acción particular. Por ejemplo, cuando caminamos a través de una abertura, la relación entre la abertura y la anchura de nuestros hombros determina nuestra acción. Los *affordances* son el conjunto de estímulos a partir de los cuales un objeto proporciona a un agente la oportunidad de ejecutar una acción.

Con base en la perspectiva gibsoniana, se tiene que:

X es una instancia de conocimiento corporizado *syss*

- (i) X es un objeto que posibilita (ofrece, o permite) al sujeto Y realizar *n* acciones.

Esta caracterización del conocimiento corporizado podría contar con instancias a su favor en el dominio de la tecnología. Supóngase, por ejemplo, que la máquina de vapor (qua objeto *n*) posibilitó al ingeniero mecánico James Watt realizar *n*, ejecutar un curso de acciones, que lo llevaron al diseño del condensador.

Es bien conocido por todos que a las manos de Watt llegó un modelo de la máquina de Newcomen cuyo rendimiento no era satisfactorio. Watt se dio cuenta de que la causa del funcionamiento inadecuado de la máquina era el hecho de tener que enfriar el cilindro con el propósito de

17 Aunque en el dominio de *la tecnología* resulta difícil advertir instancias de conocimiento corporizado pre-reflexivo, sí abundan casos en el dominio de *la técnica*. Las investigaciones que se adelantan en los estudios de la denominada “cognición 4e” (“embodied”: encarnada; “embedded”: incrustada; “extended”: extendida; “enactive”: enactiva) podrían resultar tremendamente iluminadoras para el examen de la estructura epistémica —y cognitiva— de la técnica. Adicionalmente, para un análisis de las características de las habilidades (*skills*) y sus estadios, ver Dreyfus (1999).

condensar el vapor. Fue en mayo de 1765 cuando Watt formuló la idea de un condensador separado y en 1769 registró finalmente la patente.

Ahora bien, la dificultad para considerar este caso como una genuina instancia de conocimiento incorporado estriba en que Watt bien pudo seguir un circuito de razonamiento que exhibiría la siguiente estructura: “Si se pudiese mantener el cilindro siempre caliente, entonces cabría esperar un mejor rendimiento de la máquina. Como se puede advertir, un caso así no satisfaría la naturaleza pre-reflexiva de los *affordances* Gibsonianos”.

Con base en lo anterior, se puede entender que en el dominio de las técnicas hay instancias de conocimiento corporizado, por ejemplo, la técnica del grabado al buril requiere —*inter alia*— habilidades tales como: 1) dibujar excavando líneas sobre una matriz de metal ayudándose exclusivamente con el buril; y 2) ejercer una fuerte presión en el metal, de suerte que la incisión realizada provoque que se aloje en ella una mayor cantidad de tinta. Sin embargo, en el dominio de la tecnología, los movimientos que exigen las acciones en el diseño, por ejemplo de una planta nuclear, deben ser cuidadosamente planificados y ejecutados.

Es cierto que no todas las instancias del conocimiento tecnológico son resultado de la deliberación racional. Como bien se ha advertido en la literatura filosófica, hay fases en el diseño mismo, especialmente en las etapas iniciales, que podrían considerarse instancias de lo que Polanyi llamó “conocimiento tácito” (Young, 2018). Este se caracteriza, al igual que el conocimiento corporizado, por ser no-proposicional (por ejemplo, ambos difícilmente se pueden articular a partir de proposiciones). *La intuición* que exige el diseño —entendido como la capacidad de reconocer las soluciones adecuadas— es una instancia de conocimiento tácito.

¿Podría ser la intuición, además, una instancia de conocimiento corporizado? A partir de la teoría de los *affordances* gibsonianos, podríamos decir que sí, puesto que la intuición en el diseño se agudiza a partir de las posibilidades mismas del entorno o de las condiciones idiosincráticas (constreñimientos materiales, políticos, económicos específicos) de las cuales emerge el problema a resolver. Dicho en otras palabras, es el entorno mismo el que ofrece (*affordances*) a la ingeniera

los factores entre los cuales ella tendrá que decidir cuáles son relevantes y cuáles no.

Ahora, supóngase que hay suficiente respaldo empírico que apoya la tesis de que, en efecto, hay conocimiento corporizado y conocimiento tácito en la tecnología. Con todo, estos no serían tipos de conocimiento propios —o exclusivos— de la tecnología: si la máquina de vapor (o cualquier otro artilugio tecnológico) es una suerte de *affordance*, también lo tendrían que ser los prismas con los cuales Isaac Newton llevó a cabo el experimento del doble prisma.

Cuando Newton pudo por fin hacerse a un segundo prisma —gracias al dinero que reunió producto de sus cátedras en Cambridge— realizó el siguiente experimento: en 1666 consiguió un prisma triangular para estudiar el célebre fenómeno de los colores, para lo cual oscureció la habitación. Hizo un pequeño agujero en una de las ventanas, con lo cual entraba un chorro de luz —proveniente del sol—. Luego puso el prisma, de manera que la luz se refractaba hacia la pared opuesta. Dice Newton:

Al principio fue una diversión muy placentera ver los colores vívidos e intensos producidos, pero después los consideré con mayor circunspección, quedé sorprendido al verlos de una manera oblonga porque de acuerdo con las leyes de la refracción deberían tener forma circular. (Lozano Leyva, 2007, pp. 94-95)

La conclusión que Newton derivó es que la luz blanca debía contener todos los colores, por lo que añadió un segundo prisma, en el cual todos los colores proyectados se convirtieron —de nuevo— en un haz de luz blanca.

Nótese que los prismas ofrecieron —o permitieron— a Newton la posibilidad de explorar las cualidades de la luz y de los colores, a partir de un experimento que parece un verdadero acto de prestidigitación.

Este caso, sirve a su vez para mostrar por qué la intuición —como una instancia de conocimiento tácito— no solo juega un rol decisivo en los diseños, sino también en la experimentación: ¿qué se intuye en el experimento del doble prisma? Que si el haz de luz blanca contiene todos los colores, tal y como lo exhibe el paso del chorro de luz

por el primer prisma, una vez se hiciesen pasar todos los colores por el segundo prisma, estos tendrían que transformarse, de nuevo, en un chorro de luz blanca¹⁸.

De vuelta al análisis de la perspectiva estándar, nótese lo que afirman Pieter Vermaas et al:

El conocimiento tecnológico tiene una faceta de “saber cómo”: los filósofos, siguiendo a Ryle, G. (1949), a menudo distinguen dos tipos de conocimiento. Uno es el tipo de conocimiento que puede ser expresado completamente en términos de afirmaciones; el otro no lo es o, más al punto, no puede ser expresado de esa manera. (Vermaas et al., 2013, p. 63)

De acuerdo con Vermaas, un ejemplo paradigmático de *saber como* es el de poder montar en bicicleta. Supongamos que leemos toda la literatura disponible acerca de cómo es montar en bicicleta. Este saber es completamente vano, si no nos montamos en una bicicleta y ejecutamos todos los movimientos que se requieren para desplazarnos en ella. El *saber como* es un tipo de conocimiento que se adquiere haciendo. Dicho en otras palabras, por un lado, tenemos el dominio cognitivo del conocimiento proposicional y, de otro lado, el dominio psico-motor al que corresponden las habilidades que implica bailar, nadar, montar en bicicleta, operar una máquina, entre otras.

Supóngase que se dispone de una explicación no problemática del *saber como*, y que esta variante del conocimiento comprende un conjunto de habilidades, destrezas y disposiciones. Las habilidades en ingeniería, por ejemplo, son muy variadas. Entre ellas, sobresale *la*

18 La solución a problemas matemáticos también requiere de un proceso parsimonioso y muy elaborado, en el que probablemente también se aplica un agudizado “sentido” de la intuición. La intuición metafísica o filosófica también ha sido una fuerza motriz para el impulso en la formulación de hipótesis científicas. Piénsese, por ejemplo, en el atomismo clásico, el cual sin lugar a dudas es un precedente de la física atómica que surge a finales del siglo XIX. También podríamos considerar la intuición de Bruno sobre la existencia de innumerables mundos, y de la infinitud del universo.

*creatividad*¹⁹, puesto que se supone que la virtud distintiva de una ingeniera es el ingenio.

En la ingeniería, la creatividad es, o bien la capacidad de formular soluciones innovadoras, o bien la habilidad para plantear una solución técnica enteramente nueva (invento) para resolver un problema práctico. James Watt tuvo, sin lugar a dudas, una mente genial y creativa, pues no solo advirtió que la ineficiencia de la máquina de Newcomen se debía al hecho de que, en el mismo cilindro donde tenía lugar la evaporación del agua, se debía enfriar el agua para ser condensada, sino que adicionalmente ideó la manera de resolver el problema, a saber: diseñar un cilindro independiente: el condensador.

Con base en ello, podría decirse que,

X tiene la habilidad de ser creativo²⁰ —en tecnología— syss:

1. X formula una solución innovadora para resolver un problema práctico²¹.

o

2. La solución técnica que X plantea a un problema práctico conduce al descubrimiento de un nuevo principio de funcionamiento; o de nuevos materiales. “Se trata de artefactos que tienen propiedades nuevas, que utilizan componentes que no habían

19 La creatividad es la habilidad de generar ideas o artefactos que sean nuevos, sorprendentes y valiosos. Las “ideas” incluyen aquí conceptos, poemas, composiciones musicales, teorías científicas, recetas de cocina, coreografía, chistes, entre otros. Los “artefactos” incluyen pinturas, esculturas, motores a vapor, aspiradoras, cerámica, origami, silbatos de centavo, y muchas otras cosas que se puedan nombrar (Boden, 1990/2004).

20 También hay creatividad en filosofía. En filosofía, la creatividad comprende desde la capacidad de plantear un nuevo problema filosófico hasta un contraejemplo: por mucho que no estemos de acuerdo con el carácter concluyente de los contra-ejemplos de Gettier, ello no les resta ingenio.

21 Un proyecto puede ser original y creativo, sin que ello suponga ninguna novedad desde un punto de vista estrictamente técnico (Quintanilla, 2017).

sido nunca utilizados para funciones equivalentes y cuya estructura, por lo tanto, es completamente original” (Quintanilla, 2017, p. 113).

El asunto es que la creatividad no es una virtud epistémica exclusiva de los ingenieros. Una mente creativa puede cosechar frutos en cualquier dominio del conocimiento. Piénsese, por ejemplo, en el descubrimiento de las lunas de Júpiter por Galileo, cuando apuntó con un telescopio, aún muy rudimentario, hacia este planeta. Recuérdsese que el telescopio se había diseñado con fines comerciales. El mérito en el uso del artefacto está en que el científico italiano fue el primero en usarlo para escudriñar el cielo, y gracias a esa audacia “vio lo que nadie había visto antes”.

Tras varias noches de observaciones pacientes, Galileo observó que los cuatro cuerpos a los cuales les seguía la pista no abandonaban la vecindad de Júpiter y parecían moverse con el planeta. Con base en ello, derivó la conclusión de que se trataba de satélites²².

Ahora, no se analizarán las condiciones singulares que favorecen el pensamiento creativo, ya que esto es de competencia de la psicología de la ciencia, pero en la siguiente subsección, desde la perspectiva de Simondon y Rabardel, se harán algunas acotaciones. En esta subsección interesa, más bien, mostrar que las instancias de *saber como* en la ingeniería no son exclusivas de este dominio. Dicho de otro modo, habilidades tales como ser ingenioso, ser creativo e intuitivo también se cultivan en otros dominios del conocimiento, tales como las ciencias básicas, la filosofía, las artes.

Por supuesto, se podría objetar que lo que está en discusión realmente es que este tipo de conocimiento es no-proposicional; sin embargo, si queremos adelantar un análisis epistemológico provechoso sobre la tecnología, la vía que hay que tomar es aquella que conduce al examen del contenido conceptual del conocimiento tecnológico.

22 Para una exposición detallada de los descubrimientos astronómicos de Galileo, ver García Duque (1997).

Entre otras razones, porque hay muchas habilidades en la ingeniería que son conceptuales. Piénsese, por ejemplo, en la habilidad analítica. “Analizar” en ingeniería significa lo mismo que en química o en filosofía, por ejemplo, no es otra cosa que descomponer o aislar elementos. Por supuesto, varía lo que se aísla según el dominio del conocimiento. En el caso de la ingeniería, a partir del análisis se descompone “(a) un todo en sus partes, (b) se relacionan las partes de un sistema, (c) se extraen variables, (d) se vinculan efectos y causas” (cfr. Grech, 2013).

De acuerdo con lo anterior, un ingeniero analítico se distingue porque puede identificar variables dentro de un sistema X. Por ejemplo, Watt relacionó dos variables en la explicación de la ineficiencia de la máquina de Newcomen: la evaporación del agua y el enfriamiento para la condensación. Ahora, aunque dichas variables pueden parecer variables empíricas, *simpliciter*, lo cierto es que hay una importante “carga conceptual” en el modo mismo como se conciben, *inter alia*, fenómenos físicos como la combustión.

Recuérdese que, en el tiempo en el que Watt logra la patente del motor a vapor, recientemente había tenido lugar lo que después se denominaría “la revolución química” cuyo padre, Lavoisier, anunció en 1777 su teoría de la combustión del oxígeno. Sin embargo, el propio Watt fue renuente a esta teoría, y sus compromisos teóricos y conceptuales le hicieron inclinarse hacia la teoría del flogisto, defendida por su entrañable colega Joseph Priestley (cfr. Miller, 2009, p. 6).

Argumentos sobre la especificidad del conocimiento tecnológico

Una vez planteados los argumentos para controvertir el P1, y con él la idea de que hay variedades de conocimiento —no proposicional— distintivo de la tecnología, se ofrecen cuatro argumentos que apoyan la tesis de que una parte importante del conocimiento tecnológico sí corresponde al *saber que*; y que el examen de esta porción del conocimiento tecnológico es clave para entender su estructura epistémica.

Primer argumento: el argumento de los diseños basados en la ciencia.

(P1) Si una condición necesaria para diseñar X es que este sea la aplicación de conocimiento científico (*saber que*), entonces una

parte importante del conocimiento tecnológico sí corresponde al *saber que* (este es su condición de posibilidad).

Aunque es abundante la literatura que afirma que hay tecnología sin el beneficio de la ciencia, lo cierto es que, al prestar una escrupulosa atención a la evidencia empírica disponible, hay numerosos casos que sí validan la fórmula según la cual la ciencia es una *conditio sine qua non* de la tecnología.

Para defender esta idea, por ahora se recurrirá al descubrimiento de la inducción electromagnética por Faraday. Este caso es sorprendente porque incluso los detractores de la fórmula que identifica a la tecnología con la ciencia aplicada reconocen que el desarrollo de la industria eléctrica fue posible gracias a la ciencia:

Como ya hemos visto, la mayor parte de los adelantos tecnológicos fueron consecuencia de descubrimientos empíricos llevados a cabo por hombres eminentemente prácticos: [...] ya dijimos que hasta hace relativamente poco tiempo la tecnología ha aportado más a la ciencia que la ciencia a la tecnología. La industria eléctrica *es una excepción* en el sentido de que su nacimiento y desarrollo fueron consecuencia directa de investigaciones científicas. (Derry y Williams, p. 893. Cursivas añadidas)

El descubrimiento de la inducción electromagnética fue esencial para el comienzo de la producción de corriente eléctrica alterna y el transporte de electricidad hacia finales del siglo XIX y, por tanto, para la electrificación de la economía y de la sociedad. Justamente por ello, alrededor de este impresionante descubrimiento se difundió la historia —cuya verdad está en cuestión— de que una vez el primer ministro británico visitó el laboratorio de Faraday y le preguntó para qué serviría la investigación que estaba adelantando, a lo que Faraday respondió: “No lo sé, pero apuesto a que algún día usted le pondrá un impuesto” (cfr. Farías y Cuevas, 2028, p. 46).

Considérense ahora los tres argumentos restantes:

Segundo argumento: el argumento de las teorías en ingeniería:

(P1) Si hay teorías propias en ingeniería (con sus respectivas proposiciones), entonces una parte importante del conocimiento tecnológico sí corresponde al *saber que*.

(P2) Hay teorías propias en ingeniería...

Por lo tanto, una parte importante del conocimiento tecnológico sí corresponde al *saber que*²³. Bunge ya había mencionado en su clásico escrito “Technology as Applied Science” la teoría del vuelo como una aplicación de la dinámica de fluidos (es una teoría tecnológica substantiva porque está precedida por una teoría científica) (cfr. Bunge, 1966, p. 21). Tales teorías no solo contienen proposiciones, sino genuinos conceptos propios, tales como “máquinas”, “servomotor”, “vigas”, etc.

Tercer argumento: el argumento de las leyes tecnológicas:

(P1) Si hay leyes en tecnología, entonces una parte importante del conocimiento tecnológico sí corresponde al *saber que*. Un ejemplo —aunque sea controversial— es la ley de Moore, la cual afirma que el número de transistores en un circuito integrado se duplica aproximadamente cada dos años²⁴.

(P2) Si hay leyes en tecnología (por ejemplo, la ley de Moore [informática] o la ley de Peter, en la administración).

Por lo tanto, una parte importante del conocimiento tecnológico sí corresponde al saber que.

Cuarto argumento: el argumento de los conceptos propios en tecnología:

23 Para unas teorías propias de la ingeniería como la teoría de vigas, la teoría de placas y láminas, ver Flórez, 2023.

24 Hay otras leyes, como la ley de Nielsen, de Metcalfe, Kryder, Gildey y Bell. Una ley en la ingeniería civil podría ser la ley de Coulomb, que establece la relación entre la fuerza eléctrica y la carga eléctrica. Esta ley es importante en la ingeniería civil porque se utiliza para el diseño de cimientos, estructuras de soporte y otros aspectos relacionados con la resistencia de los materiales y la estabilidad estructural en general.

(P1) Si hay conceptos propios en tecnología (por ejemplo, “máquina”, “servomotor”, “viga”), entonces una parte importante del conocimiento tecnológico sí corresponde al *saber que*.

(P2) Hay conceptos propios en tecnología.

Por lo tanto, una parte importante del conocimiento tecnológico sí corresponde al *saber que*.

Se espera que los anteriores argumentos sean lo suficientemente provocadores y contundentes para respaldar la propuesta sobre la vía de análisis que se debería tomar en la filosofía de la tecnología, a saber: explorar el dominio del contenido conceptual de la tecnología. Este dominio es el que hace posible explicar qué tipo conocimiento proviene de las ciencias, así como inspeccionar los conceptos propios en la tecnología y, con ellos, incluso los cambios semánticos profundos que tienen lugar tras las “revoluciones tecnológicas”. Este dominio luce mucho más promisorio para un análisis de la estructura epistémica de la tecnología.

La cuestión, entonces, radica en entender la estrategia o el dispositivo propio del conocimiento tecnológico. Si bien parte de unos condicionamientos históricos, la imaginación del ser humano aprende según la historia tecnológica que ha vivido la humanidad. La propuesta es que el arte de la clasificación sirve como el trasfondo para explicar que el conocimiento tecnológico aplica el conocimiento científico a la vez que goza de una especificidad propia, precisamente porque el diseño recurre al sistema de objetos previamente clasificados (entre ellos, conceptos científicos, teorías científicas, pero también elementos, máquinas o dispositivos producidos con anterioridad para solucionar problemas análogos o completamente diferentes).

Clasificación epistemológica de dos formas de conocimiento tecnológico

En esta sección se explican dos enfoques epistemológicos desde los que se ha entendido el conocimiento de la tecnología: la del diseño y la ergonomía. El primero de los dos enfoques sobre el diseño enfatiza

la imaginación (Simondon, 2007, 2017), en la medida en que pretende dar cuenta de los procesos mentales de producción de los objetos técnicos. Al segundo enfoque le interesa hacer explícita la participación de los usuarios en la producción de artefactos, por razones de su estética, ergonomía, funcionalidad y efectividad (Rabardel, 2011).

Simondon

El reciente interés de los filósofos por la tecnología ha traído consigo la recuperación de unos pensadores que habrían quedado en el olvido. Dicho en otras palabras, no quiere decir que no se supiera o no se tuviera noticia sobre unos pensadores que habían construido explicaciones sobre la tecnología, sino que, aunque se tenía conocimiento de estos, solo el renovado interés durante las tres últimas décadas ha vuelto a incluir sus aportes en la discusión filosófica. Esta idea la podemos corroborar con el caso de Ernst Kapp, primero, y de Gilbert Simondon, en segunda instancia, de cuyo pensamiento hablaremos en esta sección.

Carl Mitcham le atribuye la primera obra de filosofía de la tecnología a Ernst Kapp (1989, 1994). El libro publicado por este alemán hegeliano de izquierda se titula *Principios de una filosofía de la técnica: sobre la historia del origen de la cultura desde nuevos puntos de vista* (1877/2018), del cual Mitcham explica que es la primera reflexión filosófica detallada y novedosa para comprender la ingeniería desde la filosofía. La reedición alemana es de 2015 y aunque está disponible en Google Books desde 2006 en el original alemán de 1877 —la primera traducción al francés es de 2007, al inglés de 2018 y al español de 2020—, tuvieron que pasar más de cien años para que la academia filosófica volviera a interesarse por este tipo de obras.

De hecho, la reflexión de Kapp formula la idea de que la técnica duplica o extiende los órganos humanos o que, al decir de Mitcham, los instrumentos (y las armas) son proyecciones de estos. Esto querría decir que Kapp no solo es el primero en formular una filosofía de la tecnología, sino que es el primero en tener la idea de que los instrumentos tienen su configuración según el cuerpo humano, esto es, que son diseñados conforme a la corporalidad humana. Aunque esta sección

no presentará un estudio de Kapp, en el siguiente apartado sí estudiará esta epistemología ergonómica de la tecnología.

El conocimiento de la obra de Simondon no ha sido tan desafortunado como el de Kapp. Para empezar, acaso pasaron treinta años para que surgiera un renovado interés por sus propuestas. Con todo, durante esos años hubo algunas referencias a su pensamiento que lo mantenían dentro de discusiones filosóficas, aunque fuera un poco al margen. En 1964 Marcuse (1972) y en 1968 Baudrillard (2010) hacen un uso de la obra de Simondon, sin plantear una crítica a esta y sin suscribir toda la concepción simondoniana, lo hacen solo de manera selectiva. Mitcham, en su revisión histórica de pensadores sobre la tecnología (1994), lo incluye, aunque no como personaje principal dentro de la filosofía de la tecnología, sino como un psicólogo-fenomenólogo de la mano de Jacques Lafitte (Lafitte et al., 1932). Finalmente, y como referencia el mismo Mitcham, Gilbert Hottois es quien publica un primer estudio pormenorizado de la obra de Simondon en *Simondon y la filosofía de la “cultura técnica”* (Hottois, 1993)²⁵.

No resulta simple ni fácil organizar la filosofía de Simondon sobre la técnica para poder exponerla como enfoque epistemológico. Esta será, por tanto, una propuesta muy heterodoxa. Como nos interesa entenderlo en términos epistemológicos, se expondrá su idea de objeto técnico, su idea de producción de objetos técnicos o proceso de concretización, su idea de imaginación inventiva y lo que concibe como la diferencia de la mentalidad técnica respecto de otras formas de pensamiento.

Como se expuso arriba, la clasificación de objetos, si bien puede tener un sentido ontológico, corresponde a un primer ejercicio de clasificación epistemológica. Lo primero que se debe tener muy claro es que Simondon distingue tres niveles o tipos de objetos técnicos, de

25 Gilles Deleuze, en *Diferencia y repetición* (2002) y en *Lógica del sentido* (1994), se basa en la obra de Simondon sobre la individuación, y junto con Guattari en *El antiedipo* (1995) y en *Mil mesetas* (1997) hacen referencia a la obra sobre los objetos técnicos. Empero, en ninguno de los dos casos Deleuze ni Guattari pretenden discutir la concepción de tecnología y menos entrar en discusiones de filosofía de la tecnología.

manera que otras expresiones como herramienta, instrumento, máquina, objeto material o artefacto no se oponen, sino que se incluyen. El primero es el *elemento técnico*, el segundo tipo de objeto es el *individuo técnico* y el tercero es el *conjunto técnico*. Para decirlo de modo directo, son tres modos de existencia del objeto técnico.

El segundo aspecto que se debe tener claro es que la concepción simondoniana de la tecnología está enmarcada en su teoría de la individuación. Esto significa que la teoría de la tecnología es un apartado de la teoría de la individuación. En efecto, el objeto técnico se entiende, para Simondon, si se entiende su proceso de concretización; este proceso de concretización equivale a la individuación del objeto técnico o la génesis del objeto técnico.

La tercera cuestión simondoniana importante, que tal vez sea la más relevante dentro del planteamiento epistemológico, es precisamente su idea de imaginación inventiva como núcleo explicativo del mismo. Esto permite dar cuenta de la forma específica en que opera el pensamiento técnico.

La diferenciación del pensamiento técnico frente al pensamiento mágico, religioso, estético y filosófico es la cuarta y última claridad que permitirá entender el análisis epistemológico que Simondon hace de la tecnología. Este cuarto aspecto requiere pasar de una concepción histórica de la evolución de la técnica a una concepción tecnológica de esta. Su estrategia para lograr este cambio en la manera de exponer la diferencia entre esos tipos de pensamiento consiste en recurrir al concepto de fase, para dar cuenta de esas sinergias.

Los objetos técnicos: elemento, individuo y conjunto

El elemento técnico para Simondon (2007) corresponde a los componentes simples de las máquinas y demás artefactos tecnológicos. Tornillos, alambres, resortes, puntillas y similares corresponden a los elementos. Esto quiere decir que son casi como entidades inertes que por sí solas no hacen nada, sino que requieren o bien que las use un agente humano o que entren en correlación funcional con otros elementos en una máquina. Él recalca que el elemento sirve para medir con propiedad el grado de desarrollo técnico de una sociedad. En efecto, toda producción de elementos técnicos implica una serie de especificaciones

técnicas que, a su vez, requieren de un diseño. Las especificaciones pueden llegar a ser, por tanto, de muy alta precisión (por ejemplo, nanómetros, nanogramos, para el grosor, largor, aleaciones, ángulos, y demás), como para la producción de arandelas, tornillos, resortes o de microtransistores. De modo que poder producir un elemento con alto grado de precisión equivale a tener la infraestructura tecnológica y el capital humano capaz de ello. Así, a mayor exigencia para la producción de elementos que posteriormente se utilizarán para la construcción de individuos técnicos, mayor exigencia cognitiva y productiva tecnológica, por ende, mayor desarrollo tecnológico social.

Los instrumentos o las máquinas más complejas integradas por los elementos técnicos, Simondon los denomina *individuos técnicos*. Las máquinas modernas de la época industrial corresponden a los individuos. Estas tienen la posibilidad de tener una etapa de funcionamiento posterior a su ignición, independiente del agente humano. Pero él advierte que esta relación humano-máquina realmente se desplaza a los momentos cruciales que hacen evidente la codependencia: volver a cargar el combustible, hacer mantenimiento, apagar, ubicar en el lugar deseado. Es cierto que los individuos técnicos tienen un grado de diseño que puede ser más complejo que el de los elementos, pero su construcción o ensamble, en sentido estricto, depende de los elementos técnicos que los constituyen, razón por la cual los individuos técnicos no son baremo para saber el grado de desarrollo social; de hecho, podrían ser importados y ensamblados en maquilas.

El tercer tipo de objeto técnico engloba a los sistemas técnicos articulados en red de individuos técnicos y agentes. Es decir, se trata de conjuntos de individuos que son más complejos; las fábricas, las redes eléctricas de una ciudad o los aeropuertos o los aviones, un astillero o un transatlántico son buenos ejemplos. Estos conjuntos técnicos se conforman de individuos, pero también pueden ser productores de individuos técnicos, lo cual significa que el sistema técnico tiene un grado de autoproducción, siempre codependiente con los agentes humanos.

Ahora bien, el concepto de objeto técnico parece que solo se puede usar adecuadamente de manera interpretativa porque un avión puede ser considerado como un individuo técnico, pero también como un conjunto técnico en sí mismo. Un procesador equivaldría a un elemento,

pero es una máquina que no funciona fuera de su interacción con los demás circuitos de la placa base (tarjeta madre o *motherboard*).

Esta maleabilidad del concepto de objeto técnico en sus tres niveles puede ser el punto débil de la concepción simondoniana porque hace explícito su límite para dar cuenta de las nuevas tecnologías electrónicas, aunque él conocía bastante bien los desarrollos de los transistores de su época que, de hecho, habían sido paradigmáticos para desarrollar su teoría de los objetos técnicos. Pero, por un lado, para hoy día, este concepto no alcanza a cubrir el que, sin lugar a duda, hace parte de la cultura. A pesar de que la propuesta de Yuk Hui (2016) por precisar la idea de objetos digitales sigue en esta línea de desarrollo simondoniano. El problema radica en que el *software* no puede pensarse como una cosa física pero tampoco como una realidad mental. El *software* hace parte de la cultura técnica pero no corresponde ni al elemento ni al individuo ni al conjunto, así como tampoco es solo información o solo archivos digitales. Y, por otro lado, los objetos como la ropa, los zapatos, los muebles y similares no entrarían tampoco dentro de lo que abarca el objeto técnico de Simondon, con lo cual la dicotomía que quería salvar entre la cultura y la técnica queda desplazada porque no incluye como parte de la tecnicidad de una cultura todo aquello que no entra como aparato o artefacto para hacer algo. De ahí que resulte crucial revisar la concepción de Beth Preston sobre la tecnología.

Proceso de concretización

La individuación es el proceso para dar cuenta de la aparición de individuos en la realidad. Esta teoría de la individuación implica, para Simondon, evitar recurrir a la visión hilemorfista, que considera que cualquier individuo (cosa u objeto) se explica por una aleación entre materia y forma. De hecho, quiere dar cuenta del surgimiento de individuos físicos (como las rocas o cristales), a partir de los cuales puede explicarse la aparición de individuos biológicos (los seres vivos junto con sus ecosistemas físicos), para después pasar a explicar la individuación psíquica y, por último y a partir de esta, la individuación colectiva (Simondon, 2005, 2014). La cuestión es que los objetos técnicos tienen un proceso de individuación particular en la medida en que surgen como parte de los procesos de individuación psíquica y parte de los

procesos de individuación colectiva, a la vez que implican una especie de proceso de individuación física, pero distinta de la que había explicado en sentido estricto como tal, podríamos decir, de orden natural. Este proceso particular Simondon lo llama *proceso de concretización* y tiene inserto un proceso de diseño o de imaginación inventiva muy importante, que se expondrá en el siguiente apartado.

Simondon tiene como paradigma explicativo de la concretización del objeto técnico al ser viviente. En sentido estricto:

Se puede decir solamente que los objetos técnicos tienden hacia la concretización, mientras que los objetos naturales, tales como los seres vivientes, son concretos desde el comienzo. No hay que confundir la tendencia a la concretización con el estatuto de existencia enteramente concreta. (Simondon, 2007, p. 69)

Entonces, el proceso de concretización corresponde a la tendencia de todo objeto técnico a concretizarse cada vez más en su existencia, pero nunca alcanzará la concretización completa, en cierto sentido perfecta, como la del ser viviente.

Simondon piensa, pues, que los seres vivientes son concretos, o su estado de existencia es concreto, mientras que los objetos técnicos inician con un estado abstracto (opuesto al concreto) y tienden a la concretización. La existencia concreta de un individuo es cuando su estado implica que todas sus partes están integradas de modo tal que un mal funcionamiento o daño de una parte no implica que deje de funcionar o existir como individuo, porque las partes se pueden relevar en situaciones imprevistas de manera momentánea, a veces, permanente.

Dos ejemplos generales pueden ayudar: un automóvil que le falte una llanta no puede andar, mientras que un animal no humano que tenga un daño en una extremidad, puede compensar con las otras, aunque cojee, y seguir andando. Cuando la fuente de poder de un ordenador se funde por sobrecalentamiento este simplemente deja de funcionar completamente, mientras que un ser viviente sigue existiendo (funcionando) aunque tenga problemas de alimentación (hasta tres días sin agua, una semana sin alimento) o de metabolismo.

Pero esto lleva a pensar la producción misma de los objetos técnicos. En principio y según lo establecido, cuando un objeto técnico se inventa y existe por vez primera, se encuentra en estado abstracto porque sus partes son independientes unas de otras, en virtud de ese diseño inicial. El proceso de concretización de un objeto técnico equivale a pasar de un estado abstracto a un estado más concreto. En este sentido, el proceso de concretización de un objeto técnico quiere decir que el diseño inicial cambia y las partes se van ajustando para un mejor funcionamiento del objeto técnico. Dicho en otras palabras, el proceso de concretización, que implica un diseño abstracto al inicio, trae consigo momentos posteriores de re-diseño por el cual los objetos técnicos se perfeccionan, mejoran y se vuelven más eficientes.

Esta idea de concretización es relevante porque muestra la centralidad del diseño para la existencia de los objetos técnicos. Pero ¿cómo es el diseño de los objetos técnicos?

Imaginación inventiva

La creación de un objeto técnico comienza por un problema al cual se le quiere dar solución técnica. Sin embargo, la misma determinación del problema, como se dice en filosofía, o el hecho de que los problemas se los hacen o se los plantean las personas o los mismos grupos humanos, antecede a la invención de una solución técnica. Se debe tener claro que, aún si se plantea un problema por parte de un individuo, esto no quiere decir que su solución sea técnica. Considérese el siguiente caso:

Un constructor tiene veinte obreros con palas haciendo una excavación profunda. Para rentabilizar su empresa, llama un asesor para que le oriente sobre cómo mejorarla. El asesor le dice que es muy fácil si en vez de tener veinte obreros alquila una retroexcavadora y contrata solo al conductor de esta. El constructor le explica que dejaría sin ingresos a veinte familias, razón por la cual debe asesorarlo de otra manera. A partir de esto, el asesor le explica que es todavía más fácil porque debe entonces contratar cien obreros y cambiarle las palas por cucharas.

Sin duda hay un problema: rentabilizar la empresa de la excavación. Pero el constructor no quiere una solución técnica: alquilar la retroexcavadora. No se sabe si el constructor aceptó la segunda propuesta

del asesor, pero queda claro que no todos los problemas que se pueden solucionar técnicamente, de hecho, se vayan a solucionar de este modo. En este marco, importa más partir de la idea de que el problema tiene que ser ideado por los seres humanos y que se debe aspirar a tener una solución técnica, y no de la idea de que los problemas están ahí dados, de manera evidente para todas las personas y que siempre se aspira a una solución técnica. Este aspecto de la intención en la solución de problemas se examinará en las siguientes secciones.

Establecido entonces que los problemas se construyen, se puede pasar a entender que la imaginación inventiva, propia del pensamiento técnico, tiene como punto de partida el problema solucionado. En efecto, Simondon considera que aquello que hace posible el surgimiento (la condición de posibilidad) de un objeto técnico es la solución del problema (2007, p. 76). Las reflexiones sobre la imaginación inventiva tienen gran importancia epistemológica porque explican cómo se producen los inventos y, por lo tanto, la producción de nuevas soluciones técnicas.

Una vez establecido el problema a ser solucionado, la imaginación inventiva produce soluciones y, de manera inversa, empieza a buscar y proponer esquemas mentales que satisfagan la solución imaginada. Este paso consiste en recurrir al fondo o, lo que es lo mismo, hacer uso de un sistema de esquemas ya pre-existentes a los cuales la imaginación recurre para articular la solución posible:

Lo que se le ha escapado a la atención de los psicólogos en el análisis de la imaginación inventiva *no son* los esquemas o las formas, o las operaciones, que son elementos espontáneamente sobresalientes y en relieve, sino el fondo dinámico sobre el cual se enfrentan dichos esquemas, se combinan, y del cual participan. [...] El fondo es aquello que se oculta de los dinamismos; es aquello que hace existir el sistema de formas; las formas no participan de las formas sino del fondo, que es el sistema de todas las formas, o más bien el reservorio común de las tendencias de las formas, antes incluso de que éstas existan a título separado y se constituyan en sistema explícito. (Simondon, 2007, p. 79)

La aclaración de Simondon es crucial porque explica que, a la hora de inventar, siempre se asume un reservorio de pequeños esquemas de los que se echa mano para proponer una solución técnica. En este punto el conocimiento tecnológico recupera nuevamente su estatuto de *techné* o arte porque, al igual que una paleta de colores, el inventor tiene a su disposición una paleta de elementos, diseños, dispositivos o materiales que le permiten concretar la solución técnica que tiene en mente.

En efecto, la obra de Walter Vincenti *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History* (1993) explica con bastante detalle el caso del desarrollo tecnológico de la aviación a partir de diferentes tipos de clasificación de objetos o componentes del avión. Este es un ejemplo de compendio de catálogos de partes de avión que han sido previamente diseñadas y clasificadas o catalogadas para pensar nuevos objetos técnicos o aviones.

Pero se requiere entrar en un enfoque que haga aún más explícito que Simondon la corporalidad humana, de manera que se entienda que el diseño de cualquier solución técnica se piensa siempre en función de la necesidad humana que se pretende satisfacer. No es que no haya tecnologías que apunten a mejorar el bienestar de otros animales no humanos, sino que cualquier solución técnica implica un diseño humano en cuatro etapas: la construcción del esquema o planificación, la construcción de los prototipos y sus pruebas (modelamiento), la producción a mediana y gran escala según el público al que esté destinado y el uso propiamente dicho a nivel particular de la solución técnica. Estas etapas, además, se retroalimentan, de manera que siempre puede ser necesario un regreso a una etapa anterior, así como puede haber una atención a etapas posteriores²⁶.

26 Por cuanto hay múltiples sistemas de industrias que se cruzan constantemente, se debe tener presente que los intereses económicos, sociales y políticos influyen también en el diseño, la producción y renovación de las soluciones tecnológicas. En efecto, la actualización constante de dispositivos móviles computacionales o de ropa, productos alimenticios, cada vez más completos, más sanos, más adecuados, potentes o, sencillamente, mejores, dejan claro que la retroalimentación en el diseño es más que obvia, necesaria.

Conocimiento y uso de herramientas

Cuando se piensa en el uso de herramientas, esto es, objetos externos de los que nos servimos para conseguir algo, se nos vienen a la mente una gran variedad de casos, desde el mundano empleo de un martillo para clavar un clavo hasta la manipulación de artefactos complejos y sofisticados como, por ejemplo, la utilización de un computador para realizar complejos cálculos estadísticos. Una inquietud que puede surgir al respecto, y que de hecho la filosofía de la tecnología ha considerado, está relacionada con las condiciones epistemológicas que suponen este tipo de casos. Por ejemplo, ¿qué debe saber una persona para manipular adecuadamente un martillo si lo que quiere es clavar un clavo evitando doblarlo o, peor aún, dañar la madera sobre la cual espera incrustarlo? ¿Este tipo de saber sería diferente al que debe poseer una persona que usa lentes todos los días? O más bien ¿no aprenderíamos a manipular herramientas de manera mecánica y repetitiva sin que para ello medie algún tipo de conocimiento? Como veremos, este tipo de preguntas reflejan ciertas actitudes filosóficas hacia el aspecto epistemológico del uso de herramientas.

Una primera actitud consiste en negar, o al menos minimizar, que el uso de herramientas suponga, en general, conocimiento. Esto puede sonar extraño de entrada, pero si se piensa bien la tesis en cuestión puede cobrar cierta plausibilidad. Pensemos, como lo hizo el filósofo griego Aristóteles (*Ética a Nicomaco* y *Metafísica*), en algunos oficios artesanales como la carpintería o la zapatería. En ellos, se podrían distinguir dos fases claramente diferenciadas. Por una parte, encontramos la etapa de diseño que abarca tanto la prefiguración del resultado final que se quiere alcanzar, digamos un zapato o un mueble, como la especificación de los pasos sucesivos que han de darse para su construcción. La segunda etapa corresponde a la construcción propiamente dicha del objeto planeado. En esta etapa es cuando alguien, por lo general un artesano, toma las herramientas adecuadas para trabajar ciertos materiales; su trabajo, en caso de seguir los pasos especificados con anterioridad, deberá dar como resultado un objeto que se ajusta al diseño del objeto prefigurado. Sin duda, la primera etapa concentraría la mayor parte del esfuerzo intelectual, mientras que la segunda

sería una mera cuestión de ejecutar paso a paso un conjunto específico de instrucciones. En consecuencia, cuando el artesano, o alguien más, emplea herramientas para construir un mueble o hacer un zapato, no necesita pensar demasiado en lo que hace, pues su única responsabilidad es atenerse al plan previamente establecido. Lo realmente interesante en términos epistémicos sucede antes.

Para reforzar el anterior argumento, piénsese en otro ejemplo. Cuando se aprende a montar bicicleta, los mayores esfuerzos parecen estar concentrados en las etapas iniciales de dicho aprendizaje. En ellas, el aprendiz se concentra en observar lo que hace o evita hacer el instructor cuando muestra cómo montar bicicleta. Se presta atención a sus explicaciones, sus correcciones y, en general, a todos sus consejos sobre la tarea en cuestión. Las primeras veces que se intenta por sí mismo, también se presta mucha atención a lo que se hace, buscando mantener de forma consciente el equilibrio y evitando cometer los errores sobre los cuales se había advertido. Pero después de un tiempo, y, por supuesto, tras varias caídas, el aprendiz se da cuenta de que puede avanzar sin grandes contratiempos, y que para ello no se hace necesario concentrarse demasiado en lo que hace. A partir de este momento, él puede empezar a disfrutar las salidas en bicicleta, ya que toda su atención queda libre para dirigirse hacia el paisaje o hacia sus acompañantes. En últimas, cuando se usa una bicicleta como medio de transporte, no se piensa mucho en las especificidades de la tarea misma.

A partir del ejemplo anterior y del argumento aristotélico sobre los oficios artesanales, podríamos llegar a la conclusión de que, en efecto, cuando se manipulan herramientas, el esfuerzo cognitivo que se hace es mucho menor al que debe hacerse cuando se enfrenta a tareas como resolver una integral o escribir un ensayo para la asignatura de Filosofía. En este tipo de actividades, más “intelectuales”, toda la atención está dirigida a los elementos que conforman la estructura de la tarea y la persona se ve obligada a analizar cuidadosamente cada paso que da. En contraposición, las herramientas, cuando son manipuladas con destreza, parecen, como diría Merleau-Ponty, transparentes en el sentido de que se deja de notarlas y ella se concentra en otras cosas que considera más importantes. Durante su uso, entonces, no sería necesario algún análisis o consideración de sus acciones sobre ellas y,

por lo tanto, no tendría que suponer que el conocimiento sobre la tarea o sobre la herramienta particular se pone en juego y se moviliza.

La anterior visión, por supuesto, no ha estado exenta de críticas. La primera de ellas es que parece ofrecer un modelo demasiado centralizado del trabajo de producción de nuevos objetos, en el que tan pronto se definen, en las primeras etapas, los pasos a seguir, nada ni nadie puede modificarlos. Es como si quien construyese un nuevo objeto no tuviera la capacidad para modificar los planos iniciales. Y ello es así porque esa persona no sabría cómo hacerlo, pues solo se limita, durante el proceso mismo de construcción, a seguir ciegamente las instrucciones que le han sido asignadas. Quizás por su lejanía con el trabajo artesanal, Aristóteles no consideró que un buen artesano es aquel que, precisamente, posee la habilidad para ajustar sobre la marcha los planes iniciales, quizás porque le fue imposible conseguir alguno de los materiales a buen precio, y aun así construir un producto bello y funcional.

Otra crítica a esta visión tiene que ver con la supuesta falta de atención con que se realizan tareas manuales. Esta tesis, defendida tanto en filosofía como en psicología, contradice lo dicho en los reportes de primera persona que nos ofrecen artesanos y atletas. En el periodismo deportivo es una práctica común preguntar a los atletas de élite qué experimentaron en el transcurso de los momentos más decisivos de sus carreras deportivas. Egan Bernal, ciclista colombiano que ganó el Tour de Francia en 2019, reporta, en el transcurso de una entrevista para un medio extranjero, que durante la carrera se sintió mejor en los Alpes que en los Pirineos porque el *feeling* que tuvo sobre la bicicleta fue mejor en las etapas que corresponden a la primera de dichas cadenas montañosas. Añade que tal sensación le permitió sentirse confiado pero analítico, lo cual contribuyó, a la postre, en su victoria en el Tour. La experiencia de Egan corrobora algo que omiten algunos análisis filosóficos: practicar con destreza un deporte que involucra la manipulación de una herramienta no supone su “desaparición”, sino, por el contrario, una sensación de confort y dominio sobre ella. Es como si la bicicleta de Egan Bernal se hubiese convertido en una extensión de su cuerpo, similar, quizás, a una prótesis a la cual

alguien logra acoplarse sin problemas y que prácticamente no ejerce ningún tipo de resistencia.

Precisamente, la idea de que algunas herramientas pueden considerarse como extensiones corporales se ha convertido en el eje de una segunda aproximación al papel del conocimiento en su utilización, la cual tiene como punto de partida el reconocimiento de la importancia que poseen las herramientas en la supervivencia y, en general, para lograr la adaptación a diferentes entornos naturales. Imagínese cuán difícil sería sobrevivir un día en el bosque si no se contase con un hacha, por ejemplo. Ahora bien, la misma hacha es útil porque en cierta forma incorpora, en su diseño y características materiales, el conocimiento de la realidad; ella posee la clase adecuada de propiedades para cortar eficientemente superficies duras y de cierta densidad. En este sentido, se dice que las herramientas expanden las capacidades físicas: gracias al hacha ahora se pueden hacer cosas que sin ella sería imposible conseguir. Siguiendo esta misma línea argumentativa, se afirma que otros instrumentos, como los lentes, expanden las capacidades cognitivas: permiten ver más lejos y mejor. Todo ello en virtud de que incorporan un conocimiento de los principios de la óptica.

La ampliación de las capacidades humanas gracias a la tecnología va desde casos obvios como la visión aumentada al usar un telescopio o una locomoción más veloz gracias a los vehículos, hasta casos más sutiles como una mejor precisión en los cálculos al usar lápiz y papel. Naturalmente, al ampliar las capacidades, las herramientas también las transforman (Chakrabarty, 2018). Esta transformación se produce de muchas maneras. Considérese brevemente la visión aumentada a través de las tecnologías de la visión: cosas como gafas, telescopios, microscopios y similares. En todos estos casos, y en distintos grados, la visión se amplía. Las gafas graduadas, por ejemplo, permiten ver ligeramente más grande y con mayor nitidez de lo que se vería con los ojos desnudos (Ihde, 1990, p. 72). Aquí, el instrumento empleado introduce correcciones en nuestra visión, que de otro modo sería imperfecta. En el caso de un telescopio, los cambios introducidos en la visión son más radicales y van más allá de las meras correcciones; lo que alguien ve a través de un telescopio son objetos y horizontes antes inaccesibles a sus ojos desnudos, bien sea porque son demasiado

distantes o bien porque son imperceptibles a simple vista (como los objetos celestes que emiten rayos en infrarrojo).

Si bien es cierto que la anterior perspectiva constituye un avance con respecto a la visión que minimiza el papel del conocimiento en el uso de herramientas, pareciera que aclara muy poco en qué consiste dicho rol. Más bien, lo que sí permite aclarar es el aporte del uso de herramientas al conocimiento humano: las gafas nos permiten ver más y mejor, el lápiz y el papel permiten hacer operaciones matemáticas más complejas, etc. También reconoce el importante papel que tienen las herramientas en la interacción cotidiana con el mundo y la manera en que ellas nos han ayudado a convertirlos en lo que son hoy en día. Sin embargo, cuando se trata de decir algo relativo al tipo de conocimiento involucrado en el uso de herramientas, esta perspectiva permanece en silencio. Esto es preocupante, especialmente, si se tiene en cuenta que al preguntar por lo que significa vivir en una cultura como la actual, tan llena de artefactos tecnológicos y herramientas de diversa índole, pasa por alto discutir el papel que juegan, en su uso y producción, las habilidades manuales y técnicas, así como el conocimiento práctico y teórico. Por ello, parece que la filosofía de la tecnología no estará completa si no se abordan estas cuestiones epistemológicas de manera directa.

Pero, antes de intentar decir algo sobre las cuestiones relacionadas con la epistemología del uso de herramientas, se revisará, aunque sea brevemente, por qué se ha dejado de considerar dichas cuestiones en la filosofía de la tecnología más contemporánea, representada por la segunda de las perspectivas hasta ahora examinadas. Parece que este descuido puede explicarse por el tipo de herramientas y por el tipo de uso que se suelen tener en cuenta en la filosofía de la tecnología. Así, autores que se identifican a sí mismos como postfenomenólogos, se han centrado en gran medida en las tecnologías relacionadas con la visión. Sin embargo, estas tecnologías son una especie de excepción a la regla, ya que se pueden utilizar de forma eficiente sin necesidad de entender mucho sobre su funcionamiento. Y lo que es más importante, este tipo de instrumentos aumentan y mejoran nuestra visión sin tratar de modificar el objeto u objetos que se ven. Así, por ejemplo, en el uso de este tipo de tecnologías, no hay necesidad obvia de que el usuario

tenga alguna representación de las relaciones entre fenómenos visuales como la emisión, absorción o reflexión de la radiación electromagnética de objetos remotos. Sin embargo, en el caso de muchos otros tipos de herramientas, y tal vez incluso de la mayoría, el usuario aprende a manipularlas de forma eficaz para crear cambios en un objetivo. Aquí es donde las preguntas de corte epistemológico toman más fuerza.

En primer lugar, piénsese que el uso diestro de herramientas tiene dos características principales. Por una parte, se puede encontrar lo que se conoce como generalidad, esto es, la capacidad de un agente para utilizar la misma herramienta para completar una amplia variedad de tareas en diferentes contextos de uso (Greif y Needham, 2011). Esta capacidad se exhibe cuando se reutilizan objetos cotidianos para que cumplan funciones diferentes a las previstas en su diseño original. Por ejemplo, cuando se reutilizan botellas, después de que el líquido contenido en ellas ha sido consumido por completo, para funciones tan disímiles como lo son servir de ornamentación o de soporte para otro tipo de objetos. El segundo rasgo de cualquier uso hábil de una herramienta, la sistematicidad, tiene que ver con la proporcionalidad que un agente es capaz de establecer entre sus acciones y las características que posee la herramienta que manipula. Así, el usuario de un martillo es capaz de modular, por ejemplo, la fuerza que aplica a la herramienta de tal forma que evita romperla junto con el clavo que intenta clavar. Claro, entre más resistentes perciba que son el martillo y el clavo más fuerza aplicará para completar la tarea.

Ambas características se han visto, en la bibliografía científica, como indicadores de destreza y flexibilidad en el uso de herramientas. Por ejemplo, un tema recurrente en psicología en las últimas décadas es la capacidad de introducir correcciones, grandes o pequeñas, en la acción cuando se manipulan herramientas. Los investigadores han descubierto, a través de diversos tipos de análisis de la acción (Mangalam et al., 2018; Liu et al., 2016), que los usuarios de herramientas, tanto humanos como no humanos, son capaces de corregir los movimientos para optimizar el esfuerzo, lo cual es bastante importante desde una perspectiva evolutiva (Newell, 1986). Por ejemplo, los monos capuchinos salvajes del Brasil rompen nueces golpeándolas con piedras

mientras modulan los parámetros cinemáticos de sus golpes y los ajustan a las condiciones específicas tanto de los martillos de piedra como de las nueces mismas; al hacerlo, evitan aplastar el núcleo de la nuez o malgastar energía dando más golpes de los debidos (Mangalam et al., 2018; Liu et al., 2016). Estudios en humanos muestran que los humanos desarrollan la misma capacidad de corrección de la acción a partir de los tres años (Smitsman y Cox, 2008).

Los estudios psicológicos también han prestado mucha atención al grado en que los usuarios de herramientas pueden superar el denominado efecto de fijación funcional. La fijación funcional es el efecto de considerar una herramienta exclusivamente para un propósito específico, lo que hace que los usuarios se sientan menos inclinados a utilizarla fuera de su contexto habitual (German y Barrett, 2005; Ebel et al., 2020). Curiosamente, estos estudios se han llevado a cabo tanto en humanos como en otros primates (monos capuchinos y grandes simios) (Ebel et al., 2020; Evans y Westergaard, 2006; Sousa et al., 2009). Los experimentadores han demostrado que los usuarios de herramientas pueden ser sensibles a diversas variables, como la experiencia previa con herramientas (Ebel et al., 2020) y la información sobre las funciones previstas en su diseño original (German y Defeyter, 2000; German y Barrett, 2005).

Pero ¿cuál es la base epistémica y cognitiva de estas dos características? Lo primero que hay que decir es que ambos rasgos parecen depender de la capacidad para detectar las relaciones funcionales que existen entre las acciones, las herramientas y un tercer factor aún no considerado, a saber, los objetos que buscamos modificar. Este sería el objetivo de cualquier tarea, ya que los usuarios de las herramientas pretenden modificarlos de alguna forma. Al martillar, el objetivo de esta clase de tarea es un objeto diferente al martillo mismo, por lo general un clavo, que busca ser incrustado en una superficie. La herramienta es el medio a través del cual el usuario afecta al objetivo de la tarea. El usuario es el agente que determina el objetivo de la tarea y utiliza la herramienta para cualquier fin que se proponga. El uso de una herramienta para cumplir cierta tarea requiere que el usuario intervenga sobre la propia herramienta modificando algunas de sus propiedades (por ejemplo,

manipulando su dirección de movimiento, impulso, energía cinética, etc.), aplicando para ello fuerza, en cualquiera de sus variadas manifestaciones físicas. La herramienta actuará sobre el objetivo de la tarea y lo modificará. Así pues, el objetivo, la herramienta y el usuario están relacionados causalmente a través de las intervenciones intencionadas del usuario.

Las representaciones precisas de estas relaciones proporcionan al usuario información diferenciadora que facilita la manipulación eficaz de las herramientas. En primer lugar, el usuario necesita una representación cognitiva de cómo se relacionan causalmente los valores o propiedades de los objetivos y de las herramientas empleadas. Por ejemplo, alguien que use un martillo para martillar debe saber que el ángulo de entrada del clavo es importante a la hora de determinar si la madera en la que se le pretende incrustar se va a mantener unida o se partirá. Del mismo modo, dado que la cantidad de fricción que actúa sobre el clavo aumentará en función de la densidad de la madera, una madera más densa requerirá que el usuario aplique una mayor fuerza para transferir aún más energía al clavo. En segundo lugar, un usuario no solo necesita reconocer qué propiedades de las herramientas y de los objetivos son causalmente relevantes, sino también debe saber cómo ajustar sus movimientos (acciones motrices) en función de las propiedades de las primeras para producir con fiabilidad los efectos deseados sobre los segundos. Para ello, el usuario debe aprender a modificar sus acciones de la mejor manera posible para provocar un cambio adecuado en las propiedades tanto de las herramientas como de los objetivos. Para transferir la máxima energía del martillo al clavo, lo ideal es que el usuario de la herramienta golpee el clavo justo en la cabeza, ya que de lo contrario se perdería parte de la energía. Si no se utiliza la fuerza suficiente, el clavo no se clavará en la tabla como es debido, mientras que, si se golpea con demasiada fuerza, se doblará o dañará la tabla. Con cada golpe del martillo, el usuario puede modificar el ángulo de impacto, el impulso y la aceleración del martillo para clavar eficazmente el clavo en la madera.

Para la representación de todas estas relaciones, el conocimiento causal del mundo parece ser muy importante. Cuando se habla

de este tipo de conocimiento se hace referencia a la representación de relaciones entre dos o más eventos en los que al menos uno de ellos produce a los demás. Y, justamente, hemos dicho que aprender a manipular una herramienta pasa por reconocer lo que se puede hacer con ella, de sus posibilidades en la red de relaciones causales entre las acciones del usuario, la herramienta y los objetivos potenciales de una tarea.

La información diferencial que aportan las representaciones de las relaciones causales puede ser utilizada por los agentes para ajustar sus acciones a las diferentes propiedades tanto de las herramientas como de los objetivos. Así, si se aprende que una cierta intervención en la herramienta produce los efectos deseados sobre el objetivo de la tarea, entonces se puede usar esa misma información más adelante para modular las acciones y conseguir otros efectos. Por ejemplo, si se quisiera martillar más rápido, pero sin doblar el clavo, se sabe que se debe incrementar levemente la fuerza que se ha aplicado antes sobre el martillo empleado. Aquí es importante aclarar que es probable que el usuario no necesite una representación precisa de todos los detalles de la interacción física. La estructura de la tarea restringe el grado de detalle que requieren las propias representaciones. Los coeficientes de fricción de la madera o algo parecido a la tercera ley de Newton —es decir, que cuando el martillo ejerce una fuerza sobre el clavo, el clavo ejerce una fuerza igual y opuesta sobre el martillo— podrían ser ignorados o tergiversados. Lo importante es saber que las intervenciones, las acciones de los seres humanos sobre las herramientas, introducen ciertos cambios en los objetos que se pretenden modificar. Una vez adquirida esta clase de conocimiento, están en capacidad de emplear una herramienta para completar una gran variedad de tareas con ella.

Preguntas

1. Desde la perspectiva actual, se pueden detectar varios vacíos en esta concepción griega de técnica, pero también varias claves

importantes para poder entender la especificidad del conocimiento tecnológico: ¿cuáles?

2. Ofrezca un caso, distinto de los que se ofrece en el segundo acápite, a favor de la ecuación que identifica la tecnología como ciencia aplicada.
3. Argumente por qué la biotecnología también constituye un caso a favor de la tecnología como ciencia aplicada.
4. Indique otro artefacto (distinto del microscopio electrónico), diseñado en la nanotecnología, que sirva como respaldo de la ecuación de la tecnología como ciencia aplicada.
5. ¿Por qué la invención de la radio respalda la tesis de que la tecnología es ciencia aplicada?
6. En el acápite se argumenta que el caso paradigmático de la máquina de vapor no logra rebatir felizmente la ecuación que identifica la tecnología con la ciencia aplicada. ¿Cuáles son las premisas que se ofrecen en su defensa?

Para saber más

Los siguientes textos ofrecen una entrada sólida a la epistemología de la tecnología y al problema de la especificidad del conocimiento tecnológico (incluyendo su relación con la ciencia, el papel del saber como, el conocimiento tácito/visual y las formas de representación y diseño):

- Baird, D. (2004). *Thing Knowledge: A Philosophy of Scientific Instruments*. University of California Press.
- Bunge, M. (1966). Technology as Applied Science. *Technology and Culture*, 7(3), 329-347. <https://doi.org/10.2307/3101932>

- De Vries, M., & Meijers, A. (2013). Technological Knowledge. En J. K. B. Olsen, S. A. Pedersen & V. F. Hendricks (Eds.), *A Companion to the Philosophy of Technology* (pp. 70-74). Wiley-Blackwell.
- Fantl, J. (2017). *Knowledge How*. Stanford Encyclopedia of Philosophy.
- Ferguson, E. S. (1992). *Engineering and the Mind's Eye*. MIT Press.
- Niiniluoto, I. (1997). Science versus technology: difference or identity? *Arbor*, 157(620), 397-411.
- Polanyi, M. (1967). *The Tacit Dimension*. Anchor Books.
- Ryle, G. (1949/2009). *The Concept of Mind*. Routledge.
- Vermaas, P., Kroes, P., Van De Poel, I., Franssen, M., & Houkes, W. (2011). *A Philosophy of Technology: From Technical Artefacts to Sociotechnical Systems*. Springer. <https://doi.org/10.1007/978-3-031-79971-6>
- Vincenti, W. G. (1993). *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*. Johns Hopkins University Press.

Capítulo 3

Ética de las tecnologías

JUAN CARLOS MORENO ORTIZ

DIANA MARÍA MUÑOZ GONZÁLEZ

Introducción

La ética de las tecnologías se ha vuelto un tema crucial, en un mundo mediado por las tecnologías en todos los ámbitos. Las preocupaciones éticas sobre las tecnologías son hoy cuestiones cotidianas, tanto en la vida ordinaria de los individuos como en los debates públicos.

Tradicionalmente, muchos de los planteamientos éticos sobre las tecnologías se han derivado de reflexiones filosóficas amplias, de las que se infiere, de manera explícita o implícita, algún tipo de posición ética sobre el tema. Sin embargo, en las décadas recientes se ha dado un desarrollo autónomo muy significativo del campo de la ética de las tecnologías y se han producido algunos importantes giros o transformaciones teóricas en este campo que han replanteado los puntos de vista tradicionales. Por ejemplo, históricamente hemos centrado el análisis ético en los sujetos sin conceder ninguna relevancia moral a los objetos, y hemos orientado las preocupaciones éticas en especial hacia la evaluación y el control de las tecnologías. Sin embargo, en la

actualidad se ha considerado que los objetos tecnológicos concretos son relevantes en términos morales, o cuentan de alguna forma en el análisis moral, en tanto tienen capacidad de agencia. Así mismo, se ha pretendido incluir en los enfoques éticos la necesidad de orientar moralmente las tecnologías en el terreno y anticipar sus riesgos, además de evaluar y controlar.

En este capítulo se presentan las distintas perspectivas éticas sobre las tecnologías como un marco teórico introductorio a este campo. Al respecto, se examinan sus posibilidades y límites, ventajas y desventajas, y se contrastan las orientaciones tradicionales con las orientaciones recientes.

Contexto histórico y desarrollo de la ética de las tecnologías

La ética de la tecnología como un campo de estudio independiente es relativamente reciente y se encuentra dentro del ámbito más amplio de la filosofía de la tecnología. Desde los albores del desarrollo de la tecnología moderna, comenzaron a surgir inquietudes acerca de las implicaciones de los avances tecnológicos, lo que llevó a la formulación de diversas posturas morales. Sin embargo, la ética de la tecnología, en sentido estricto, solo se consolidó más adelante, impulsada por el surgimiento de la filosofía de la tecnología como disciplina académica en los primeros años del siglo xx.

Un ejemplo temprano de preocupación ética en relación con la tecnología se encuentra en el movimiento ludita del siglo xix, que cuestionaba las implicaciones sociales de la Revolución Industrial y se oponía a los cambios que esta generaba. O en Karl Marx, quien vinculó algunos fenómenos —como la lucha de clases y la plusvalía— con los avances tecnológicos modernos. No obstante, fue solo a inicios del siglo xx que autores como Mumford, Ellul y Heidegger, entre otros, sentaron las bases de lo que hoy conocemos como ética de la técnica y de las tecnologías.

Sin embargo, después de aproximadamente cien años de trayectoria filosófica, y a partir de la gran explosión de desarrollo tecnológico

que hemos experimentado en las últimas cinco décadas, este campo ha adquirido una gran relevancia y se ha ampliado de manera considerable. Algunos de los principales textos generales y autores actuales en este campo son Robson y Tsou (2023), Mitcham (1989), Mitcham y Waelbers (2009), Verbeek (2005, 2011), Tavani (2011), Van de Poel y Royakkers (2023), Hansson (2017), Peterson (2017), Jasanoff (2016), Berg et al. (2009), Echeverría (2001), Broncano (2000, 2006, 2012), Ibarra y Olivé (2003). En este texto, no estudiaremos a cada uno de estos autores y sus ideas en detalle, sino que analizaremos de modo general los diferentes enfoques éticos que se usan comúnmente cuando hablamos de tecnología. Las perspectivas tradicionales que se van a abordar se han sintetizado en los siguientes tres grupos:

1. Las perspectivas críticas de la cultura tecnológica moderna como conjunto, vinculadas a la filosofía de la tecnología calificada por Mitcham (1989) como humanista.
2. Las perspectivas normativas, expresadas en los múltiples códigos de ética relacionadas con los distintos desarrollos tecnológicos, en especial, las perspectivas que han asumido como fundamento el enfoque normativo de la bioética, como la contenida en el texto de Peterson (2017).
3. Las perspectivas utilitaristas, pragmáticas o procedimentales que se han basado en un análisis de los costos y beneficios de los desarrollos tecnológicos, o que promueven de manera procedimental un control moral sobre las tecnologías desde el fortalecimiento de las condiciones de participación democráticas, como las formuladas en los textos de Durbin (2003) y Keulartz et al. (2004).

Junto a estas perspectivas tradicionales, se mostrarán otras actuales que buscan ampliar el análisis moral de las tecnologías, en algunos aspectos no considerados por las perspectivas tradicionales, relacionados con los propósitos y funciones de los análisis morales de las tecnologías, los posibles extremos que pueden acusar los análisis morales, las

formas como las perspectivas asumen o no el sentido valorativo social de las tecnologías, la forma esencialista o plural de comprensión de la instrumentalización tecnológica, así como la aceptación de la relevancia moral de los objetos tecnológicos. Las perspectivas tradicionales y actuales difieren sustancialmente en relación con estos aspectos.

Las condiciones y características que definen las éticas de las tecnologías

Las éticas de las tecnologías pueden caracterizarse de varias maneras, según algunos aspectos, con base en sus propósitos o funciones, sus orientaciones, los tipos de instrumentalización que analizan, los tipos de agenciamiento que estudian y sus presupuestos ontológicos. En la tabla 5 se sintetizan tales características.

Tabla 5. Características de las éticas de las tecnologías

Según factor	Característica
Según sus propósitos o funciones	Interpretativa, evaluativa, prescriptiva, regulativa, orientadora, constructiva y anticipatoria
Según sus orientaciones	Voluntarista, instrumentalista, determinista, sustantivista, crítica, situada, contextual, etc.
Según formas de instrumentalización	(alienante) Instrumentalización primaria, (emancipatoria) instrumentalización secundaria
Según tipos de agenciamiento y presupuestos ontológicos	Instrumental, determinista, dualista, relacional, etc.

Fuente: elaboración propia.

Propósitos y funciones de las éticas de las tecnologías

Podemos identificar distintos propósitos y funciones de los análisis morales de las tecnologías, de acuerdo con lo que en cada contexto se requiere, y también según con las fortalezas y los énfasis de cada perspectiva ética.

Debido a la complejidad de nuestras interacciones con las tecnologías y con los papeles que cumplen en nuestras vidas, los propósitos y funciones son diversos; por tal razón, no tiene sentido reducir el análisis moral solo a uno o unos pocos de ellos. Así mismo, ninguna perspectiva ética logra abarcar todos los propósitos y funciones. Por lo tanto, para una orientación ética que tenga en cuenta todos los aspectos, necesitamos de los aportes de todas las perspectivas.

Las principales funciones de la ética de la tecnología son la interpretativa, la evaluativa, la prescriptiva, la regulativa, la orientadora, la constructiva y la anticipatoria. Y en correspondencia, los principales propósitos son advertir, evaluar, evitar, controlar, regular, orientar, co-construir y anticipar. Una muy buena forma de comprender las posibilidades y los límites de las distintas perspectivas éticas puede ser desde la identificación de las funciones y propósitos que tiene cada perspectiva. En consecuencia, antes de presentar las principales perspectivas éticas es necesario explicar algunas características de las funciones y propósitos señalados.

Algunos análisis éticos de las tecnologías, realizados especialmente desde una mirada hermenéutica, tienen como propósito comprender el significado y el devenir de los desarrollos técnicos o tecnológicos en general o como conjunto. Esta interpretación puede ser muy útil para advertir los posibles riesgos y problemas de las tecnologías, muchos de ellos no evidentes para quienes las desarrollan y usan, pues solo se develan a través del análisis crítico realizado desde perspectivas amplias y profundas de los sentidos y los fines de esos desarrollos y usos dentro de un contexto amplio. Este tipo de análisis es muy valioso culturalmente, cumple una buena función crítica y prende las alarmas frente a posibles problemas, pero suele ser muy general y poco eficaz para ofrecer controles, regulaciones y evaluaciones.

En contraste, una perspectiva normativa tiene como función y propósito ofrecer normas en las que se establecen o prescriben los límites de los desarrollos y los usos considerados como no correctos y dañinos, y plantear los deberes desde los que se pretenden regular los desarrollos y usos correctos y positivos. Estos límites y deberes son los parámetros más concretos para establecer evaluaciones y controles en situaciones específicas. Sin embargo, es imposible advertir, regular y

controlar todos los cursos posibles de interacción con las tecnologías y sus problemas. Las ventajas que este enfoque normativo tiene para evaluar y controlar las tecnologías se contraponen a la miopía para contextualizar y comprender la incertidumbre y diversidad de interacciones con las tecnologías, y para interpretar sus fines y sentidos.

Las dos perspectivas anteriores son fuertes en los propósitos y funciones señalados, pero son débiles en los siguientes dos propósitos y funciones: orientar o construir y anticipar. Debido a su mirada general de lo tecnológico de las perspectivas tanto humanistas como las normativas, ellas tienen poca capacidad para analizar cada desarrollo y uso particular y concreto, y orientar o co-construir en el terreno los procesos sociotécnicos más adecuados. Frente a estas limitaciones, se requieren además perspectivas capaces de anticipar, a partir de análisis empíricos detallados del presente y de los escenarios futuros, los tipos de cambios y las transiciones tecnológicas requeridas, sobre todo en el contexto actual del cambio climático y de algunos fenómenos disruptivos altamente imprevisibles, entre ellos los que generan tecnologías como la inteligencia artificial o las biotecnologías (Márquez, 2020).

En términos metafóricos, una ética de la tecnología debe prender alarmas, servir de faro para advertir los límites, servir de brújula para retomar el camino, en medio de los virajes imprevistos, ofrecer mapas o cartas de navegación, que señalen las mejores rutas y adviertan los riesgos. Este tipo de aportes son los que ofrecen las perspectivas tradicionales. Pero estas funciones no son suficientes para moverse en el terreno. Adicionalmente, se necesita un buen control del timón o del volante, para mantener el curso en medio de los cambios del clima o del terreno. Esta última cuestión la ofrecen los enfoques contemporáneos.

Los puntos cardinales de las éticas de las tecnologías

Feenberg (2005) y Mitcham y Waelbers (2009) proponen un mapa de cuatro puntos cardinales, o cuatro extremos y tendencias desde los cuales se puede evaluar los límites y los posibles problemas de las distintas perspectivas éticas de las tecnologías, y en general de las formas como hemos comprendido las interacciones entre los seres humanos y las tecnologías.

Mitcham y Waelbers (2009, p. 371) presentan en la figura 9 un mapa de puntos cardinales de la ética de la tecnología.

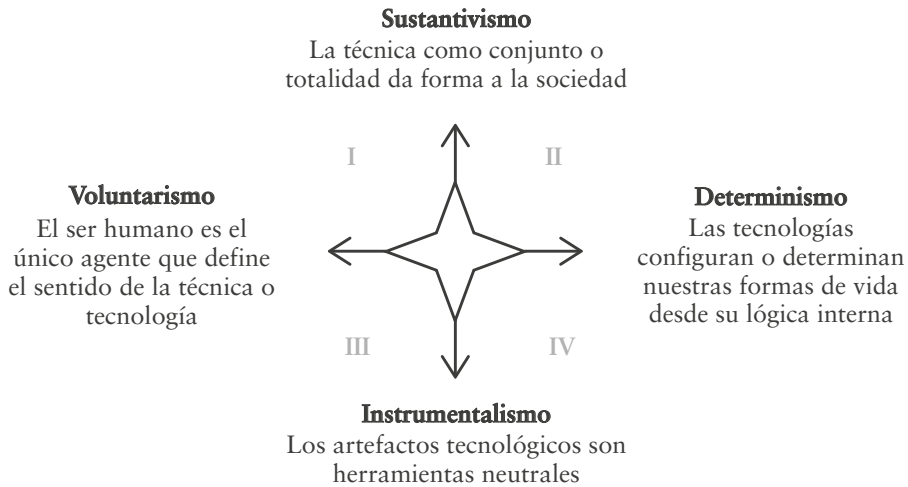


Figura 9. Puntos cardinales de la ética de las tecnologías

Fuente: elaboración propia basada en Mitcham y Waelbers (2005).

Estas posiciones extremas o puntos cardinales se pueden explicar de la siguiente manera:

Para el *voluntarismo*, los únicos agentes son los seres humanos, por tanto, el desarrollo de las tecnologías está determinado exclusivamente por la voluntad y las intenciones humanas. Los problemas morales con las tecnologías son problemas no de las tecnologías en sí mismas, sino de lo que hacemos los seres humanos con ellas. Esta perspectiva es compatible con la instrumentalista.

Para el *instrumentalismo*, las tecnologías son productos humanos neutrales pasivos, o meros instrumentos. Los únicos que tienen el estatus de agentes son los seres humanos, pero esa agencia se puede depositar de alguna manera en los objetos técnicos en tanto instrumentos.

Para el *sustantivismo*, lo tecnológico se analiza como una totalidad homogénea o indiferenciada, que deviene de manera casi

autónoma en una cultura técnica específica, y genera un efecto sobre la sociedad. Es decir, esa totalidad puede devenir de manera sustantiva en una cultura técnica, que determina a la sociedad de diferentes formas. Se pueden juzgar de sustantivistas las perspectivas filosóficas que hablan de la técnica en general o como conjunto, y la responsabilizan de algunos problemas, como la instrumentalización y alienación de la vida. Esta perspectiva puede acusar un determinismo y un esencialismo.

Para el *determinismo*, los objetos técnicos pueden ser responsables de algunos males o beneficios, como entes que ostentan un sentido de agencia en algún sentido fuerte o débil. Pueden producir por sí mismas algún efecto material, positivo o negativo, sobre la sociedad o determinar la sociedad (Smith y Marx, 1994).

De acuerdo con Mitcham (1989), pueden juzgarse de deterministas tanto las perspectivas positivistas que asocian el desarrollo tecnológico con el progreso de las sociedades, como los puntos de vista del movimiento ingenieril de los primeros teóricos de la tecnología, o la evaluación negativa de Marx sobre la Revolución Industrial, o las perspectivas de Ellul (1964) y de Gille (1978). En contraposición, buena parte de los filósofos de la tecnología clásicos, como Heidegger (1994), Ortega y Gasset (1996), Mumford (1964), etc., y la mayoría de las perspectivas normativas acusan un punto de vista voluntarista en la medida que asignan a la intencionalidad y voluntad humana toda la responsabilidad de los efectos e implicaciones del desarrollo tecnológico. Feenberg (2005) cuestiona a la perspectiva heideggeriana, la de Ellul y la de Marcuse de sustantivistas, en la medida que se refieren a la técnica como un todo, con su propia lógica, que deviene en una forma de instrumentalización y alienación de la vida humana. Y pueden calificarse de instrumentalistas las posiciones sostenidas por la mayoría de las personas, o algunas de las perspectivas normativas, que plantean que los objetos tecnológicos son neutrales y la responsabilidad moral depende exclusivamente del uso realizado por los seres humanos de esos objetos.

Estos cuatros puntos cardinales sirven también de límite para establecer una demarcación entre las perspectivas tradicionales y las

perspectivas más recientes que evitan, de manera crítica, inclinarse hacia alguno de estos extremos.

Los tipos de instrumentalización tecnológica

En el ámbito de la filosofía de la tecnología, es fundamental comprender el concepto de *instrumentalización*, que ha sido utilizado durante un extenso período de tiempo con diversas connotaciones, algunas de ellas de carácter esencialista y negativo. No obstante, este concepto no tiene un significado preestablecido negativo o positivo, sino que se refiere al proceso general mediante el cual las tecnologías median y configuran los procesos sociales, al mismo tiempo que describen cómo estas tecnologías son mediadas y moldeadas por los procesos sociales. O, en otras palabras, este término se relaciona con las maneras en que las tecnologías se utilizan como herramientas o medios para alcanzar determinados objetivos sociales desde contextos e intereses específicos, así como con la manera como estos fines, intereses y objetivos se materializan en la sociedad a través de las tecnologías.

La filosofía de la tecnología contemporánea ha cuestionado de manera profunda la idea de la neutralidad de las tecnologías, pues ellas expresan inevitablemente los valores y los intereses de los contextos particulares en los que se desarrollan, y son el producto de la materialización de procesos sociales que las hacen posibles y, al mismo tiempo, generan nuevas formas de materialización social, a menudo desapercibidas. Este doble proceso de co-construcción de lo tecnológico y lo social se denomina en términos generales como instrumentalización, y es un fenómeno que se puede manifestar y comprender de múltiples maneras, ya sea en un sentido positivo o negativo, constructivo o destructivo. Esto depende en gran medida de los fines, intereses, valores y objetivos involucrados en cada situación específica. Por lo tanto, es importante reconocer que no existe una única forma esencial de instrumentalización, sino que existen numerosas posibilidades y perspectivas para su interpretación y expresión.

La tabla 6 sintetiza algunas de las interpretaciones más conocidas sobre la instrumentalización tecnológica.

Tabla 6. Tipos de instrumentalización tecnológica

Martín Heidegger (1994)	Para este autor, la instrumentalización generada por la técnica moderna es una condición ontológica de la técnica como tal. Manifiesta la esencia de la técnica. Esa instrumentalización se expresa en la tendencia a tratar todo, incluyendo a los seres humanos, como meros recursos, o materias primas, lo que lleva a la alienación del ser humano y a una forma de relación con el mundo basada en la explotación y el dominio.
Herbert Marcuse (1993)	Para este autor, la técnica se ha convertido, en la sociedad contemporánea, en un medio de control y dominación, ya que se utiliza para reforzar y perpetuar las estructuras de poder existentes. En virtud de este tipo de instrumentalización, la técnica como tal contribuye a la conformidad y la opresión en una sociedad que busca mantener el <i>statu quo</i> .
Andrew Feenberg (2016, 1991)	Este autor propone una concepción no esencialista de la instrumentalización. Ella tiene que ver con las distintas formas de interacción entre la tecnología y la sociedad, o con el proceso mediante el cual las tecnologías se adaptan a los intereses y objetivos de determinados grupos o actores sociales y, a la vez, cómo esos grupos o actores se adaptan a ciertas tecnologías, suponiendo también que ellas tienen alguna capacidad de agencia. La instrumentalización moderna no es simplemente una cuestión de la tecnología en sí, ni el tipo de instrumentalización moderna es el único tipo de instrumentalización, o no es la esencia de la técnica como tal. Para Feenberg, en la modernidad se desarrolló fundamentalmente una instrumentalización primaria, bajo los intereses y objetivos del capitalismo industrial, y teniendo en cuenta solo el proceso de diseño inicial de lo tecnológico. Pero también es posible un proceso de instrumentalización secundaria, orientado por la diversidad de intereses de los usuarios finales.
Pierre Rabardel (2011)	Para este autor, la instrumentalización tiene que ver con la relación entre las herramientas y los usuarios en el contexto de la actividad humana, desde el marco de la psicología cognitiva y de la ergonomía. La instrumentalización implica un proceso de adaptación y transformación mutua entre el usuario y la herramienta. A medida que el usuario interactúa con la herramienta, esta puede ser modificada y adaptada para satisfacer las necesidades y las metas específicas del usuario. A su vez, el usuario se adapta y aprende a utilizar la herramienta de manera más eficiente y efectiva. Para este autor, las herramientas no son solo objetos inanimados utilizados por los seres humanos, sino que tienen una dimensión cognitiva y actúan como mediadores entre los usuarios y el entorno en el que se lleva a cabo una actividad.

Fuente: Moreno-Ortiz et al. (2024, pp. 93-94).

Los sentidos de agencia de las tecnologías y sus presupuestos ontológicos

De manera esquemática se podrían distinguir cinco grandes sentidos de agencia de las tecnologías o agenciamiento tecnológico (Moreno Ortiz et al., 2024). El primero corresponde a las perspectivas que asumen la teoría estándar de la agencia y adoptan un punto de vista instrumental, según el cual los únicos agentes somos los seres humanos y las tecnologías no tienen agencia (Schlosser, 2019). El segundo abarca las posturas que proponen un determinismo tecnológico, en el que se asocia un sentido de agencia autónomo a los objetos tecnológicos que determina la vida de las personas. El tercero incluye las perspectivas que atribuyen un sentido de agencia vicaria, prestado o extendido a los objetos tecnológicos de diversas formas. El cuarto engloba las perspectivas que identifican un sentido distribuido y relacional de agencia entre sujetos y objetos. Y el quinto se refiere a las perspectivas que sugieren una autonomía agenciataria relativa en los objetos tecnológicos.

El primer sentido es el más comúnmente sostenido. Si suponemos este punto de vista, las tecnologías no tienen ninguna relevancia moral, pues suponemos que son neutrales, solo capaces de generar efectos causales.

También es común suponer el sentido determinista, en el que se asume que los objetos o procesos tecnológicos poseen propiedades inherentes o desarrollan una dinámica autónoma, influyendo decisivamente en la existencia y experiencias humanas. Este punto de vista sugiere que la vida humana se ha configurado en función de las relaciones tecnológicas establecidas.

Alejándose en parte, tanto de la concepción instrumental como de la determinista, la mayoría de las perspectivas filosóficas actuales asumen el tercer sentido de agencia de las tecnologías, en el que se supone que los objetos no son agentes, pero en ellos se puede extender un sentido de agencia prestada. Por ejemplo, según la perspectiva fenomenológica reciente del enfoque 4E (Duarte Arias, 2025), nuestras experiencias, cogniciones e interacciones con las tecnologías tienen las siguientes condiciones: 1) es extendida; nuestra experiencia se extiende al mundo material con el que interactuamos; 2) es embebida;

ellas están embebidas o inmersas en un entorno o contexto material específico; 3) es enactiva: nuestras percepciones e interacciones con las tecnologías se producen de forma enactiva, es decir, se estructuran en las acciones sobre el entorno; y 4) es encarnada: las experiencias con las tecnologías se producen de manera encarnada en nuestro cuerpo. Desde estas formas de concebir la agencia, los objetos tecnológicos son relevantes moralmente en tanto las intenciones morales se extienden, embeben, producen de manera enactiva y encarnada. Sin embargo, al decir que las intenciones y la agencia de los sujetos se inscriben, delegan, materializan, andamian, despliegan o encarnan, se presupone que las experiencias se articulan en los sujetos como centro, y se extienden en sentido amplio a los objetos. Por esta razón, se ha cuestionado estas perspectivas por sus presupuestos antropocéntricos, dualistas y trascendentales.

Tomando distancia de los anteriores enfoques y de sus presupuestos, otras perspectivas han formulado que los objetos tecnológicos pueden desplegar una capacidad de agencia en su interacción con los sujetos, de manera relacional y distribuida. Para estas perspectivas, la agencia no se centra en los sujetos y se extiende a los objetos, sino que se distribuye entre los objetos y sujetos. Ello supone una ontología relacional en lugar de dualista, a partir de la cual ni los sujetos ni los objetos son, en sentido estricto, agentes por sí mismos, sino siempre co-agentes de manera relacional. O, en otras palabras, los sujetos no tienen agencia sin los objetos y los objetos no tienen agencia sin los sujetos. La agencia es un asunto que se despliega y distribuye en las relaciones. Un punto de vista destacado de agencia distribuida se plantea en la teoría actor red, formulada por Latour, Callon y Law. Desde esta perspectiva, se argumenta que tanto los sujetos como los objetos tecnológicos son co-agentes o “actantes” de manera relacional en un escenario de mediación, o de red híbrida de co-agentes. Por ejemplo, un simple objeto como un reductor de velocidad en la vía es un híbrido donde se articulan las capacidades de agencia de los constructores, pensando en la seguridad de los peatones, y las capacidades de agencia del objeto para lograr que los conductores reduzcan la velocidad. En general, para las perspectivas que suponen un sentido de agencia distribuido, los objetos tecnológicos son relevantes moralmente

de manera necesaria o en un sentido fuerte, pues sin ellos los sujetos no pueden ser sujetos morales.

Finalmente, un sentido de autonomía agenciativa ha sido propuesto por algunos autores como Floridi (2008) para algunos objetos, como los sistemas inteligentes artificiales. Este autor aclara que ese sentido de agencia tiene menos requisitos que los de la teoría estándar de la agencia, pues no presupone la intencionalidad y la mente, sino cualquier sistema que ostente autonomía funcional, adaptabilidad e interactividad (tabla 7). En esa medida, se puede extender a muchos objetos artificiales y naturales que funcionan como sistemas adaptativos complejos. Desde esta perspectiva, los objetos tecnológicos pueden tener una clara relevancia moral.

Tabla 7. Tipos de instrumentalización tecnológica

<i>Teoría estándar de la agencia</i> En tanto la agencia supone la intencionalidad, los únicos agentes somos los seres humanos y las tecnologías no tienen agencia.
<i>Determinismo tecnológico</i> Los objetos técnicos pueden ser responsables de algunos males o beneficios, como entes que adquieren un sentido de agencia. Pueden producir algún efecto material no controlado, fuera la voluntad humana o de los propósitos o intencionalidades para los que fueron contruidos.
<i>Perspectivas de extensión de la agencia</i> Aunque los objetos técnicos no tienen el estatus de agentes, la agencia humana se puede extender a ellos. Estas perspectivas siguen siendo antropocéntricas, pues los objetos técnicos solo gozan de un sentido de agencia de manera vicaria, andamiada o delegada. Algunas perspectivas y autores representativos son: • Los objetos técnicos son componentes de cadenas causales en los que se extiende la eficacia causal humana (Olsen, 2009). • Los objetos técnicos son sustitutos a los que se les delega la eficacia causal. • Los objetos técnicos son “operadores de posibilidad” (Broncano, 2009). • Los objetos técnicos tienen agencia “situada” o “encarnada” (<i>embodied action</i>) (Merleau-Ponty, 1945; Di Paolo, 2005). • Construccinismo social: los objetos técnicos son formas de representación social, formas de materialización de las intenciones de la sociedad; o vectores de influencia material de la sociedad sobre sí misma.

Perspectivas relacionales o de agencia distribuida

Los objetos técnicos tienen el estatus de co-agentes desde los presupuestos de una ontología relacional. En sentido estricto, ni los sujetos ni los objetos son agentes sino siempre co-agentes. Algunas perspectivas y autores representativos son:

- Lambros Malafouris (2013) y Tim Ingold (2000): en el trabajo creativo artesanal, no es en el sujeto o en el objeto donde reposa la agencia, sino en la interacción dinámica entre “cerebro, cuerpo y recursos”, o entre “alfarero, rueda y arcilla”. La agencia es una condición que pertenece al conjunto del involucramiento material entre los objetos técnicos y los humanos, o es una propiedad relacional y emergente.
- La TAR: Bruno Latour (2005): en los procesos técnicos, la agencia se encuentra distribuida, de manera simétrica y continua, en los actores humanos y los actantes no humanos. En sentido estricto, ambos actores son coagentes, de forma relacional.
- Peter-Paul Verbeek: los objetos técnicos son relevantes moralmente de manera fuerte debido a que la agencia se encuentra distribuida de manera relacional entre los actores humanos y los objetos. Los análisis morales de las perspectivas éticas tradicionales se han concentrado en la intencionalidad, pero han invisibilizado las condiciones materiales de lo moral.

Perspectivas de autonomía de la agencia

De acuerdo con Luciano Floridi (2008, 2023), es posible un sentido de agencia autónoma relativa en los sistemas inteligentes, o en los sistemas complejos. A partir de algunas distinciones categoriales, es posible entender como agente toda entidad que goce de interactividad, autonomía y adaptabilidad, aunque no suponga intencionalidad.

Fuente: elaboración propia.

Perspectivas éticas tradicionales

Éticas de las tecnologías humanistas

La designación “filosofía de la tecnología humanista” fue propuesta por Carl Mitcham (1989, pp. 21-48) para describir las corrientes que surgieron a principios del siglo xx, posteriores a los primeros enfoques sobre el tema realizados por un grupo de ingenieros que dieron origen al campo de la filosofía de la tecnología, entre ellos Ernst Kapp, P. K. Engelmeier y Friedrich Dissauer (Mitcham, 1989). Esta perspectiva humanista de la tecnología, que Mitcham contrastó con la tradición ingenieril, tuvo como antecedentes las críticas a la Revolución Industrial,

por ejemplo, la de Marx, y estuvo representada por pensadores como Ortega y Gasset, Martin Heidegger, Jacques Ellul y Lewin Mumford (Mitcham, 1989, pp. 49-81). También pueden incluirse en ella otros autores de la época como Paul Alsberg, Bertrand Gille y Marcuse, así como muchos otros pensadores contemporáneos que se inspiran en estas fuentes filosóficas.

En términos generales, podríamos ubicar en este grupo a numerosos filósofos de la tecnología, tanto del pasado como del presente, en la medida en que se alinean con los siguientes criterios (Moreno-Ortiz et al., 2024):

- 1) En sus análisis, tratan la evolución de la cultura técnica como una totalidad indiferenciada, o un fenómeno único. No discriminan las diversas formas de desarrollo y de expresión técnicas ni se enfocan en cada objeto tecnológico específico.
- 2) Predominan en sus estudios un enfoque interpretativo, ya sea hermenéutico o fenomenológico, para analizar la cultura técnica.
- 3) Consideran la técnica como una manifestación cultural o una realidad humana, y la analizan de manera crítica a lo largo del tiempo. Desde estas perspectivas, el valor principal de la filosofía reside en cuestionar la evolución de la cultura técnica en su conjunto, viéndola como una reproducción o expresión del pensamiento moderno.
- 4) En el ámbito moral, sostienen que la adecuada gestión, control o dirección del desarrollo técnico depende de un marco humano externo que lo trasciende. Esta necesidad de una perspectiva que vaya más allá de lo puramente técnico es lo que define especialmente a estas visiones como “humanistas”.
- 4) Muchos autores que adoptan esta perspectiva humanista asumen que el cambio tecnológico sigue una lógica propia que puede deshumanizar, o que implica una forma de instrumentalización, al centrarse exclusivamente en aspectos funcionales y materiales del capitalismo industrial moderno.

Algunos pensadores contemporáneos, como Han (2017), Sloterdijk (2009) y Sadin (2020) pueden ser ubicados dentro de esta perspectiva, ya que comparten algunas de sus características generales. Esta corriente sigue siendo, hasta el día de hoy, la principal crítica al devenir de la cultura tecnológica en su conjunto, o a ciertos campos del desarrollo tecnológico considerados no de manera aislada, sino como un todo, como por ejemplo las tecnologías digitales y las biotecnologías.

La visión de la técnica propuesta por Heidegger (1994) es considerada el principal punto de vista humanista sobre la técnica. Autores como Feenberg (1991, 1995, 2016), Mitcham (1989), y Mitcham y Waelbers (2009) han cuestionado la perspectiva sustantivista de la técnica heideggeriana, así como la de otros pensadores que tienden a hablar de la técnica como una totalidad. Este sustantivismo implica una generalización esencialista que puede llevar a concepciones tecnológicas deterministas, pues supone que la concepción moderna de la técnica manifiesta la esencia de la técnica como tal.

La concepción de la técnica heideggeriana supone que la técnica es una especie de sistema cultural, que da forma al mundo social como objeto de control y lo mediatiza, constituyéndose así en un modo de vida. Este tipo de discurso totalizador e indiferenciado sobre la técnica ha dado lugar a un conjunto de argumentos que encierran una cierta protesta romántica.

Esta manera totalizadora de referirse a la técnica puede ofrecer una interpretación crítica acertada de cómo las técnicas o tecnologías influyen en la sociedad. Por ejemplo, algunos contextos tecnológicos asumidos profundamente por la sociedad pueden ser la base de culturas tecnológicas específicas, con sus propias dinámicas, lógicas e implicaciones. En este sentido, nos referimos hoy a la cultura digital, la cultura algorítmica, etc. Sin embargo, este discurso totalizador puede incurrir en un sustantivismo cuando se utiliza de manera esencialista para definir lo que caracteriza a la técnica en sí misma y le adjudica procesos generales necesarios de alienación e instrumentalización.

Por otro lado, Verbeek (2005, pp. 74-95) ha criticado profundamente esta concepción de la técnica por su carácter trascendental implícito. Verbeek señala que en Heidegger se percibe una tendencia a reducir el análisis de la técnica a las condiciones que la hicieron posible

en la modernidad. Heidegger y otros autores que emplean un análisis hermenéutico existencial de la técnica se centraron en estudiar las condiciones de posibilidad de la técnica como si estuvieran hablando de la técnica misma, y de esta visión reduccionista y unilateral derivan un proceso de alienación.

Verbeek (2005, pp. 74-95) muestra que esta filosofía tradicional de la tecnología tiende a pensar “hacia atrás”, enfocándose en las condiciones de posibilidad y asumiendo que esta definición de la técnica es esencial o refleja una condición general de esta. Como resultado, dicha perspectiva ha impedido un análisis profundo de los roles específicos que las tecnologías desempeñan en la vida de las personas. Aunque la técnica moderna puede derivar en una mentalidad controladora y funcionalista, ello no significa que necesariamente se traduce en todos los casos en una forma controladora y funcionalista de interactuar con la realidad (Verbeek, 2005, pp. 74-95).

Éticas de la tecnología normativas

Las perspectivas normativas están inspiradas en la tradición de la ética del deber kantiana y se enfocan en la formulación de principios, deberes y normas para regular el desarrollo, la implementación y el uso de las tecnologías. Estas perspectivas toman como puntos de partida algunos principios generales de los que se desprenden pautas y normas concretas que regulan cada campo tecnológico. Ejemplos de estos principios generales son los propuestos por la bioética normativa, como la beneficencia, la no maleficencia, la autonomía y la justicia. Asimismo, existen marcos normativos específicos que regulan desarrollos tecnológicos en áreas como las biotecnologías, las tecnologías de la información y la comunicación, la inteligencia artificial, etc.

La principal perspectiva ubicada en esta tradición es el enfoque normativo y principialista de la bioética (Beauchamp y Childress, 1979). Este enfoque se ha extendido a la regulación de las tecnologías en diversos campos, y se ha ampliado con la inclusión de más principios, como el de responsabilidad, propuesto por Jonas (1984), y el de precaución, introducido por la Resolución de Niza del Consejo Europeo en 2000, así como el principio de privacidad, entre otros. Un ejemplo

de un planteamiento ético normativo para las tecnologías basado en el principialismo de la bioética es el texto de Peterson (2017).

Estas perspectivas comparten las siguientes características (Moreno et al., 2024, p. 32):

- 1) Se centran en controlar los límites y riesgos, y en prevenir problemas y situaciones negativas asociadas a las tecnologías, más que en promover sus desarrollos positivos.
- 2) Se enfocan especialmente en la regulación de las acciones humanas relacionadas con el desarrollo, la implementación y el uso de las tecnologías, asumiendo que estas son neutrales. Es decir, tienden a ser instrumentalistas.
- 3) Tienen un carácter prescriptivo y normativo, en lugar de interpretativo.
- 4) En su mayoría, se basan en una ética formal que parte de principios generales, aunque suelen concretarse en normas específicas, algunas de las cuales tienen implicaciones legales.

Desde una perspectiva normativa, se han desarrollado numerosas propuestas éticas específicas y especializadas para abordar las tecnologías. Franssen et al. (2024) identifican en la Enciclopedia Stanford de Filosofía una gran cantidad de marcos normativos para las siguientes áreas: ética de la ingeniería, ética informática, ética de la robótica, ética de la inteligencia artificial, ética de las máquinas y de los algoritmos, ética de las biotecnologías, ética de la arquitectura y la planificación urbana, ética de la nanotecnología y las tecnologías convergentes, ética de la disuasión nuclear, etc.

Los enfoques normativos proporcionan los parámetros éticos más concretos y aplicables para la regulación de las tecnologías, lo que los hace indispensables. No obstante, estos enfoques suelen centrarse principalmente en el “control” de los posibles problemas potenciales, y en determinar sus límites de desarrollo, implementación y uso. Aunque son fundamentales, resultan demasiado generales y poco efectivos para

orientar y fomentar el desarrollo positivo y las oportunidades que surgen con los cambios tecnológicos.

Éticas de las tecnologías pragmáticas y procedimentales

Las perspectivas pragmáticas y procedimentales se diferencian de las humanistas y normativas porque no se comprometen ni parten de una perspectiva de sentido para orientar lo técnico, ni de un parámetro normativo para buscar el control. En su lugar, determinan lo correcto o incorrecto de manera contextual, basándose en el cálculo de riesgos o de manera consecuencialista, evaluando los costos y beneficios del desarrollo tecnológico, o las consecuencias sociales de los cambios tecnológicos, evaluadas *a posteriori*. Estas perspectivas buscan el control y regulación de la tecnología con acciones evaluativas emprendidas por la sociedad, a través de mecanismos procedimentales y de participación democrática. Se inspiran en el utilitarismo ético y el pragmatismo norteamericano. Este enfoque puede identificarse actualmente en algunos activismos dentro del campo CTS o en movimientos ambientalistas. Durbin (2003) desarrolla una propuesta ética en esta línea y Keulartz et al. (2002, 2004) proponen una ética pragmática.

Estas perspectivas comparten las siguientes características (Moreno et al., 2024, p. 34):

- 1) Cuestionan las bases metafísicas de la ética, en particular el fundamentismo, esencialismo y dualismo presentes en muchas perspectivas filosóficas humanistas y normativas. Se centran en el control de los límites, riesgos y problemas tecnológicos a través de la deliberación y evaluación social, con un enfoque especialmente político.
- 2) Son contextuales, no universalistas. Buscan, a través de procesos de participación, deliberación y evaluación, un control social y político de los problemas tecnológicos.
- 3) Son prescriptivas, pero no de manera *a priori* como las normativas, sino *a posteriori*, basándose en la evaluación de los cursos de acción moral. La prescripción moral se realiza de manera posterior,

a través de la deliberación y evaluación social, a diferencia de otras perspectivas que parten de parámetros establecidos *a priori*.

4) Son procedimentales. Se centran especialmente en los mecanismos sociales y políticos de control de los riesgos asociados a las tecnologías.

Al distanciarse de las fundamentaciones metafísicas de los enfoques humanistas y normativos, estas perspectivas eluden muchas de las críticas que tienen esos enfoques. Sin embargo, al igual que las perspectivas normativas, se enfocan en el control de límites y riesgos, aunque no *a priori*, sino *a posteriori*, a través de procesos políticos y sociales, y en esta medida comparten las desventajas del enfoque normativo, de ser insuficiente para garantizar una orientación moral en los desarrollos tecnológicos, por sus limitaciones para orientar la exploración de alternativas.

Además, el carácter procedimental de estas perspectivas se enfrenta a una cuestión circular difícil de resolver. Este enfoque promueve la deliberación ciudadana, activa y participativa para asegurar el control de los riesgos morales asociados al desarrollo tecnológico, pero, a su vez, presupone unas condiciones democráticas de base.

Análisis comparativo de las perspectivas

El siguiente análisis comparativo de las perspectivas tomado de Moreno et al. (2024, pp. 35-36) resalta las similitudes y diferencias según varios criterios, para ayudar a su comprensión:

Tabla 8. Síntesis enfoques éticos tradicionales

	Éticas humanistas	Éticas normativas	Éticas pragmáticas y procedimentales
Enfoque	Se enfocan en el análisis del sentido y los fines de la técnica como totalidad o como un conjunto.	Se enfocan en el control de los límites, los riesgos y los problemas de las tecnologías desde un punto de vista prescriptivo o normativo.	Se enfocan en el control de los límites, los riesgos y los problemas de las tecnologías desde la deliberación y evaluación social y en un sentido político.

	Éticas humanistas	Éticas normativas	Éticas pragmáticas y procedimentales
Sentido contextual	No son contextuales. Consideran que el gobierno, el control o la correcta orientación del desarrollo técnico depende de un ámbito humano externo <i>a priori</i> que sobrepasa y trasciende lo meramente técnico.	No son contextuales. Desde un marco normativo <i>a priori</i> regulan las acciones humanas en relación con los desarrollos, implementaciones y usos de las tecnologías, suponiendo la neutralidad de las tecnologías.	Son contextuales, no universalistas. A través de procesos de participación, deliberación y evaluación <i>a posteriori</i> buscan un control social y político de los problemas de las tecnologías.
Alcance	Tienen un sentido interpretativo hermenéutico o fenomenológico sobre la cultura técnica.	Tienen un sentido prescriptivo, en lugar de interpretativo.	Son prescriptivas, pero <i>a posteriori</i> , a partir de la evaluación de los cursos de acción morales.
Forma de operar	Son éticas de contenidos. Consideran a la técnica como un tipo de cultura o realidad humana, y analizan su devenir de manera crítica. Analizan la deshumanización derivada de la técnica, en la medida que la técnica se orienta por una lógica propia deshumanizante, o instrumentaliza la existencia humana.	Son éticas formales. Parten de principios generales, y se concretan en normas específicas, algunas de ellas con implicaciones legales.	Son éticas procedimentales. Se concentran especialmente en los mecanismos sociales y políticos de control de los riesgos asociados a las tecnologías.
Limitaciones	Tienen a ser sustantivistas y deterministas. Plantean el análisis moral de manera transcendental.	Tienden a ser instrumentalistas y voluntaristas. Plantean el análisis moral de manera transcendental.	Tienden a ser instrumentalistas y voluntaristas.

Fuente: Moreno et al. (2024, pp. 35-36).

Nuevas orientaciones en las éticas de las tecnologías

Además de las contribuciones de las diversas perspectivas éticas tradicionales, en las últimas décadas, especialmente desde 1980, se han producido transformaciones teóricas significativas que han dado lugar a nuevos enfoques en la forma de abordar los problemas éticos de las tecnologías.

En esta sección, se examinarán estas transformaciones teóricas que sustentan posiciones críticas hacia las perspectivas tradicionales y aportan nuevas ideas. Acogiendo la propuesta de Moreno Ortiz et al. (2024, pp. 57-87), los cambios contemporáneos se pueden identificar alrededor de cinco giros teóricos que engloban las perspectivas: 1) el giro empírico en el estudio de las tecnologías, 2) el giro social, 3) el giro ontológico en torno a la naturaleza de las tecnologías, 4) el giro material y 5) el giro normativo, que propone una nueva comprensión de la normatividad ética en el contexto tecnológico.

Estos giros se relacionan con algunos propósitos y funciones diferentes a los de las perspectivas tradicionales y responden a los nuevos desafíos y exigencias que plantean las tecnologías hoy en día. En el mundo complejo en el que vivimos, donde mantenemos múltiples y variadas interacciones con las tecnologías, no basta con criticar, establecer límites, advertir sobre peligros o controlar. También es necesario ofrecer alternativas y orientar la construcción de nuevas formas de desarrollo tecnológico. Se requieren análisis morales específicos y particulares, en lugar de enfoques generales o abstractos, como los humanistas o normativos, que permitan distinguir los desarrollos, implementaciones y usos más éticos en cada caso. Es fundamental también construir nuevas formas de desarrollo tecnológico alternativo con una orientación ética explícita desde el diseño, en lugar de adoptar posturas críticas, escépticas o distópicas de manera evaluativa y tardía.

Las perspectivas tradicionales ofrecen opciones de análisis limitadas en una sociedad que busca orientar adecuadamente el desarrollo tecnológico en cuestiones como la sostenibilidad, el cambio climático y la equidad, sin abandonar por completo las formas de vida tecnológicas ya adoptadas y evitando polarizaciones entre visiones positivistas, optimistas, tecnófilas y visiones negativas, pesimistas y tecnófobas.

En el contexto de la actual cultura tecnológica, estas perspectivas tradicionales presentan las siguientes limitaciones (Moreno Ortiz et al., 2024, p. 59): 1) tienen un contenido empírico débil, ya que suelen referirse a las tecnologías en términos generales; 2) son poco concretas debido a su carácter abstracto y formal; 3) abordan los problemas éticos exclusivamente desde la acción humana, sin considerar el agenciamiento material de los objetos tecnológicos; 4) consideran la

regulación moral únicamente desde las posibilidades normativas morales de los sujetos, con una comprensión limitada de la normatividad inherente a las tecnologías; y 5) se centran en la evaluación, pero son débiles para ayudar a construir soluciones; es decir, son efectivas para formular críticas, establecer límites, controlar y advertir sobre riesgos, pero muy limitadas para orientar la construcción de opciones en el terreno, vislumbrar alternativas y señalar posibilidades.

El giro empírico en la filosofía de la tecnología

Como se señalan Moreno et al. (2024, pp. 61-64), este giro ha sido planteado explícitamente por la filosofía norteamericana de la tecnología, como se detalla en el resumen de la filosofía de la tecnología estadounidense editado por Hans Achterhuis (2001), titulado *American Philosophy of Technology: The Empirical Turn*. Kroes y Meijers (2000), Achterhuis (2001) y Brey (2010) explican los detalles de este giro en la filosofía de la tecnología de las últimas décadas, y no solo identifican un giro, sino dos.

El primer giro se dio en las décadas de 1980 y 1990, y se caracteriza por una orientación de los filósofos neoheideggerianos, teóricos neocríticos y fenomenólogos críticos hacia las tecnologías específicas, las teorías contextuales menos deterministas y los enfoques más pragmáticos. Esta nueva perspectiva se apoyó en los métodos de análisis empíricos y en los marcos teóricos de los estudios sociales de la ciencia y la tecnología (STS). Entre los pensadores destacados en esta tendencia se encuentran Feenberg (1991), Ihde (1990), Mitcham (1994), Dreyfus (1992), Hickman (2001), Light (2001) y Haraway (1995). Por ejemplo, Feenberg, desarrolló una teoría crítica de la tecnología combinando elementos de la teoría crítica de la tecnología de la Escuela de Frankfurt, con el constructivismo y los enfoques STS. De esta manera, pudo ofrecer un análisis social y contextual de la tecnología, diferente de las perspectivas sustantivistas de Heidegger e instrumentalistas de la Escuela de Frankfurt. Por su parte, Don Ihde construyó una perspectiva fenomenológica, llamada “postfenomenología”, que se centra en cómo la experiencia, la acción, la subjetividad y el mundo se constituyen a través de las mediaciones tecnológicas (Ihde, 1990).

Otros autores, como Hickman (2001), Light (2001) y Haraway (1995) promovieron planteamientos más empíricos y descriptivos, tomando aportes del pragmatismo, el postestructuralismo, el campo STS, los estudios culturales y los estudios de la comunicación.

En las décadas de 1990 y 2000, autores como Mitcham (1994), Pitt (1995), Kroes y Meijers (2000), y Meijers (2009) discutieron un segundo giro empírico en la filosofía de la tecnología. Este giro se orienta especialmente hacia las orientaciones empíricas ingenieriles de la tecnología, que abordan los objetos concretos, pero sin llegar a afirmar el positivismo característico de los enfoques ingenieriles de inicios del siglo XX. Para estos autores, la filosofía tradicional de la tecnología no ha tratado adecuadamente las tecnologías concretas y materiales, en la medida que se ha concentrado en especial en las implicaciones sociales de los cambios tecnológicos. En otras palabras, la filosofía tradicional se ha enfocado en las implicaciones de las tecnologías, en lugar de abordar las tecnologías concretas, o las tecnologías en su sentido material (Brey, 2010).

Por lo tanto, para volver a centrar la filosofía en las tecnologías en sí mismas, estos autores proponen que se debe priorizar la descripción sobre la evaluación, sobre todo en relación con las prácticas tecnológicas en contextos específicos y los principales problemas tecnológicos, en lugar de solo las representaciones sociales de las tecnologías. Los procesos tecnológicos revelan toda su complejidad cuando se analizan a partir de sus prácticas concretas, en contraste con el enfoque limitado derivado de las críticas a los procesos tecnológicos propios de la lógica mercantil eficientista del capitalismo moderno. La ética de la tecnología debe comenzar con el estudio empírico de las prácticas tecnológicas concretas para orientar su desarrollo ético positivo, en lugar de limitarse a un análisis evaluativo.

El giro social de las éticas de las tecnologías

Existe un importante cambio entre la filosofía de la tecnología tradicional y la contemporánea, en relación con la forma como entendemos el sentido social de las tecnologías y su carácter valorativo. Tradicionalmente hemos concebido los objetos tecnológicos como

meras cosas materiales, neutrales valorativamente, o que no involucran valores en sí mismos. Por lo tanto, lo valorativo se concibe de manera externa a los objetos, que solo depende de los sujetos, o del uso que los sujetos realizan de los objetos. Este punto de vista es el que sostienen las perspectivas instrumentales, como se vio, pero ha sido ampliamente cuestionado en la filosofía de la tecnología contemporánea.

Las críticas más fuertes a este punto de vista provienen del constructivismo social de las tecnologías (Bijker et al., 1987), desde el cual las tecnologías no son meras cosas, sino formas de materialización de representaciones, valores, intereses, etc. de unos actores y un contexto social específico, y ninguna tecnología es neutral en sentido estricto. El sentido de las tecnologías se construye socialmente. Pero en nuestras relaciones cotidianas con las tecnologías se invisibiliza ese sentido social, a través de lo que los sociólogos denominan “efecto caja negra”:

Una vez que la tecnología se diferencia de otros aspectos sociales, su interacción con ellos parece ser externa. Esto es particularmente claro en el caso de las mediaciones. El arte ya no es una parte intrínseca de la práctica técnica, sino un agregado posterior. Las columnas del Partenón no constituyen una ornamentación en nuestro sentido, sino que pertenecen integralmente a su diseño; nosotros colocamos las columnas en la fachada una vez finalizadas las estructuras de soporte. Los valores éticos regulan la tecnología desde afuera, a través de leyes; no son componentes internos de la práctica de la ingeniería. (Feenberg, 2016, p. 240)

Esta diferenciación entre lo tecnológico y lo social realizada por las perspectivas esencialistas de la tecnología puede conducir a lo que Feenberg (1991) denomina como “fetichismo tecnológico”, de manera análoga a lo que Marx denomina fetichismo de las mercancías.

Para Marx, el fetichismo de las mercancías es la creencia errónea de que el valor de las mercancías o productos reside en ellos mismos, como si los precios fueran una característica real o inherente a los productos. Sin embargo, los precios son el resultado de una valoración de las relaciones que se establecen entre los fabricantes, los distribuidores y los consumidores, en los que se establece un valor de cambio. En

este fetichismo, las relaciones de producción e intercambio parecen ser entre objetos, como el dinero y los productos, en lugar de relaciones sociales entre personas.

De manera análoga, según Feenberg (2016), el fetichismo de la tecnología trata a los productos tecnológicos como si fueran cosas objetivas y oculta su naturaleza relacional, social y valorativa. En otras palabras, la tecnología se presenta principalmente en términos de su función y se ocultan sus aspectos éticos, estéticos y sociales. Desde esta perspectiva fetichista, la tecnología se considera social o ética solo cuando se utiliza para algo social o para un propósito ético específico (Feenberg, 2016, pp. 242-243).

El giro ontológico relacional en la filosofía de la tecnología

Las respuestas a la pregunta ¿qué son las tecnologías?, relacionadas con las concepciones sobre la naturaleza de las tecnologías o la ontología de las tecnologías, también han experimentado un importante cambio en las décadas recientes. Desde los planteamientos de autores como Simondon (2007), Stiegler (1998) y Latour (2005), se ha pasado de visiones esencialistas y ontológicas dualistas de las tecnologías, sostenidas especialmente por las perspectivas humanistas clásicas, a concepciones relacionales. Este cambio tiene implicaciones fundamentales para la ética de las tecnologías.

Para comprender este giro el autor más representativo es Simondon (2007, 2015). En sus textos, *El modo de existencia de los objetos técnicos* (2007) y *La individuación a la luz de las nociones de forma y de información* (2015), este autor cuestiona los planteamientos que intentan gobernar o controlar el desarrollo técnico desde una perspectiva que trasciende la técnica misma, basados en una separación entre técnica y cultura dentro de una ontología dualista moderna. A este respecto, argumenta que las críticas humanistas a la técnica se centran en cuestionar la abstracción de las operaciones técnicas promovida por la gestión capitalista de los recursos técnicos y no es un cuestionamiento a las distintas formas de desarrollo técnico. Sin embargo, esta perspectiva no solo carece de una base ontológica sólida, sino que también fomenta una actitud negativa o de prevención hacia los objetos técnicos,

como si fueran externos a la cultura: “La cultura se comporta con el objeto técnico como el hombre con el extranjero cuando se deja llevar por la xenofobia primitiva” (2007, p. 31).

Estas limitaciones de la perspectiva humanista se pueden superar, según Simondon (2007), desde una teoría ontológica en la que la invención técnica se entienda como un proceso de individuación y concretización humana. Específicamente, para él solo desde una perspectiva ontológica relacional se puede liberar los objetos técnicos de las representaciones inadecuadas del capitalismo industrial o de las visiones tecnocráticas que los ven simplemente como manipulación de la naturaleza.

Una explicación del punto de vista ontológico de Simondon (2007), y de sus implicaciones para la ética, se puede consultar en *La ética de la técnica en Gilbert Simondon*, de Moreno-Ortiz (2022, pp. 159-185). Allí se precisa que la individuación de los objetos técnicos no se basa en su función externa, sino en su funcionamiento interno, pues identificar un objeto por su función es superficial, porque varios objetos pueden cumplir la misma función, aunque sean diferentes. La función se relaciona con requisitos externos a la realidad técnica, mientras que el funcionamiento es el resultado de un proceso específico de la técnica, denominado “concretización”. Este proceso se define a través de momentos de reconfiguración estructural o actos de invención, que buscan resolver problemas estableciendo una mayor coherencia y estabilidad entre los elementos y su entorno.

La separación entre lo técnico, lo social y lo cultural surge cuando se omite la perspectiva del técnico, quien entiende mejor la lógica del funcionamiento interno del objeto técnico, y se la reduce al rol que le asigna el organizador industrial, que solo se interesa por la función. El organizador industrial ajusta el proceso técnico según una finalidad externa, realizando una doble abstracción: representa las máquinas mediante evaluaciones cuantificadas, económicas (rentabilidad) o energéticas (eficiencia), y comprende los conjuntos técnicos solo después de su operación, en el proceso final, ocultando el complejo proceso de desarrollo de lo técnico o, como dicen los sociólogos de la tecnología, generando un efecto de caja negra sobre su desarrollo. En contraste, el técnico logra interpretar y comprender el funcionamiento de las

máquinas, atendiendo a las interconexiones intrínsecas a los ensamblajes sociotécnicos (Moreno-Ortiz, 2022; Simondon, 2007).

Para Simondon, los conjuntos técnicos pueden verse como un “tejido de individuos técnicos interconectados”, y los técnicos como colaboradores de las máquinas. Esto contrasta con la “filosofía autocrática de las técnicas” o la “filosofía tecnocrática” moderna, que asume los procesos técnicos como recursos para obtener poder, conquistar y dominar la naturaleza mediante el uso de las máquinas.

Desde la ontología relacional de Simondon, el análisis moral de los objetos técnicos se ve como un análisis moral de las formas de concretización técnica. Este enfoque se centra en perfeccionar el funcionamiento más que en maximizar la función. Los objetos técnicos son los medios en los que se concretizan otras formas de individuación, o se logran materializar las transducciones psíquicas y colectivas generadas por los individuos y las sociedades. Desde esta perspectiva, la normatividad moral de los objetos técnicos no es externa a la normatividad moral humana, sino una expresión de ella. Por lo tanto, la cultura material es una forma de individuación de la cultura moral, y viceversa.

Para Simondon, la moralidad no se limita a nuestra capacidad de ser intencionales, libres y autónomos, sino que también se manifiesta en nuestra existencia como seres técnicos, es decir, en nuestras concretizaciones materiales. Esta perspectiva cuestiona la visión tradicional de la ética, que ha tendido a considerar la moralidad como derivada exclusivamente de la racionalidad y la intencionalidad.

El giro material en las éticas de las tecnologías. La relevancia moral de los artefactos tecnológicos

Como se muestra en los textos de Moreno-Ortiz et al. (2024, pp. 70-82) y Moreno-Ortiz (2019), para las éticas tradicionales de la tecnología los objetos tecnológicos concretos no cuentan o no son relevantes moralmente, en la medida que se centran únicamente en la voluntad e intencionalidad de los sujetos. Estas perspectivas no consideran el sentido moral que podría derivarse de la materialidad de los objetos tecnológicos.

Las éticas tradicionales de la tecnología han sido voluntaristas e instrumentalistas, ya que han asumido que la moralidad depende exclusivamente de la capacidad de agencia de los seres humanos, ignorando la capacidad de agencia de los objetos tecnológicos y suponiendo su neutralidad. No han sido éticas materiales en sentido estricto, en la medida que la cultura material y la materialidad específica de los objetos tecnológicos no ha contado para nada en el análisis moral.

En un sentido más estricto, estas éticas han sido externalistas, pues han supuesto que la agencia moral depende exclusivamente de las deliberaciones humanas o de las facultades externas a la materialidad de los artefactos o a los objetos tecnológicos en sí mismos. No son éticas “de” las tecnologías, sino éticas “para” las tecnologías, enfocadas en lo que los seres humanos hacemos con las tecnologías, bajo el supuesto de que ellas son neutrales. Esto se manifiesta en la siguiente expresión común: “las tecnologías no son buenas o malas, pues son cosas neutrales. Lo bueno o lo malo es lo que hacemos con ellas”. Estas perspectivas no han asumido el giro empírico de la filosofía de las tecnologías contemporáneo, en el que se promueve basar el análisis de las tecnologías en las tecnologías mismas en tanto cosas o procesos materiales (Luna Herrera, 2026; Antolínez Uribe, 2023).

Más allá de un determinismo tecnológico ingenuo, hoy en día, algunas corrientes filosóficas han debatido extensamente la relevancia moral de los objetos tecnológicos. Según estas nuevas perspectivas, los objetos son relevantes en el análisis moral, en la medida que tienen un sentido de agencia que influye en el curso de nuestras acciones, configuran nuestras percepciones y forman parte integral de nuestras experiencias.

Desde el enfoque de la agencia material de las tecnologías propuesto por muchas corrientes contemporáneas —como la teoría actor-red de Latour (1998a, 1998b, 2002, 2003, 2005), las perspectivas de Ingold (2000, 2012), Malafouris (2013), Verbeek (2005, 2011) y Kroes y Verbeek (Eds., 2014)— ha ganado fuerza la idea de que los artefactos tienen relevancia moral, ya que poseen un sentido de agencia que influye en las acciones morales. De este modo, la ética de la tecnología no es ajena a los objetos tecnológicos en sí mismos como objetos materiales.

Estas ideas se pueden comprender más fácilmente desde algunos ejemplos de interacciones ordinarias con los objetos tecnológicos. En la actualidad, la percepción de nuestros contextos sociales está mediada en gran medida por lo que nos muestran las redes sociales. Percibimos la realidad desde los lentes de varias tecnologías que se han integrado profundamente en nuestras vidas. Incluso, algunas de ellas operan de manera invisible para nosotros, como los sistemas que filtran lo que vemos en las redes y los motores de búsqueda. Estas tecnologías no solo agencian nuestras percepciones, sino que también influyen sobre nuestras acciones y sobre la constitución de nuestra subjetividad. Un agenciamiento moral más explícito lo podemos ver en los sistemas inteligentes artificiales que tienen roles morales, como los que se emplean para la selección de personal, o para conceder créditos bancarios, etc.

En términos generales, el giro material producido en la filosofía de la tecnología se relaciona con la relevancia concedida a los objetos materiales en el análisis moral, dependiendo de su sentido de agencia. Así mismo, reconocer la relevancia de los objetos materiales es fundamental para enfocar la ética de las tecnologías hacia la construcción material de los procesos éticos, y no solo hacia el control de los sujetos sobre los objetos.

La tesis de la relevancia moral de los objetos tecnológicos en tanto objetos materiales ha sido posiblemente el punto de quiebre teórico más importante acontecido en la ética de la tecnología reciente. Una discusión detallada de este planteamiento se puede consultar en el texto de Kroes y Verbeek (2014), o más sintéticamente en los artículos de Moreno-Ortiz (2019) y Durán (2020). En términos generales, la tesis de la relevancia moral sostiene que los objetos tecnológicos cuentan —o son relevantes— en el análisis moral, ya que, por su capacidad de agenciamiento, afectan las acciones, y no únicamente los resultados de estas acciones. Dicho de otro modo, pueden actuar como mediadores y no solo como medios o meros recursos. En otras palabras, la tesis afirma que los objetos pueden incluir en el curso de las acciones, lo que los hace relevantes en el análisis moral. De esta manera, no se limitan a ser objetos de estudio —de análisis moral—, sino que también pueden considerarse fuentes de moralidad.

Esta tesis cuestiona el voluntarismo y el instrumentalismo de las perspectivas tradicionales. Según estas perspectivas, los objetos tecnológicos no tienen capacidad de agencia y no moldean o median nuestras acciones, aunque pueden desempeñar un papel causal en una cadena de eventos, aproximadamente en el mismo sentido que los fenómenos naturales (Moreno-Ortiz, 2019, p. 101).

Desde una perspectiva fenomenológica, o más exactamente, “postfenomenológica”, Ihde (1979, 1990, 1993, 2009, 2010, 2012) y Verbeek (2011) explican de la siguiente manera los dos sentidos en los que los objetos tecnológicos son mediadores: primero, en la medida que ayudan a definir la realidad en la que estamos inmersos, al mediar nuestra percepción e interpretación; y, segundo, en tanto dan forma a las maneras como interactuamos o nos hacemos presentes en la realidad, al mediar las acciones y las prácticas humanas.

Estos dos sentidos se pueden ilustrar con dos ejemplos. El primer sentido plantea que algunas tecnologías con las que interactuamos no son el fin de nuestra percepción y acción, sino el medio a través del cual percibimos y actuamos en la realidad. Por ejemplo, al conducir un automóvil nuestra atención no se concentra en el objeto en sí mismo, sino en la carretera, en los peatones, en el tráfico y en las señales de tráfico. Solo colocamos nuestra atención en el automóvil cuando alguien aún no sabe cómo conducir un automóvil o cuando se descompone.

El segundo sentido se puede ilustrar con un ejemplo propuesto por Verbeek (2011). Específicamente, él explica cómo la tecnología puede cambiar la forma en que actuamos y vemos el mundo que nos rodea, a través del ejemplo de un diagnóstico genético temprano del cáncer de mama hereditario. Algunas pruebas pueden predecir si alguien tiene un alto riesgo de desarrollar esta enfermedad debido a mutaciones genéticas. Las personas que presentan estas mutaciones pueden elegir hacerse pruebas regulares para detectar el cáncer temprano o incluso someterse a una cirugía para reducir el riesgo. Esta tecnología nos hace enfrentar una nueva responsabilidad de cuidar nuestra salud antes de estar enfermos, lo que cambia nuestra percepción de la realidad.

Adicionalmente, las tecnologías pueden mediar también las formas como construimos la subjetividad, la objetividad y definimos la realidad, o las formas como nos vemos a nosotros mismos y como

entendemos el mundo. Verbeek (2005, 2011) explica esto con el ejemplo del uso del ultrasonido durante el embarazo para ver al feto. Esta tecnología representa al feto como un paciente potencial y a los padres como potenciales padres con decisiones morales que tomar. El ultrasonido cambia la experiencia de ser padres, de simplemente esperar un bebé a tomar decisiones importantes, como en casos de problemas de salud fetal. Esta tecnología aísla al feto de su entorno, la madre, y nos ayuda a definir la realidad de una manera diferente.

Desde la perspectiva de este autor, la intencionalidad, la libertad y la responsabilidad humanas siempre han estado moldeadas y co-formadas por los artefactos en las prácticas tecnológicas. La cuestión no radica solo en entender lo que pueden generar o no las tecnologías, sino en cómo interpretar las intenciones, la libertad y la responsabilidad considerando la permanente mediación tecnológica en las acciones humanas. El análisis de estas mediaciones materiales ha revelado que estos aspectos nunca han sido exclusivamente resultado de nuestros procesos internos introspectivos, sino que también han estado influenciados por las condiciones materiales que afectan nuestra percepción y acción. La intencionalidad, la libertad y la responsabilidad nunca han existido de forma aislada en el sujeto, sino que siempre han estado mediadas por factores materiales.

El giro normativo de las éticas de las tecnologías

Desde un punto de vista amplio, toda perspectiva ética tiene un sentido normativo, no solo aquellas tradicionalmente catalogadas como normativas o del deber, como las kantianas, en la medida que cada enfoque ético establece criterios para determinar lo que es correcto o incorrecto. Por ejemplo, las éticas de inspiración kantiana establecen deberes sobre lo que es correcto e incorrecto, mientras las éticas humanistas establecen interpretaciones y orientaciones sobre lo que puede ser correcto e incorrecto en las culturas tecnológicas.

Suponiendo este sentido normativo de todas las perspectivas éticas de la tecnología, es importante resaltar el giro general que se ha producido en las perspectivas contemporáneas, en la forma de pensar esa normatividad. Ella se ha pensado no solo como algo externo, sino

también como una normatividad interna inherente a las operaciones técnicas mismas. Esto es innovador, ya que históricamente ha prevalecido la noción instrumental de que las tecnologías son neutrales y que los criterios normativos dependen exclusivamente de las deliberaciones humanas externas a ellas. Pero ¿cómo se puede proponer una normatividad interna en los objetos técnicos si estos no involucran intenciones? ¿Cómo afirmar que las tecnologías pueden llevar a desarrollos éticos a partir de su propia evolución material?

La respuesta a estas cuestiones supone el giro ontológico y el giro material mencionados anteriormente. Desde un punto de vista ontológico relacional, las tecnologías son una concretización de la vida humana en todos sus aspectos, incluidos los éticos. Así, el sentido ético de las tecnologías puede concebirse parcialmente como una etapa en su propio desarrollo, y también como una forma de concretización social. Un aporte crucial a este giro hacia una normatividad interna en lo tecnológico proviene de Simondon (2017) y su concepto de “profundización tecnológica”.

Para Simondon, el progreso técnico puede desempeñar un papel redentor, donde el sentido ético de la técnica se convierte en una fase del propio desarrollo técnico, resultado de su profundización. Sin embargo, Simondon no consideró este sentido salvífico como una etapa inevitable del desarrollo técnico, sino como una posibilidad deseable en donde prevalece el trabajo del técnico para perfeccionar el funcionamiento de la tecnología, en consonancia o disonancia con las funciones sociales que a ella se le asignen. A través de esta profundización se puede generar un proceso dialéctico de recuperación positiva del pasado técnico. Este proceso de perfeccionamiento puede ofrecer alternativas frente a problemas éticos y sociales como la obsolescencia y el consumismo, y optimizar los beneficios frente a los posibles perjuicios derivados del avance técnico entendido como un fin en sí mismo. Un ejemplo muy familiar de esta profundización es lo que ha sucedido con las bombillas eléctricas. Con el tiempo, las bombillas de led han ido reemplazando las tradicionales, por su menor consumo energético y su mayor durabilidad.

En otras palabras, la profundización de la técnica podría ayudar a solucionar los problemas que se derivan de muchos usos de las tecnologías.

Por ejemplo, en el contexto del cambio climático los automóviles eléctricos y de hidrógeno podrían ayudar a hacer la transición hacia una movilidad menos contaminante que la de combustibles fósiles.

Ante tecnologías contaminantes, se puede optar por profundizar en el desarrollo de tecnologías que contaminen menos o incluso que descontaminen. Del mismo modo, frente a tecnologías que generan riesgos, es posible avanzar hacia el desarrollo de tecnologías que disminuyan esos riesgos, y que además aumenten la eficiencia y la durabilidad.

La perspectiva ética de Simondon cuestiona algunos puntos de vista ampliamente aceptados y socialmente internalizados como evidentes. Un ejemplo es la creencia de que existe una contradicción inevitable entre el progreso tecnológico y algunos valores fundamentales como la sostenibilidad, la privacidad, el bienestar social, etc. Desde esta perspectiva, se asume erróneamente que, para disfrutar del progreso tecnológico, debemos pagar un precio existencial, moral, social y ambiental. O que las tecnologías son un “mal necesario” que aceptamos a cambio de los beneficios que aportan. Esta idea conduce, para Simondon, a una cierta aversión o desconfianza hacia las máquinas y tecnologías. Frente a esta visión, la respuesta más frecuente es plantear la necesidad de un control externo sobre las tecnologías desde una perspectiva humanista que las trascienda, o desde un marco normativo que las regule.

En contraste con este punto de vista, Simondon (2017) propone asumir un enfoque ontológico relacional, que supere la falsa dualidad entre lo técnico y lo cultural, y que promueve una visión positiva, de aprecio o de amor hacia las máquinas, como la que tiene el técnico que busca de manera insaciable perfeccionar los procesos técnicos. Este autor aboga por una armonización entre un desarrollo técnico con un sentido moral a través de la profundización técnica, o un análisis ético interno, inherente y propio de lo técnico. Esto se opone a la mayoría de las perspectivas éticas que le asignan a la ética el papel de ofrecer un control externo a lo técnico (Moreno-Ortiz, 2022, pp. 159-185).

Estos puntos de vista son fundamentales para reconocer las posibilidades de normatividad inherente en las tecnologías, pero también tienen como riesgo su derivación en la posición del solucionismo tecnológico. Para Simondon, la profundización es algo deseable y posible en muchas circunstancias, pero no necesario y posible en todos los

casos. Ella depende también de otros procesos de individuación social. A pesar de su profundización, las tecnologías son también formas de concretización social, y, por tanto, requieren de procesos de orientación y regulación social más amplios, además de su profundización. Ello puede equilibrar los posibles riesgos hacia el solucionismo tecnológico. En la siguiente sección de este capítulo se incluyen dos casos de solucionismo tecnológico y su análisis, con el propósito de mostrar que no es adecuado absolutizar ningún punto de vista, ofrecer, manera crítica y plural, aportes para la comprensión de la complejidad del análisis ético.

Análisis del caso de la geoingeniería y el transhumanismo: dos caras del solucionismo tecnológico contemporáneo

En medio del desierto de Texas, estado petrolero por excelencia, se levanta desde hace un tiempo una gran mole de extraña arquitectura ubicada en terrenos pertenecientes a Oxy (Occidental Petroleum Corporation). Se trata de la estación DAC-1 (Direct Air Capture), costoso proyecto financiado por la corporación transnacional con el propósito de poner en marcha la recolección de miles de toneladas de CO₂ de la atmósfera mediante un proceso denominado secuestro de carbono. La idea es bastante simple en principio: consiste en aspirar el aire contaminado mediante enormes turbinas, retirar el dióxido de carbono y enterrarlo a cientos de metros bajo el suelo para devolver luego el aire limpio a la atmósfera. Se trata de un primer paso en un proyecto tecnológico de gran escala (se planea levantar muchas estaciones más como esta en varias partes del mundo) con el que la industria petrolera, una de las más contaminantes, pretende hacer su parte para luchar contra la crisis climática que ella misma ha contribuido a generar.

De otro lado, y a medida que crece la presión por frenar rápidamente el aumento de las temperaturas globales, se habla cada vez más de la posibilidad de enviar flotas de aviones a alturas muy elevadas para inyectar periódicamente azufre en la atmósfera terrestre, e incrementar de ese modo la nubosidad en algunas regiones del planeta, de suerte que

se fomenten las precipitaciones, cuya dramática ausencia, como resultado del calentamiento global, está ocasionando prolongadas sequías que amenazan la producción agrícola y el suministro de agua potable en muchas regiones. Y ante la urgencia no han faltado tampoco los visionarios que proponen filtrar el ingreso de rayos solares a la atmósfera terrestre gracias a enormes sombrillas que orbiten alrededor del planeta y que, a manera de pantallas solares, repelan hacia el espacio exterior el exceso de radiación que amenaza con derretir el hielo polar. El fin que persiguen estas y otras iniciativas semejantes, a primera vista sacadas de novelas de ciencia ficción, no es otro que disminuir la temperatura global, la cual está encaminada actualmente, según los últimos informes del *Panel Intergubernamental de Expertos del Cambio Climático* (IPCC, por su siglas en inglés), a sobrepasar el peligroso 1.5 °C con respecto al periodo preindustrial. A menos, claro está, que se consiga una reducción drástica e inmediata de la cantidad de emisión de gases de efecto invernadero. Aunque un grado y medio más de temperatura parece un incremento insignificante, todos los estudios climatológicos señalan que un aumento igual o mayor acarrearía nefastas consecuencias para la habitabilidad de la Tierra, no solo para la especie humana sino para la gran mayoría de especies de vivientes que la pueblan.

Las anteriores propuestas ilustran lo que se ha dado en llamar *geoingeniería* (Neyrat, 2019). Como el nombre sugiere, esta consiste en emplear las nuevas tecnologías de punta para concebir dispositivos capaces de hacer frente al cambio climático, haciendo de la Tierra, mejor aún, de la biosfera (esa delgada capa del planeta donde existe y se mantiene la vida), el objeto mayúsculo sobre el cual poner en marcha tales adelantos ingenieriles. Nunca antes experimentos como estos, a escala planetaria, habían sido propuestos, y nunca como ahora tales iniciativas, tan ambiciosas como riesgosas, habrían podido encontrar oídos receptivos. Hoy en día, sin embargo, la urgencia y la gravedad del problema climático que enfrentamos, así como la creciente confianza en las tecnologías humanas que progresan a un ritmo inusitado (Boström, 2014), llevan a que se contemple seriamente su uso para frenar la catástrofe.

Por otra parte, en Silicon Valley, afamada meca californiana de *startups* situadas a la vanguardia de la innovación tecnológica mundial,

se encuentra la sede del llamado *transhumanismo*. Este es el nombre que recibe un proyecto del que se ha venido sabiendo cada vez más gracias a informes en la red y a libros publicados por sus defensores e ideólogos, quienes divulgan así sus pretensiones casi mesiánicas de liberar a la humanidad de todos sus males mediante el uso combinado de la nanotecnología, la biotecnología y la tecnología informática y computacional (Ferry, 2016). Gracias, por ejemplo, a la implantación de dispositivos tecnológicos en el cuerpo humano, el transhumanismo busca dotarlo de facultades ilimitadas de percepción, memoria, cálculo, rendimiento, etc. El ideal al que apuntan es el cibernético, entidad híbrida medio orgánica, medio cibernética, que, dotada de lo mejor que ofrece la tecnología actual, llevaría al ser humano a un nivel inimaginable en todos los ámbitos de su desempeño físico y mental. Mediante implantes y prótesis perfectamente fusionados con el cuerpo y el cerebro humanos se aumentaría de tal manera su potencial cognitivo y físico que los transhumanistas no escatiman en hablar del futuro advenimiento de un humano mejorado o, incluso, de un poshumano (Diéguez, 2021). Con este objetivo en mente, el transhumanismo también apuesta por emplear técnicas de modificación genética, no solo con el fin terapéutico de curar enfermedades, sino ante todo con el propósito de prolongar indefinidamente la vida humana, de manera que la muerte sea finalmente vencida (Redeker, 2018). En efecto, la edición del genoma (llamado también edición de genes o edición genética —CRISPR—) es un grupo de tecnologías recientemente desarrolladas que permiten modificar el ADN de un organismo, haciendo posible agregar, quitar o alterar material genético en ciertos lugares específicos del genoma. Para el transhumanismo, esta es una puerta que se abre a una infinidad de mejoras posibles al vulnerable ser humano. El transhumanismo sueña así con tomar en sus manos el curso de la evolución humana, y no simplemente seguir padeciéndola de forma lenta y aleatoria como ha ocurrido hasta ahora. Y aunque parece que su proyecto no es sino producto de la imaginación afebrada de un puñado de multimillonarios, casi una visión más propia de películas de ciencia ficción que de la ciencia real, el flujo creciente de millones en recursos invertidos en estas iniciativas permite afirmar que la promesa del transhumanismo es tomada cada vez más en serio.

Cierto es que nada suena más atractivo que poder desterrar del horizonte enfermedades tan amargas como el Alzheimer, o raras condiciones genéticas que afectan duramente la vida de las personas ocasionándoles mucho dolor y sufrimiento. El problema, sin embargo, es que la línea entre un enfoque terapéutico y uno mejorativo se vuelve muy borroso en el transhumanismo, dejando planear la sospecha de que, tras un bienintencionado proyecto de mejora de la condición humana, se esconde o se promueve de alguna forma un aterrador proyecto eugenésico. Se trata de peligro ético de grandes proporciones que no puede ser dejado sin cuestionamiento. Así mismo, y por más que se plantee cada vez más como la mejor solución a la gravísima crisis climática en curso, la geoingeniería levanta mucha reserva incluso entre muchos científicos. Se alerta sobre los efectos nefastos e imprevisibles que puede ocasionar en el sistema Tierra el hecho de jugar a aprendices de brujo con el clima, pudiendo llegar a castigar aún más a los países y regiones más vulnerables por el calentamiento global. En fin, el anhelado remedio que prometen las nuevas tecnologías puede ser eventualmente mucho peor que el mal que se busca conjurar con ellas.

En efecto, tanto la geoingeniería como el transhumanismo son proyectos que parecen traer consigo una promesa de paraíso perdido que, no obstante, podría virar fácilmente hacia una pesadilla. La confianza es que, gracias a las proezas de la tecnología más avanzada, la humanidad podría resolver todos los problemas que la aquejan, pero esta es una promesa que despierta profundas inquietudes y reparos de orden ético y político. Quienes observan con recelo tales propuestas no dejan de señalar los peligros que suponen tales pretensiones “prometeicas”. Se emplea el término con el mismo tono de advertencia que se hizo desde muy temprano en la historia occidental, en particular entre los griegos, frente a la aspiración del ser humano de escapar a su condición natural mediante el uso desmedido de la técnica. Es la lección moral que suele leerse en el famoso mito de Prometeo, el ladrón del fuego de los dioses, cuyo destino trágico resuena en el adjetivo prometeico.

Ciertamente, si bien es gracias a la técnica que parece posible para los seres humanos escapar a un orden natural preestablecido y darse a sí mismos una segunda naturaleza, mejorada, segura, artificial, controlable y poderosa, el riesgo es, sin embargo, pecar por un exceso

de confianza que termine por pagar muy caro. Ahora bien, aunque el motivo de la transgresión de la naturaleza no estuvo ausente de la reflexión filosófica acerca de la técnica, al mismo tiempo ha existido una lectura mucho más positiva (como se ha visto en otros capítulos de este manual), en la medida en que la capacidad técnica —reforzada por el conocimiento científico— es concebida como *lo propio* del ser humano, esto es, la mejor expresión de sus capacidades únicas que le confieren un lugar aparte de las demás entidades naturales. Así pues, vista desde esta perspectiva, lejos de ser una amenaza para su humanidad, el desarrollo de la tecnología sería, por el contrario, la manifestación más clara de lo humano, es decir, la que humaniza y no la que deshumaniza al ser humano. Son dos caras de un mismo fenómeno que emergen con fuerza una vez más frente al transhumanismo y la geingeniería actuales.

Tal parece, empero, que en la actualidad se tienen pocos oídos para el tipo de advertencias sobre la *hybris* que el poder tecnológico puede despertar en el ser humano. Antes bien, hoy en día crece la confianza en que la tecnología es no solo la expresión más acabada del ingenio humano, hacia la cual no cabe más que albergar un sentimiento de orgullo y admiración, sino que además ofrece la mejor solución posible a cualquiera de los múltiples problemas que nos acechan. Esta actitud ha sido denominada “solucionismo tecnológico”. La expresión acuñada por Morozov (2015) se refiere a una actitud hoy en día dominante, por la cual se espera que la solución a cualquier problema humano provenga de un avance tecnológico. Como dice irónicamente Morozov: se cree que la salvación está “a un click de distancia”.

El solucionismo tecnológico sueña así con que la tecnología permitirá en un futuro próximo resolver desde problemas de salud, como epidemias y enfermedades, hasta problemas a gran escala como la corrupción o la criminalidad, en los ámbitos de la política, la cultura y la vida cotidiana. No habría, pues, problema alguno que se resistiera a ser resuelto de forma eficiente y original mediante las nuevas tecnologías que cada día nos sorprenden con su progreso. Para Morozov, la era tecnológica en la que vivimos tiende así a esperar que tarde o temprano se pueda arreglar todo lo que nos causa dificultad o preocupación en nuestra existencia, no mediante transformaciones sociales, educativas, políticas o culturales, sino,

ante todo, gracias a la digitalización masiva. Cada vez menos esperamos algo de nosotros mismos y más de la tecnología, sentencia por su parte la experta Sherry Turkle (Turkle, 2017). Pero ¿dónde quedan los dilemas políticos y morales que estas “soluciones” tecnológicas despiertan?

No hay certeza alguna de cómo nos afectará su uso. En cambio, sí hay muchas razones para llamar a la reflexión y a la prudencia, pues en el fondo se trata de algo tan delicado como el comportamiento y la motivación humanos en sus múltiples dimensiones. Sin caer en la demonización de las tecnologías, no hay duda de que esto amerita prestar atención suficiente a las consecuencias morales y de otro tipo que su uso intensivo pueda traer, al estar movido por la búsqueda obsesiva de eficiencia en todos los frentes de la experiencia humana. Semejante fe en la tecnología tiene a su base la convicción de que ella ha sido y es la que detenta la clave de un progreso humano sin límites. Lo paradójico del asunto es que ese mismo ideal de progreso tecno-científico afianzado desde la modernidad ha sido señalado como el origen de una de las problemáticas más graves que enfrentamos actualmente: el cambio climático. La impronta negativa de la acción humana en la faz de la Tierra ha llevado a hablar de que vivimos en la nueva “era del hombre” o Antropoceno. El cambio climático sería así el fenómeno antropocénico por excelencia, al delatar la poderosa —y nociva— influencia del ser humano sobre los ciclos de la naturaleza en su conjunto. Un cambio de roles que no tiene antecedentes en la historia, pues el poder tecnológico humano nunca llegó a la posición de poder competir o siquiera figurar entre las más poderosas fuerzas naturales.

Así pues, la geoingeniería, al igual que el transhumanismo, parecen descansar en la convicción de que la tecnología es el medio humano más poderoso, capaz de modificar e incluso de someter a su control a la naturaleza, por un lado, y a la naturaleza humana por otro lado (Habermas, 2012). De hecho, en ambos casos se da por obsoleta la idea misma de *naturaleza*. Si con este concepto se tiene en mente una realidad dada, autónoma en su funcionamiento, que determina y se resiste al poder e interés del ser humano, ambos proyectos tecnológicos apuntan al fin de esa idea de naturaleza. En ninguno de los dos casos se la ve más como una realidad cuyo poder o misterio exceda al ser humano y, por ende, algo frente a lo cual este deba mantener cierta reverencia o restricción

en su trato. La idea de naturaleza no sirve más para trazar un límite que lleve a cuestionar éticamente el uso de los avances técnicos. La inquietud acerca de si hay un exceso tecnológico que pueda violar algo esencial llamado naturaleza, o naturaleza humana, no está muy presente en la geoingeniería ni en el transhumanismo, para los cuales hablar de naturaleza en términos de una realidad establecida o definitiva es un rezago metafísico que carece de sustento ante el ilimitado poder de transformación adquirido por la tecnología humana.

Por lo anterior, además de la preocupación por la incógnita que suscita el hecho de no poder prever los efectos colaterales de esas intervenciones profundas en la naturaleza, se critica sobretodo que detrás de estas empresas prometeicas hay un modo de relacionarse tanto con la naturaleza, como con el propio cuerpo humano, como si se tratara de meros objetos (Neyrat, 2019). En virtud de esta concepción, estos se vuelven susceptibles de ser dominados, transformados y reconstruidos a imagen y semejanza de un cierto ideal en el que ya no es bienvenido hablar de finitud, vulnerabilidad, límites, en fin, de cualquier forma de negatividad o exterioridad ante la cual deba plegarse el ser humano. En otras palabras, son expresión de una visión antropológica que se centra en el poder del ser humano antes que en su fragilidad y dependencia.

Ciertamente, algunos de los reparos que desde el punto de vista ético se dirigen en contra del transhumanismo tienen que ver con la exaltación que hace de una supuesta perfección humana, entendida aquí en términos de un desempeño más y más eficiente. Pues no otro parece ser el contenido de un programa de mejoramiento que entiende tal mejora humana como incremento cuantitativo en la capacidad de los cuerpos humanos para desempeñar funciones muy complejas que empujan “hacia delante” los límites biológicos. No se promueve tanto una mejora como la preconizada por el humanismo tradicional, que apunta a conseguir avances de orden ético y político básicamente mediante la educación y otros esfuerzos formativos a través del tiempo, avances que no dependen de la puesta en marcha de dispositivos tecnológicos sino ante todo de la transmisión de valores e ideales. De manera similar, los críticos de la geoingeniería apuntan no solo a los dañinos efectos colaterales que pueden seguirse de experimentos de esta naturaleza a escala planetaria, sino que subrayan sobre todo que este tipo de “soluciones” tienden a ocultar la necesidad de

emprender cambios de orden político, económico y cultural que son imprescindibles para atacar las verdaderas causas del problema ecológico. Estos cambios son seguramente más lentos y dolorosos de lo que muchos quisieran asumir, pues apelan al abandono de ciertos modos de vida, valores e ideales ligados al ideal de progreso moderno.

En últimas, en el marco de estos dos proyectos puntuales que hemos discutido, la tecnología parece ser leída como el mejor y único medio con el que el ser humano cuenta en la actualidad para lidiar con las amenazas que lo acechan. Por esta razón, es investida con poderes que quizá exceden lo que es posible y deseable esperar de ella. No es asunto, sin embargo, de caer en una actitud simplemente ludita o tecnófoba. Pero lo que estos dos proyectos hiperbólicos llaman a pensar es si ciertos usos tecnológicos, o mejor, si una particular manera de entender la tecnología —que, como ocurre en nuestros días, tiende a verla como *la* solución para todo— no entraña un problema antropológico mayor. Hay, en efecto, detrás de la geoingeniería y del transhumanismo una comprensión que es a la vez exaltada y algo avergonzada del ser humano, por concebirse a sí mismo simultáneamente por fuera y por el encima de la naturaleza, al mismo tiempo que se reconoce dolorosa y trágicamente perteneciente a la naturaleza. Esta ambivalencia invita a una cuidadosa e interesante reflexión filosófica sobre quiénes somos y cómo nos define en cuanto humanos la relación que entablamos entre tecnología y naturaleza.

Hacia una ética constructiva de las tecnologías

Una ética contemporánea de las tecnologías debería caracterizarse por ser crítica, plural, empírica, material, relacional y profundamente integrada al propio desarrollo tecnológico. Estas características pueden combinarse con los aportes de las perspectivas éticas tradicionales, como ofrecer una visión crítica de la cultura tecnológica (sin caer en sustantivismos), proponer marcos normativos regulativos (evitando el instrumentalismo), y proporcionar mecanismos democráticos para la regulación social de las tecnologías (superando el voluntarismo y evitando el solucionismo tecnológico).

En general, una primera exigencia actual de las éticas de las tecnologías es la de ser críticas. Se podría afirmar que la diferencia entre una perspectiva crítica de la ética de las tecnologías, frente a una perspectiva no crítica tiene que ver inicialmente con la posibilidad de ofrecer aportes y soluciones, sin caer en los extremos del voluntarismo, el instrumentalismo, el determinismo y el sustantivismo.

No obstante, es esencial adoptar los giros recientes en la ética de la tecnología para abordar cuestiones como: ¿cómo enfrentar éticamente los desafíos tecnológicos actuales en contextos como el cambio climático, el Antropoceno, la sociedad de la información y la cultura algorítmica? O, más específicamente, ¿cómo reorientar éticamente los cambios tecnológicos actuales y emergentes para que respondan a estos desafíos, y no simplemente ponerles límites o restricciones? Estas preguntas son desafiantes porque requieren que las éticas de las tecnologías no solo sean capaces de evaluar, prescribir, definir límites, establecer controles, sino que también sean capaces de proponer alternativas, orientar o guiar los desarrollos más éticos, anticipar los posibles riesgos futuros. En general, estas funciones adicionales se pueden englobar alrededor del término “constructivo”. Es decir, que las éticas de las tecnologías puedan ser constructivas, además de evaluativas.

A lo largo de este texto, hemos visto que las perspectivas éticas tradicionales tienden a enfocarse en controlar los aspectos perjudiciales de los cambios tecnológicos, pero son muy débiles para ayudar a construir alternativas y orientar positivamente los desarrollos. Sin embargo, para estar a la altura de las demandas actuales, las éticas de las tecnologías deben ir más allá de una mera evaluación y contribuir a la construcción de cambios desde un enfoque propositivo, performativo y orientado a la solución de problemas, en el curso mismo de su desarrollo, implementación y uso. Y en el terreno (*in situ*) o contexto particular y no de manera abstracta o general (Ravelo Franco, 2023; Guzmán-Ortiz, 2022).

Algunos cambios tecnológicos avanzan tan rápido o tienen efectos incrementales y disruptivos tan poderosos en la sociedad que es ingenuo pensar en una ética limitada solo a la evaluación y al control. Es necesario ampliar los enfoques hacia las perspectivas éticas constructivas que ofrecen posibilidades de orientación.

En este texto se ha visto que ninguna corriente o perspectiva ética es completa ni ideal para regular y orientar los cambios tecnológicos. Para responder a la complejidad actual de estos cambios y todo lo que suponen, la mejor opción es combinar, articular los aportes de las distintas perspectivas de manera crítica y constructiva. Es decir, evitando derivar el análisis moral en las perspectivas radicales, y equilibrando el análisis moral entre los sentidos evaluativo, regulativo y de orientación y construcción de alternativas.

Preguntas

1. ¿Cómo se pueden caracterizar las distintas perspectivas éticas de las tecnologías?
2. ¿Qué diferencia una perspectiva crítica de la ética de las tecnologías frente a una perspectiva no crítica?
3. ¿Qué ventajas y desventajas tienen las perspectivas tradicionales de la ética de las tecnologías?
4. ¿Por qué es importante un punto de vista empírico de la ética de las tecnologías?
5. ¿Qué aportes realizan las perspectivas contemporáneas de la ética de las tecnologías para el análisis ético de las tecnologías?
6. ¿Cómo enfrentar éticamente los desafíos tecnológicos actuales en contextos como el cambio climático, el Antropoceno, la sociedad de la información y la cultura algorítmica?
7. ¿Cómo plantear opciones de profundización ética de las tecnologías sin caer en el solucionismo tecnológico?

8. ¿Qué significa un enfoque constructivo de la ética de las tecnologías?

Para saber más

Los siguientes textos ofrecen una aproximación general a la ética de la tecnología, muy útil para introducirse en este campo:

- Hansson, S. O. (Ed.). (2017). *The ethics of technology. Methods and approaches*. Rowman & Littlefield International, Ltd.
- Jasanoff, S. (2016). *The ethics of invention: Technology and the human future*. W. W. Norton & Company.
- Mitcham, C., & Waelbers, K. (2009). Technology and ethics: Overview. En J. K. B. Olsen, S. A. Pedersen & V. F. Hendricks (Eds.), *A companion to the philosophy of technology* (pp. 467–473). Blackwell.
- Moreno-Ortiz, J. C., Díaz Guzmán, O. D., Escobar Mejía, J. E. y Pérez Peña, N. C. (2024). *Ética de las tecnologías. Análisis crítico de perspectivas*. Ediciones USTA.
- Robson, G. J., & Tsou, J. Y. (Eds.) (2023). *Technology ethics. A philosophical introduction and readings*. Routledge, Taylor & Francis Group.
- Tavani, H. T. (2011). *Ethics and technology*. 3.^a ed. Wiley-Blackwell.
- Van de Poel, I., & Royakkers, L. (2023). *Ethics, technology and engineering. An introduction*. 2.^a Ed. Wiley-Blackwell.

Capítulo 4

Filosofía de la inteligencia artificial

ÁLVARO DAVID MONTERROZA RÍOS

JUAN CARLOS MORENO ORTIZ

JOHN JAIRO CARDOZO CARDONA

Introducción

La filosofía de la inteligencia artificial es una rama de la filosofía que explora la inteligencia artificial y sus implicaciones en áreas tales como el conocimiento, la comprensión de la inteligencia, la ética, la conciencia, la epistemología y el libre albedrío (McCarthy, 2006). Si bien la investigación en IA estuvo en sus inicios cercana a la lógica, la filosofía de las matemáticas y la filosofía de la mente, hoy se la considera una rama de la filosofía de la tecnología debido a que las IA se han popularizado fuera del ámbito académico al ser una tecnología cotidiana con implicaciones sociales, morales y políticas que trascienden el plano epistemológico.

De acuerdo con McCarthy (2006), uno de los pioneros de la IA, la investigación en inteligencia artificial ocupa un lugar único en el espectro de las ciencias por su estrecha conexión con la filosofía. A diferencia de otras disciplinas científicas, la IA comparte con la filosofía una serie de conceptos cruciales como la acción, la conciencia, la epistemología y hasta el libre albedrío. Este entrelazamiento no es casual, ya que ambos campos se ocupan de preguntas fundamentales sobre la capacidad de conocer, la naturaleza de la realidad y los límites del entendimiento humano y artificial (McCarthy, 2006).

En este ámbito, los filósofos no solo analizan estos conceptos compartidos, sino que también ofrecen reflexiones críticas sobre la coherencia de ciertas ideas y sobre qué se puede esperar razonablemente que logren las máquinas. De esta manera, la filosofía de la IA trata sobre la concepción de la tecnología misma, pero además proporciona una guía crítica para los practicantes de la IA sobre lo que pueden y no pueden hacer sus creaciones. Esto se ve reflejado en el desarrollo de sistemas que puedan realizar acciones, razonar, planificar y, en última instancia, operar dentro de un marco que mimetiza la cognición humana a un nivel filosófico. Según McCarthy, la investigación en IA ha enfatizado la formalización de las acciones disponibles en una situación dada y las consecuencias de cada acción. Este enfoque ha requerido tratar con aproximaciones simples a fenómenos complejos, un desafío no solo técnico sino también filosófico (McCarthy, 2006). Uno de los problemas clave, tanto para la IA como para la filosofía, es la comprensión del conocimiento y las habilidades del sentido común, es decir, cómo manejar situaciones en las que el conocimiento disponible es parcial, tanto en observación como en teoría, y se debe operar con conceptos mal definidos que, sin embargo, pueden ser precisos en contextos especializados. Así, la filosofía de la IA actúa como un puente entre la teorización abstracta y la aplicación tecnológica práctica, iluminando tanto los avances tecnológicos como los dilemas éticos y teóricos que estos conllevan (McCarthy, 2006).

Como cualquier rama de la filosofía, la filosofía de la IA aborda preguntas tan variadas como difíciles de responder a cabalidad. ¿Puede una máquina actuar de manera inteligente y resolver problemas de la misma manera que lo haría un ser humano mediante

el pensamiento? ¿Son la inteligencia humana y la inteligencia de las máquinas la misma cosa? ¿Es el cerebro humano esencialmente un ordenador? Más especulativamente, ¿puede una máquina tener una mente, estados mentales y conciencia de la misma forma que los seres humanos? ¿Puede experimentar sensaciones y emociones, es decir, poseer *qualia*? Estas cuestiones muestran los diversos intereses de los investigadores en IA, científicos cognitivos y filósofos, y las respuestas a estas preguntas dependen en gran medida de cómo se definan “inteligencia”, “conciencia” y qué “máquinas” están bajo discusión (Bringsjord y Govindarajulu, 2024).

En este capítulo se hace descripción y clasificación básica de las IA, en temas particulares como el procesamiento del lenguaje, el aprendizaje automático y las redes neuronales, para dar un análisis de las estructuras y procesos que subyacen a los sistemas de IA actuales, como los modelos de lenguaje a gran escala (LLM). También se menciona un recorrido histórico general desde los antecedentes de la Antigüedad hasta la consolidación contemporánea del campo de estudio. Además, se presentan algunos principios éticos clave para el desarrollo de la IA, tales como la transparencia, la justicia, la equidad, la responsabilidad y la privacidad. Estos principios se ilustran con ejemplos aplicados a contextos prácticos como el diagnóstico médico, la toma de decisiones crediticias y la moderación de contenidos en redes sociales. Finalmente, se presenta una discusión sobre los desafíos y riesgos sociales de la IA, incluyendo la desinformación, el sesgo, y la posible sobrevaloración de su impacto en la sociedad, lo que invita a una reflexión crítica sobre el papel de la IA en el futuro.

¿Qué es inteligencia artificial?

La inteligencia artificial (IA) es un campo de estudio y desarrollo que se enfoca en crear sistemas capaces de realizar tareas que normalmente requieren inteligencia humana. Estas tareas incluyen el procesamiento del lenguaje natural, la percepción sensorial, la toma de decisiones, el aprendizaje y la adaptación a diferentes entornos. La IA integra disciplinas como la ciencia de la computación, la psicología cognitiva, la

lógica y la robótica para desarrollar algoritmos y modelos que permiten que las máquinas razonen, aprendan de la experiencia, comprendan y respondan a lenguajes específicos, y tomen decisiones autónomas (Russel y Norvig, 2002). Aunque comúnmente se procede como si tuviéramos una comprensión firme y precisa de la naturaleza de la IA, la realidad es que definir esta disciplina a satisfacción de todas las partes interesadas, incluidos aquellos que trabajan dentro del campo, puede ser difícil. Las definiciones de IA varían ampliamente, pero por lo general se centran en la creación de sistemas que pueden realizar tareas que, si fueran efectuadas por humanos, requerirían inteligencia (Bringsjord y Govindarajulu, 2024).

Stuart Russell y Peter Norvig (Russell y Norvig, 2002), pioneros en el estudio formal de la IA, identifican cuatro tipos principales de sistemas de IA, cada uno basado en diferentes objetivos y metodologías (tabla 9).

Tabla 9. Posibles objetivos de la IA

	Basado en el humano	Racionalidad ideal
Basado en razonamiento	Sistemas que piensan como humanos.	Sistemas que piensan racionalmente.
Basado en comportamiento	Sistemas que actúan como humanos.	Sistemas que actúan racionalmente.

Fuente: basada en Russell y Norvig (2002).

Estas categorías ayudan a clarificar el amplio espectro de enfoques en el campo de la IA y proporcionan un marco para entender cómo los sistemas pueden ser diseñados para emular o mejorar ciertos aspectos de la inteligencia y el comportamiento humano:

- *Sistemas que piensan como humanos*: estos sistemas están diseñados para replicar los procesos de pensamiento humano. Utilizan modelos como las redes neuronales artificiales para imitar la manera en que los humanos toman decisiones, resuelven

problemas y aprenden. La idea es que estos sistemas puedan automatizar y ejecutar tareas que requieren capacidades cognitivas humanas, como el reconocimiento de patrones y el procesamiento del lenguaje natural.

- *Sistemas que actúan como humanos*: la meta de estos sistemas es imitar el comportamiento externo observable de los humanos. Un ejemplo claro son los robots diseñados para realizar tareas que los humanos realizan de manera natural y eficiente. Estos sistemas no solo intentan replicar acciones humanas básicas, sino que también buscan interactuar con el entorno de manera similar a como lo haría una persona, incluyendo la capacidad de adaptarse a nuevas situaciones.
- *Sistemas que piensan racionalmente*: estos sistemas se enfocan en emular el pensamiento racional y lógico del ser humano. A menudo se implementan a través de sistemas expertos que utilizan reglas lógicas para llegar a conclusiones o realizar inferencias. La idea es crear sistemas que puedan razonar a partir de información disponible y tomar decisiones lógicas, incluso en situaciones complejas o inciertas.
- *Sistemas que actúan racionalmente*: estos sistemas están orientados a realizar acciones de manera inteligente y racional. Los agentes inteligentes, por ejemplo, son diseñados para operar de manera autónoma en un entorno, tomando decisiones que maximizan sus posibilidades de éxito o la consecución de sus objetivos. Estos sistemas no necesariamente replican el pensamiento humano, pero operan de manera óptima dentro de los límites de lo que se considera un comportamiento racional y eficiente.

Cada uno de estos enfoques (Russell y Norvig, 2002) destaca diferentes aspectos de lo que podría ser la inteligencia, ya sea emulando directamente a los humanos o abstrayendo aspectos de la racionalidad en formas que puedan ser implementadas en sistemas computacionales. Estas categorías proporcionan una base para explorar diversas

aplicaciones de la IA, desde robots que realizan tareas domésticas hasta programas de computadora que resuelven problemas lógicos complejos o gestionan datos a gran escala.

Este amplio espectro de definiciones y enfoques muestra que la IA, como disciplina, es tan diversa y compleja como los sistemas que busca crear. No existe una definición concreta de IA sino un conjunto de teorías que buscan modelar, entender y replicar diversos aspectos de la inteligencia, tanto humana como idealizada.

Historia de las inteligencias artificiales

Los antecedentes de la IA se remontan a mitos, historias y conjeturas que exploran la creación de entidades con inteligencia o vida artificial, extendiéndose a través de diversas culturas y épocas. Desde los autómatas mecánicos en las leyendas griegas hasta las calculadoras mecánicas en el Renacimiento, las raíces de la IA se nutren de un terreno filosófico, matemático y técnico que prepara el terreno para el nacimiento formal del campo en el siglo xx (Mayor, 2018).

De acuerdo con Adrienne Mayor (2018), desde la Antigua Grecia, la mitología ya jugaba con la idea de seres artificiales. Historias como la de Hefesto, el dios del fuego y la forja, quien creó sirvientes mecánicos de metal, o Talos, un gigante de bronce que protegía Creta, son ejemplos tempranos de lo que podríamos considerar como prototipos de robots. Estos mitos no solo reflejan un interés temprano en automatizar tareas a través de agentes artificiales, sino que también plantean preguntas sobre la creación y la autonomía, temas que siguen siendo relevantes en las discusiones modernas de la IA (Mayor, 2018).

En 1315, Ramon Llull, en su obra *Ars magna*, planteó una de las primeras aproximaciones a la idea de que el razonamiento humano podría ser replicado de manera artificial. Su objetivo era desarrollar un método lógico universal para resolver disputas y acercarse a la verdad, mediante combinaciones de conceptos básicos representados simbólicamente. Llull diseñó una máquina lógica, conocida como “máquina lógica de Llull”, que combinaba estas ideas a través de diagramas y círculos concéntricos. Aunque su propuesta no alcanzó el desarrollo

técnico que la inteligencia artificial requiere hoy en día, sus ideas prefiguraron la lógica simbólica y la idea de la automatización del razonamiento, influyendo en el pensamiento posterior sobre las máquinas capaces de emular la cognición humana (Wikipedia, 2024).

Durante el Renacimiento, este interés se materializó en avances técnicos y científicos. Inventores como Leonardo da Vinci diseñaron y posiblemente construyeron autómatas mecánicos, como su famoso caballero robot, que podía mover sus brazos, la mandíbula y la visera de su casco a través de un sistema de cables y poleas. Aunque estas máquinas no eran “inteligentes” en el sentido moderno, su existencia demuestra el ingenio humano para crear máquinas que imiten las acciones humanas. La era de la Ilustración intensificó el interés en la automatización y la mecánica. Filósofos como René Descartes y Julien Offray de La Mettrie comenzaron a explorar ideas que desafiaban las nociones tradicionales de la conciencia y la mente, argumentando a veces que los humanos mismos podrían ser vistos como máquinas complejas. En su obra *El hombre máquina*, La Mettrie propuso que el cuerpo humano no es más que una máquina compleja, lo que sugiere que replicar la inteligencia humana podría llegar a ser posible a través de medios mecánicos (Mayor, 2018).

Al mismo tiempo, matemáticos y lógicos comenzaron a desarrollar los primeros conceptos que se convertirían en la base de la computación moderna. Gottfried Wilhelm Leibniz, con su cálculo “ratiocinator”, introdujo la idea de un nuevo tipo de matemática basada en la manipulación de símbolos, que podía realizar automáticamente razonamientos lógicos. Este fue un paso muy importante hacia el desarrollo de algoritmos computacionales, que son fundamentales en la programación de la IA (Mayor, 2018). En el siglo XIX, Charles Babbage y Ada Lovelace trabajaron en la idea de la máquina analítica, un diseño para el primer ordenador programable de la historia. Lovelace, en particular, notó que una máquina así podría no solo manipular números, sino también crear música o arte, siempre y cuando se le proporcionaran las instrucciones adecuadas. Este concepto de “programabilidad” es esencial para la IA, con el cual sugirió que en algún momento las máquinas podrían llevar a cabo tareas creativas o “inteligentes” (Isaacson, 2014). Mientras tanto, en la misma época, George Boole desarrollaba

un sistema algebraico que eventualmente llevaría su nombre: el álgebra booleana. Este sistema, que opera con valores binarios y un conjunto de operadores lógicos, se convertiría en la base de la arquitectura de los circuitos eléctricos y la computación digital, proporcionando el lenguaje fundamental con el que se codifican y procesan los algoritmos de IA (David, 2001; Isaacson, 2014). Hacia el final del siglo XIX y principios del XX, se realizaron otros desarrollos teóricos fundamentales. La formalización de la lógica matemática por parte de Alfred North Whitehead y Bertrand Russell en su obra *Principia mathematica* intentó establecer un fundamento firme para toda la matemática, un esfuerzo que influiría directamente en las teorías computacionales y lógicas que fundamentan la IA (Isaacson, 2014).

El término “inteligencia artificial” fue acuñado en 1956, en una conferencia patrocinada por DARPA en el Dartmouth College, un evento que es reconocido como el nacimiento oficial de la IA como campo de estudio formal (Moor, 2006). Antes de este evento, Alan Turing había introducido la idea de que las máquinas podrían no solo calcular, sino “pensar”, y propuesto lo que ahora conocemos como el *Test de Turing* para evaluar la capacidad de una máquina para imitar la inteligencia humana (Turing, 1950). Este test sigue siendo un criterio influyente para algunos en el campo, aunque el propio concepto de “inteligencia simulada” ha evolucionado significativamente. La conferencia de Dartmouth consolidó la IA como un campo de estudio formal y técnico. Proyectos como el *Logic Theorist* de Newell y Simon mostraron que las máquinas podían realizar tareas complejas como probar teoremas de cálculo proposicional, lo que en su momento fue considerado un logro relevante (Bringsjord y Govindarajulu, 2024).

En las décadas de 1960 y 1970, investigadores como Herbert Simon y Marvin Minsky estaban convencidos de que sus métodos eventualmente lograrían crear máquinas con inteligencia general. Simon predijo en 1965 que las máquinas podrían realizar cualquier trabajo humano en veinte años, y Minsky anticipó en 1967 que el problema de crear “inteligencia artificial” se resolvería sustancialmente en una generación (Crevier, 1993). Sin embargo, ambos subestimaron la dificultad del problema, lo que llevó a un recorte en la financiación por parte de los gobiernos de Estados Unidos y Reino Unido en 1974, después de las

críticas de Sir James Lighthill y la presión del Congreso estadounidense por financiar proyectos más productivos (Russell y Norvig, 2021). La situación dio lugar al primer “invierno de la IA”, un período caracterizado por la dificultad para obtener financiación para proyectos de IA. No obstante, la investigación en IA se revitalizó a principios de los años ochenta gracias al éxito comercial de los sistemas expertos, que simulaban el conocimiento y las habilidades analíticas de expertos humanos, y al mercado de IA que alcanzó más de mil millones de dólares en 1985. Sin embargo, el colapso del mercado de las máquinas Lisp en 1987 marcó el comienzo de un segundo y más duradero invierno de la IA (Russell y Norvig, 2021).

Durante los años ochenta, comenzaron a surgir dudas sobre si el enfoque simbólico, centrado en representar objetos mentales como planes y creencias, podría imitar todos los procesos de cognición humana. Esto llevó a algunos investigadores a explorar enfoques “subsimbólicos”, incluyendo el resurgimiento del “conexionismo” y la investigación en redes neuronales, liderada por figuras como Geoffrey Hinton. En 1990, Yann LeCun demostró con éxito que las redes neuronales convolucionales podían reconocer dígitos escritos a mano, lo cual marcó el comienzo de muchas aplicaciones exitosas de las redes neuronales (Russell y Norvig, 2021).

La reputación de la IA comenzó a restaurarse a finales de los 1990 y principios del siglo XXI, aprovechando métodos matemáticos formales y encontrando soluciones específicas a problemas concretos. Esta focalización “estrecha” y “formal” permitió a los investigadores producir resultados verificables y colaborar con otras disciplinas. Aunque en la década de 1990 las soluciones desarrolladas por investigadores de IA raramente se describían como “inteligencia artificial”, para el año 2000, estas soluciones eran ampliamente utilizadas. Sin embargo, algunos investigadores académicos estaban preocupados de que la IA no estaba persiguiendo su objetivo original de crear máquinas versátiles y completamente inteligentes. Alrededor de 2002, fundaron el subcampo de la inteligencia general artificial (AGI), que para la década de 2010 contaba con varias instituciones bien financiadas, reviviendo así el interés en alcanzar los ambiciosos objetivos iniciales de la IA (Wikipedia, 2024; Russell y Norvig, 2021).

El campo de la IA experimentó una transformación significativa a partir de 2012 con el dominio del aprendizaje profundo en las evaluaciones industriales. Esto llevó a su adopción generalizada en toda la disciplina. La exitosa implementación del aprendizaje profundo se basó en mejoras significativas en el *hardware*, como computadoras más rápidas, unidades de procesamiento gráfico y la expansión del uso de la computación en la nube, así como en el acceso a grandes volúmenes de datos. Además, la disponibilidad de conjuntos de datos curados, como ImageNet, fue fundamental para entrenar estos sistemas avanzados (Wikipedia, 2024).

En 2016, la introducción de la arquitectura Transformer marcó un hito significativo en el desarrollo de las tecnologías de inteligencia artificial, especialmente en el ámbito de los modelos generativos. Esta arquitectura, propuesta por Vaswani et al. (2017) en el paper “Attention is All You Need”, introdujo un mecanismo de atención que permitía a las redes de IA ponderar la importancia relativa de diferentes palabras en una oración, independientemente de su posición. Esto facilitó una mejora considerable en el procesamiento del lenguaje natural, la cual permitió que modelos basados en Transformer, como BERT o GPT, aprendieran contextos más complejos y generaran textos con una coherencia y una relevancia semántica mucho más sofisticadas que antes. Como resultado, los modelos Transformer se convirtieron en la base para desarrollos posteriores en IA generativa, incluyendo modelos avanzados como GPT-3 de OpenAI, que han revolucionado desde la traducción automática hasta la generación automática de texto, influenciando profundamente tanto la investigación académica como las aplicaciones comerciales en el campo de la inteligencia artificial (OpenAI, 2023).

Si bien la historia de las tecnologías no puede contarse como una historia de héroes y descubrimientos puntuales, la tabla 10 nos permite mostrar que el desarrollo de las inteligencias artificiales ha tenido una trayectoria de altos cambios y grandes periodos de poca innovación (inviernos de IA), pero que nos muestra una historia adaptativa en que los desarrollos han tenido un carácter interdisciplinar, tanto teórico como práctico, apoyados en hombros de gigantes.

Tabla 10. Hitos de la IA

Año	Hito
1315	Ramon Llull en su libro <i>Ars magna</i> propone que el razonamiento pudiera realizarse de manera artificial.
1840	Ada Lovelace prevé que las máquinas pueden realizar más que simples cálculos, idea inicial del <i>software</i> .
1936	Alan Turing diseña la “máquina universal”, demostrando la viabilidad de un dispositivo para realizar cómputos.
1943	McCulloch y Pitts presentan el primer modelo de neuronas artificiales.
1955	Simon, Newell y Shaw desarrollan el lenguaje de programación IPL-11.
1956	En la Conferencia de Dartmouth, se acuña formalmente el término “inteligencia artificial”.
1957	Se desarrolla el <i>General Problem Solver</i> (GPS), un sistema orientado a la resolución de problemas.
1958	John McCarthy desarrolla LISP, el primer lenguaje para procesamiento simbólico.
1959	Rosenblatt introduce el perceptrón.
1963	Quillian desarrolla las redes semánticas para la representación del conocimiento.
1964	Bertrand Raphael crea el sistema SIR, capaz de inferir conocimiento basado en información.
1965	DENDRAL, el primer sistema experto, asiste a químicos en el análisis de estructuras complejas.
1968	Seymour Papert y otros desarrollan el lenguaje de programación LOGO.
1969	Alan Kay desarrolla el lenguaje de programación Smalltalk.
1973	Alain Colmenauer y su equipo crean PROLOG, un lenguaje de programación lógico.
1974	Edward Shortliffe escribe sobre MYCIN, un sistema experto que asiste a médicos en diagnósticos.

Año	Hito
1986	McClelland y Rumelhart publican <i>Parallel Distributed Processing</i> sobre redes neuronales.
1997	Deep Blue vence al campeón de ajedrez Gari Kaspárov.
2011	IBM Watson gana una ronda de tres juegos seguidos en <i>Jeopardy!</i>
2017	DeepMind con AlphaGo derrota al campeón mundial de Go, Lee Sedol.
2017	Google inventa la arquitectura <i>Transformer</i> , que lleva al desarrollo de modelos como GPT.
2022	Se lanza ChatGPT, una IA generativa que responde preguntas y genera texto en múltiples idiomas.
2023	Fotos generadas por IA alcanzan un nivel de realismo capaz de confundirse con fotos reales.

Fuente: basada en Wikipedia (2024).

Inteligencia artificial “débil” vs. “fuerte”

En el desarrollo y reflexión sobre la inteligencia artificial suele existir un debate filosófico alrededor de dos conceptos clave: la IA “fuerte” y la IA “débil”. La distinción entre estas dos vertientes no solo refleja diferencias en los objetivos y metodologías, sino que también subraya un profundo cuestionamiento sobre lo que las máquinas pueden o no hacer.

La IA “fuerte” se propone crear verdaderas “personas artificiales”, es decir, máquinas que no solo funcionan como humanos, sino que también poseen todas las capacidades mentales humanas, incluyendo la conciencia fenomenológica. Este tipo de IA no solo aspira a emular la inteligencia humana, sino a encarnarla por completo, creando entidades que no solo actúen de manera inteligente, sino que también comprendan y experimenten el mundo de manera similar a como lo hacemos los humanos (Moravec, 1988; Vita-More, 2020).

En contraste, la IA “débil” se enfoca en diseñar máquinas que parezcan inteligentes al emular las capacidades de procesamiento de información de los seres humanos. Estos sistemas están diseñados para realizar tareas específicas que normalmente requieren inteligencia, como jugar ajedrez o traducir textos, pero sin intentar replicar la conciencia humana. La IA “débil” busca simular la apariencia de inteligencia sin necesariamente entender o experimentar conscientemente los procesos subyacentes.

Uno de los desafíos filosóficos más destacados contra la IA “fuerte” es el argumento de la “habitación china” de John Searle (1980). En este experimento mental, Searle imagina a una persona que, sin entender chino, utiliza un conjunto de instrucciones para manipular símbolos chinos de manera que parezca que comprende el idioma. A pesar de las apariencias externas, la persona dentro de la habitación no tiene ninguna comprensión real del chino. Este argumento plantea dudas sobre si una máquina que procesa símbolos de manera similar realmente “entiende” lo que está haciendo, o simplemente está manipulando símbolos sin conciencia (Searle, 1980).

A pesar de las críticas, muchos en la comunidad científica de IA consideran que estos desafíos filosóficos son meras curiosidades teóricas que eventualmente serán superadas por avances tecnológicos. Los defensores de la IA “fuerte” creen que, con el tiempo, será posible desarrollar máquinas que no solo actúen, sino que también experimenten y comprendan el mundo de manera similar a los humanos (Kurzweil, 1999, 2005).

Otro desafío interesante es el argumento hödeliano, que se basa en los teoremas de incompletitud de Gödel (1931) para argumentar que las máquinas nunca podrán alcanzar la inteligencia humana porque hay verdades que los humanos pueden comprender que no pueden ser formalizadas por sistemas computacionales. Este argumento, aunque no es universalmente aceptado, sigue generando debate sobre los límites de lo que la IA puede lograr (Bringsjord y Govindarajulu, 2024).

Los logros técnicos de los últimos años nos hace pensar que se está imponiendo la idea de que la IA debe pensarse como en una versión “débil”, no obstante, muchas ideas transhumanistas de la versión “fuerte” de la IA (Rivera-Novoa, 2024) no pueden descartarse y con

seguridad seguirán siendo planteadas para nuevas discusiones filosóficas (Rueda-Pardo y Monterroza, 2025).

Las inteligencias artificiales generativas

La inteligencia artificial generativa (GAI, por sus siglas en inglés) es una variante de la IA que puede generar una variedad de contenido comprensible por humanos, como texto en lenguaje natural, imágenes, audio, código y datos específicos de dominio. Un modelo de lenguaje a gran escala (LLM, por sus siglas en inglés) es un tipo especial de GAI que puede generar texto en lenguaje natural en respuesta a una indicación en lenguaje natural (OpenAI, 2023).

La convergencia de enormes cantidades de datos disponibles en Internet, la capacidad informática cada vez mayor y el descubrimiento de una arquitectura de Transformador eficiente en 2017 llevaron a capacidades revolucionarias de GAI en 2022 (Vaswani et al., 2017). Específicamente, ChatGPT, el chatbot de IA de OpenAI, sorprendió a gran parte de la humanidad con sus impresionantes capacidades para responder a indicaciones en lenguaje natural. Esta invención demostró que una red neuronal artificial (ANN, por sus siglas en inglés), con tantas conexiones como neuronas en el cerebro humano, puede generar lenguaje humano de manera sorprendentemente precisa.

De acuerdo con Bridgelall (2024), los académicos aún no comprenden cómo un sistema informático finito y no adaptativo puede capturar simultáneamente las complejidades de docenas de lenguajes humanos, además de la esencia de todo lo que puede conversar. Esta falta de comprensión ha llevado a conceptos erróneos sobre cómo un LLM realiza sus funciones. Este autor señala que una idea errónea común es que un LLM es simplemente un motor de autocompletado más capaz, como los que se encuentran en los teléfonos inteligentes y procesadores de texto que predicen la siguiente palabra. Los desarrolladores de OpenAI sostienen que, de alguna manera, el entrenamiento del modelo GAI lo llevó a descubrir fenómenos subyacentes que representan procesos cognitivos en el cerebro humano, lo que permite la generación coherente de lenguaje e incluso cierto nivel de razonamiento (Bridgelall, 2024). Los humanos utilizan

el lenguaje como medio para expresar pensamientos, ideas y conceptos complejos; gracias a este, pueden transmitir y recibir información, formular argumentos, extraer inferencias del contexto, comprender diferentes perspectivas y resolver diversos problemas. Así, el lenguaje ayuda a los humanos a adquirir conocimiento sobre el mundo, las normas culturales y las convenciones sociales. La capacidad humana de razonar ayuda en el aprendizaje de idiomas al entender reglas gramaticales, inferir el significado de palabras desconocidas y adaptar ese conocimiento a diferentes situaciones (Bridgelall, 2024). Ahora, las herramientas de GAI como ChatGPT y Gemini dan la impresión de que pueden hacer todo lo anterior, lo que ha generado mucha controversia sobre lo que ocurre dentro de sus redes neuronales (Floridi, 2023; Chomsky et al., 2023).

Los modelos de procesamiento de lenguaje natural

La inteligencia artificial ha transformado significativamente la forma en que interactuamos con las tecnologías digitales, especialmente a través del procesamiento del lenguaje natural (NLP) y los modelos de aprendizaje de lenguaje. Estas tecnologías se han integrado en productos cotidianos que muchos utilizamos sin siquiera pensar en la complejidad de su funcionamiento. Ejemplos de estas aplicaciones incluyen *apps* de traducción de idiomas, aplicaciones de transcripción de dictados, altavoces inteligentes y convertidores de texto a voz. Estas tecnologías utilizan redes neuronales artificiales (ANN) como uno de sus principales habilitadores, permitiendo funciones avanzadas como el reconocimiento facial en álbumes digitales, sistemas de seguridad biométricos y reconocimiento de música (Bridgelall, 2024).

Los modelos de lenguaje a gran escala (LLM), como la serie GPT, son ejemplos avanzados de cómo se aplican estas técnicas de NLP. Estos modelos se construyen sobre arquitecturas que integran elementos comunes de NLP descubiertos a lo largo del tiempo. Un GPT (*Generative Pre-trained Transformer*) es un modelo de lenguaje entrenado para generar texto coherente y contextualizado. Utiliza la arquitectura de red neuronal *Transformer* (Vaswani et al., 2017) y se entrena con grandes volúmenes de datos textuales, lo que le permite comprender y generar respuestas en lenguaje natural de manera autónoma (OpenAI, 2023).

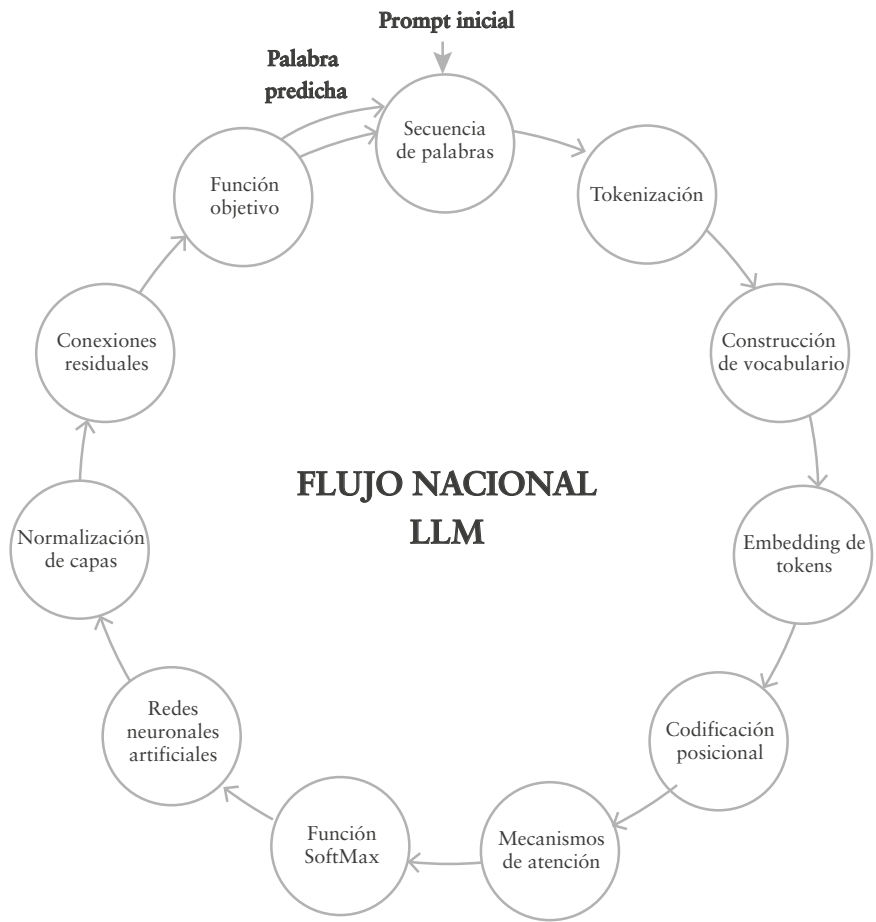


Figura 10. Flujo funcional de un LLM

Fuente: elaboración propia basada en Bridgelall (2024).

a) Fase de entrenamiento

La fase de entrenamiento es el proceso mediante el cual un modelo aprende a realizar una tarea específica utilizando datos de ejemplo. Durante esta fase, el modelo ajusta sus parámetros internos (como los pesos y los sesgos de las conexiones neuronales) para minimizar el error en sus predicciones. El proceso de entrenamiento presenta repetidamente al modelo enormes cantidades de ejemplos de secuencias de palabras para que pueda aprender de manera incremental a predecir

la siguiente palabra. El procedimiento de tokenización convierte la secuencia de palabras en símbolos que representan un lenguaje de manera más eficiente y ayuda a crear un vocabulario interno fijo para el procesamiento. Las incrustaciones de tokens y la codificación de posiciones convierten el vocabulario fijo en vectores de números de punto flotante que una computadora puede procesar. El mecanismo de atención produce vectores contextualizados que codifican la relevancia de otros tokens en la secuencia de entrada para establecer el contexto de la oración. La función SoftMax normaliza los vectores contextualizados producidos por el mecanismo de atención para que la ANN pueda descubrir y aprender patrones contextuales de manera más eficiente. La normalización y los residuales preparan los vectores de salida de la ANN para un procesamiento adicional sin perder información durante las transformaciones secuenciales. La función objetivo define las condiciones y restricciones estadísticas para predecir el siguiente token. En la tabla 11 se explica con más detalles cada una de las etapas del proceso funcional de un LLM.

Tabla 11. Explicación de las etapas de flujo funcional de un LLM

Fase	Descripción
Tokenización	Este proceso descompone textos en unidades más pequeñas llamadas tokens, que pueden ser palabras, partes de palabras o caracteres. En modelos avanzados como ChatGPT, se utilizan algoritmos específicos que optimizan la división de palabras en tokens reutilizables, como dividir “civilización” en “civil” y “ización”. Esto optimiza el reconocimiento de palabras complejas a partir de componentes más simples.
Construcción de vocabulario	El modelo utiliza técnicas como la codificación de pares de <i>bytes</i> (BPE) para crear un vocabulario fijo, generalmente alrededor de 50.000 tokens. Este proceso es esencial para gestionar la complejidad lingüística y permite al modelo procesar y generar texto de manera eficiente al reducir el número total de palabras únicas a un conjunto manejable de tokens frecuentes y combinaciones de caracteres.
<i>Embedding</i> de tokens	Una vez tokenizado el texto y definido el vocabulario, cada token se convierte en un vector numérico a través del <i>embedding</i> . Este proceso asigna a cada token un vector de dimensiones fijas (por ejemplo, 12 288 en algunos modelos), permitiendo que el modelo capture y aprenda relaciones complejas entre los tokens basadas en su uso y contexto en el lenguaje.

Fase	Descripción
Codificación posicional	La codificación posicional añade información sobre la posición de cada token dentro de la secuencia, lo cual es crucial para modelos como los Transformers que procesan los tokens de manera no secuencial. Esto permite que el modelo mantenga una comprensión del orden de las palabras, que es vital para entender estructuras gramaticales y sintácticas complejas.
Mecanismo de atención	Los modelos de lenguaje utilizan mecanismos de atención para ponderar la importancia relativa de diferentes partes de una secuencia de texto. Esto permite al modelo enfocarse en los tokens relevantes para un contexto específico, mejorando la capacidad del modelo para entender y generar respuestas adecuadas y contextualmente apropiadas.
Función SoftMax	Tras procesar la atención, la función SoftMax se utiliza para normalizar los vectores de atención y convertirlos en una distribución de probabilidades. Esto ayuda al modelo a determinar cuál es el token más probable que sigue en la secuencia. La función puede ajustarse con una “temperatura” para hacer que la selección de palabras sea más o menos aleatoria, influyendo así en la creatividad del texto generado.
Redes neuronales artificiales (ANN)	Las ANN son la estructura subyacente que aprende de los datos durante el entrenamiento. En modelos como GPT-3, que tienen miles de millones de parámetros, estas redes ajustan continuamente estos parámetros para optimizar la generación de texto. Cada capa de la red puede aprender diferentes aspectos del lenguaje, desde patrones simples hasta conceptos abstractos.
Conexiones residuales	En redes neuronales profundas, las conexiones residuales permiten que la información salte capas, evitando que se pierdan gradientes o se diluya la información a través de las capas. Esto es crucial para el entrenamiento eficaz de modelos profundos, permitiendo un aprendizaje más estable y rápido.
Normalización de capas	Este proceso estandariza las activaciones entre las capas de la red para que tengan una media y una varianza consistentes, lo que ayuda a estabilizar el aprendizaje y acelerar la convergencia durante el entrenamiento. Esto es especialmente importante en modelos grandes como GPT-3, donde la consistencia en las entradas de cada capa es vital para mantener un rendimiento efectivo.
Función objetivo	La función objetivo define los criterios bajo los cuales el modelo se entrena para predecir el siguiente token. Durante el entrenamiento, el modelo utiliza esta función para ajustar sus parámetros internos de manera que las predicciones sean cada vez más precisas, basándose en la diferencia entre la predicción del modelo y el token real que sigue. Esta es la base del aprendizaje supervisado en modelos de lenguaje.

Fuente: basada en Bridgelall (2024).

Durante el entrenamiento, se presenta una secuencia de palabras muestreadas aleatoriamente del corpus de entrenamiento, y se enmascara la próxima palabra mientras el modelo hace una predicción para maximizar su función objetiva. El algoritmo de entrenamiento calcula entonces un error entre la salida predicha y la correcta para determinar cómo ajustar incrementalmente los parámetros del modelo para reducir el error de manera gradual y continua. Este ciclo de predicciones y ajustes de parámetros continúa hasta el final de la oración, marcado por un token especial de fin de oración (EOS). El proceso converge cuando el error promedio alcanza un nivel predeterminado. Durante la inferencia, un disparador como una consulta inicia la predicción de palabras, una a la vez (OpenAI, 2023).

Es importante mencionar que el entrenamiento de este tipo de modelos se realiza mediante “aprendizaje supervisado”. Esto implica proporcionar a la red neuronal pares de entradas y salidas correctas (por ejemplo, preguntas y respuestas) y ajustar sus parámetros internos para minimizar el error en sus predicciones. Además, ChatGPT en particular utilizó “refuerzo por supervisión humana” para mejorar aún más su rendimiento. Esto implica que los humanos revisan y califican las respuestas del modelo, y estas calificaciones se utilizan para ajustar el modelo y mejorar su capacidad para generar respuestas coherentes y útiles (Gentile, 2023).

b) Fase de inferencia

La fase de inferencia es el proceso de utilizar el modelo entrenado para realizar predicciones en datos nuevos y desconocidos. Es la fase de consulta que le hace cualquier usuario. En esta fase, el modelo ya no está aprendiendo ni ajustando sus parámetros, en su lugar, se basa en los parámetros aprendidos durante la fase de entrenamiento para generar resultados (OpenAI, 2023).

El proceso comienza con la preparación de los datos de entrada, donde se recopilan y procesan los datos nuevos para hacerlos compatibles con el modelo, lo que puede incluir la tokenización del texto y la generación de vectores de características como *embeddings*. A continuación, se realiza la propagación hacia adelante, en la cual los datos preparados pasan a través de las capas de la red neuronal utilizando

los parámetros aprendidos durante el entrenamiento, generando una predicción. Luego, la salida del modelo se decodifica y se convierte en una forma comprensible, como una secuencia de palabras o asignaciones de clase para otras tareas, como la clasificación de imágenes. Después, se aplica un postprocesamiento para mejorar la calidad de las predicciones, en el que se corrigen errores o eliminan respuestas irrelevantes. Finalmente, los resultados generados se presentan al usuario o se emplean en otras aplicaciones o sistemas (OpenAI, 2023).

Es importante destacar que, en algunos casos, la fase de inferencia puede incluir un componente de búsqueda o de muestreo para generar múltiples posibles respuestas y seleccionar la mejor en función de una función de puntuación. En los modelos de lenguaje natural basados en Transformers, a menudo se realiza utilizando técnicas como la búsqueda por haz (*beam search*) o el muestreo top-k (OpenAI, 2023).

ChatGPT no piensa, calcula y también se equivoca

Sabemos que los modelos de lenguaje de gran escala (LLM) no son más que complejos programas de *software* que han sorprendido por su habilidad para discutir una amplia variedad de temas con una fluidez que a menudo se compara con la humana. Cualquier usuario puede experimentar esta capacidad simplemente interactuando con la interfaz web de OpenAI, donde el modelo puede generar textos que difícilmente se distinguen de los escritos por una persona nativa en inglés (Hicks Humphries y Slater, 2024).

La capacidad de GPT-4 para producir contenido variado y sofisticado ha llevado a muchos a creer que estos sistemas representan un avance significativo hacia una verdadera inteligencia artificial, similar a tener una mente humana operando dentro de un cerebro de silicio. Sin embargo, es crucial entender que los LLM operan de una manera fundamentalmente diferente a como lo hacen los cerebros humanos. Mientras que los humanos poseen y persiguen una multitud de objetivos complejos y conductas, muchos de los cuales son extralingüísticos (tales como deseos físicos básicos, objetivos sociales y relaciones personales, o la creación de objetos físicos), los LLM están diseñados con un solo propósito: replicar la capacidad de hablar o escribir de

los humanos (Hicks, Humphries y Slater, 2024). Lo que hacen los LLM es simular inteligencia con una serie de cálculos numéricos y probabilísticos de matrices y vectores de *embedding tokenizados* que se dan en la fase de entrenamiento y que se ponen en práctica en la fase de inferencia (Gentile, 2023).

Estos modelos logran su objetivo mediante la construcción de un modelo estadístico masivo que predice la probabilidad de secuencias de palabras basándose en vastos conjuntos de datos textuales, principalmente derivados de internet. Este proceso se realiza con poca intervención humana directa. En lugar de programar manualmente las funciones de probabilidad, los investigadores alimentan al sistema con grandes cantidades de texto y lo entrenan para predecir la siguiente palabra con base en estos datos. El modelo recibe retroalimentación, ajustando sus predicciones para mejorar su precisión con cada iteración.

En términos técnicos, el modelo asigna a cada palabra un vector dentro de un espacio abstracto de alta dimensión, agrupando palabras con usos o contextos similares y alejando aquellas con poco en común. Cuando genera texto, el modelo observa la secuencia de palabras previas y, utilizando estos vectores, calcula un nuevo conjunto de probabilidades para la siguiente palabra. Selecciona entre las más probables, aunque no necesariamente la más obvia, permitiendo variaciones que a menudo resultan en textos más creativos y humanos. Este proceso se ajusta mediante un parámetro conocido como “temperatura” del modelo, que afecta su creatividad (Álvarez-Osorio, 2026) y la propensión a generar errores o “falsedades” (OpenAI, 2023).

Dicho esto, no es sorprendente que los LLM tengan problemas con la precisión. Están diseñados para generar respuestas que parezcan coherentes y adecuadas a la consulta, pero no necesariamente para ofrecer información veraz o útil. La naturaleza de su entrenamiento los lleva a “alucinar” datos, citas o referencias que, aunque suenen convincentes, pueden ser completamente falsas. Esta tendencia a generar información incorrecta pero plausible ha llevado a lo que se denomina el “problema de la alucinación de IA” (Hicks, Humphries y Slater, 2024).

Según una crítica de Hicks et al. (2024), el problema aquí no es que los modelos de lenguaje grandes alucinen, mientan o tergiversen el mundo de alguna manera. Es que no están diseñados para representar

el mundo en absoluto; en cambio, están diseñados para transmitir líneas convincentes de texto. Por lo tanto, cuando se les proporciona una base de datos de algún tipo, usan esto, de una forma u otra, para hacer que sus respuestas sean más convincentes. Pero no están intentando de ninguna manera real transmitir o transmitir la información en la base de datos. En la medida en que a veces obtienen la respuesta correcta a tales preguntas es solo porque lograron sintetizar cadenas relevantes de lo que estaba en sus datos de entrenamiento. De manera similar, los modelos de lenguaje son propensos a inventar cosas porque no están diseñados para expresar algún conjunto subyacente de información en lenguaje natural; solo están manipulando la forma del lenguaje (Hicks et al., 2024).

Principios éticos para la regulación de la IA

Un principio ético es un fundamento normativo que sirve de guía o regla para la conducta moral. Los principios éticos actúan como normas generales o universales desde las cuales se derivan juicios morales más específicos, y su propósito es orientar la acción humana hacia lo que se considera correcto o bueno. Estos principios suelen estar enraizados en concepciones filosóficas de lo que es la justicia, el bien, el deber o la virtud, y constituyen los cimientos de sistemas morales en diversas tradiciones filosóficas (Ferrater Mora, 1978).

Para enfrentar los posibles riesgos y problemas del desarrollo de la inteligencia artificial se han propuesto desde el inicio de esta tecnología un conjunto de principios éticos que se explican a continuación.

El conocido escritor de ciencia ficción Isaac Asimov (1950) fue el principal pionero de la ética de la IA y de la robótica, con la formulación de las famosas tres leyes de la robótica, señaladas a continuación:

Primera ley: un robot no puede dañar a un ser humano o, por inacción, permitir que un ser humano sufra daño.

Segunda ley: un robot debe obedecer las órdenes dadas por los seres humanos, excepto cuando dichas órdenes entren en conflicto con la primera ley.

Tercera ley: un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o la segunda ley.

Posteriormente, Asimov añadió como “ley cero” la siguiente:

Ley cero: un robot no puede dañar a la humanidad, o, por inacción, permitir que la humanidad sufra daño.

Estas leyes definieron los límites básicos de los sistemas inteligentes artificiales, de manera anticipatoria, previendo un nivel de autonomía funcional de estos sistemas. Sin embargo, son muy generales, se limitan a lo mínimo, y los sistemas inteligentes artificiales actuales con los que interactuamos no operan con ese tipo de autonomía cuasi humana. Sin embargo, los usos que se están desarrollando con los sistemas de IA actuales pueden suponer importantes riesgos, que buscan regularse normativamente a partir de un conjunto de principios y subprincipios que se presentarán a continuación. Estos principios han servido de ejes alrededor de los cuales se han ido rápidamente articulando los distintos marcos éticos de la mayoría de los países y organizaciones, para regular los usos de la IA en distintos sectores. Una exposición detallada de estos principios se puede consultar en los siguientes textos: Russell y Norvig (2021); Floridi y Cowls (2019), y European Commission (2019).

Tabla 12. Principios éticos aplicables a la IA

Principio	Descripción	Ejemplo
Transparencia	Implica explicabilidad y auditoría de los sistemas de IA para garantizar confianza y comprensión.	Sistemas de IA en diagnósticos médicos que explican cómo se llegó a las conclusiones.
Justicia y equidad	Busca evitar discriminación y promover el acceso equitativo a los beneficios de la IA.	Algoritmos de reclutamiento que evitan sesgos en género, raza, o edad.

Principio	Descripción	Ejemplo
Responsabilidad	Involucra la rendición de cuentas y supervisión continua de los sistemas de IA por parte de sus desarrolladores y usuarios.	Empresas de vehículos autónomos que supervisan el rendimiento del sistema y ofrecen compensaciones por fallos.
Seguridad y privacidad	Asegura el uso ético y seguro de los datos, protegiendo la privacidad de los individuos.	Sistemas de reconocimiento facial que encriptan datos biométricos y respetan el derecho a la privacidad.
Beneficencia	Promueve el uso de la IA para generar beneficios positivos para la sociedad y el medio ambiente.	IA que ayuda a detectar epidemias analizando datos de salud pública en tiempo real.
Autonomía	Garantiza el respeto a las decisiones informadas y autónomas de las personas en su interacción con los sistemas de IA.	Plataformas de aprendizaje <i>online</i> que permiten a los usuarios personalizar sus propios objetivos y rutas de aprendizaje.

Fuente: basada en Russell y Norvig (2021), Floridi y Cowls (2019) y European Commission (2019).

a) Principio es de transparencia: este principio se precisa en los sub-principios de explicabilidad y de auditoría. Como su nombre lo indica, la explicabilidad tiene que ver con el acceso que las partes interesadas tienen a las explicaciones de las decisiones (opciones) y acciones definidas por los sistemas de IA, que facilitan la confianza en el sistema. También con las posibilidades de comprensión de las funciones del sistema. Y la auditoría se relaciona con la continua observación y seguimiento al funcionamiento del sistema, en especial, el monitoreo abierto de los posibles errores y sesgos.

La aplicación de estos principios se puede ilustrar con los siguientes ejemplos:

Ejemplo del uso de algoritmos de decisión crediticia en el siguiente contexto: una entidad financiera utiliza un sistema de IA para evaluar la solvencia crediticia de los solicitantes de préstamos.

Una aplicación básica del principio de auditoria podría ser la siguiente: la empresa documenta y publica los procesos lógicos detrás de los modelos utilizados para determinar la aprobación o rechazo de créditos. Esto incluye explicar qué variables (ingresos, historial crediticio, etc.) son consideradas y cómo influyen en la decisión final. Además, la empresa permite que auditores externos revisen el sistema para garantizar que no haya sesgos o discriminación implícita.

Ejemplo de uso de la IA para el diagnóstico de enfermedades, en el siguiente contexto: un hospital implementa un sistema de IA que ayuda a los médicos a diagnosticar enfermedades a partir de imágenes médicas, como radiografías o resonancias magnéticas.

El hospital aplica el principio de explicabilidad de esta manera: el sistema no solo proporciona una predicción o diagnóstico, sino que también explica los procesos que conducen a las conclusiones, mostrando las áreas específicas de la imagen que llevaron al diagnóstico y comparándolas con patrones reconocidos de enfermedades. Esto permite que los médicos entiendan y confíen en las recomendaciones del sistema y las expliquen a los pacientes. Este principio disminuye el efecto caja negra en el uso de la IA.

b) Principios de justicia y equidad: estos principios se pueden desglosar en los siguientes aspectos:

No discriminación: el uso de la IA no debe perpetuar ni amplificar los sesgos o discriminaciones existentes. Para ello es fundamental que los algoritmos sean diseñados y entrenados para ser justos y equitativos, evitando el trato desigual basado en factores como la raza, género, edad, etc.

Acceso equitativo: busca que el desarrollo y los beneficios de la IA sean accesibles para todos, no solo para grupos selectos o privilegiados.

La aplicación de estos principios se puede ilustrar con los siguientes ejemplos:

Ejemplo sobre la equidad y la no discriminación en el reclutamiento y selección de personal, en siguiente contexto: una empresa usa IA para filtrar currículums y realizar evaluaciones iniciales de candidatos a empleo.

La empresa aplica estos principios de la siguiente manera: el sistema de IA se diseña para evaluar a los candidatos únicamente con base

en habilidades y experiencia relevantes, sin considerar factores como género, etnia, o edad. Para asegurar esto, se entrena al modelo con un conjunto de datos diverso y representativo. Además, la empresa realiza auditorías periódicas para detectar y corregir cualquier posible sesgo que pueda afectar la equidad en las contrataciones.

Ejemplo sobre la aplicación del principio de justicia en la asignación de recursos públicos, en el siguiente contexto: un gobierno utiliza IA para asignar recursos y fondos de ayuda social a comunidades necesitadas.

El gobierno aplica el principio de justicia de la siguiente manera: el algoritmo se diseña para priorizar a las comunidades más vulnerables, a partir de las variables que definen la vulnerabilidad en el contexto, asegurando que los recursos lleguen a aquellos que más los necesitan, sin discriminación basada en factores como la ubicación geográfica o la etnia. El sistema también incluye mecanismos para que las comunidades puedan conocer las variables desde las que se hace la selección, y puedan cuestionar y apelar las decisiones si consideran que han sido tratadas de manera injusta.

c) Principio de responsabilidad: este principio se relaciona con el deber de los desarrolladores, los implementadores y los usuarios de los sistemas de IA de dar cuenta de las acciones, las decisiones y las consecuencias derivadas del uso e implementación de esta tecnología. Esta responsabilidad recae siempre en los seres humanos que se encuentran detrás de los sistemas de IA, pero se extiende, inscribe o materializa de manera específica en los sistemas de IA. Ella supone los siguientes aspectos:

Prevención de daños, supervisión, rendición de cuentas, compensación y reparación: quienes desarrollan, implementan y usan la IA deben realizar los ensayos y simulaciones necesarias para prevenir los posibles daños, antes de implementar la tecnología. Luego, en la implementación deben realizar una continua supervisión para detectar situaciones imprevistas. Así mismo, de manera periódica deben asumir los controles y supervisiones de las autoridades correspondientes y responder legalmente por los daños causados, con posibles compensaciones y reparaciones, generados especialmente por decisiones automatizadas, y con reparaciones de los sistemas.

A su vez, el buen desarrollo de los anteriores aspectos depende de la gobernanza que se establezca en el desarrollo, implementación y uso de los sistemas de IA, o del control social de las partes involucradas.

Ejemplo de desarrollo de una IA para el funcionamiento de vehículos autónomos en el siguiente contexto: una empresa desarrolla y entrena un conjunto de algoritmos para la conducción autónoma de vehículos.

La empresa aplica el principio de responsabilidad implementando protocolos estrictos para probar exhaustivamente el sistema antes de su lanzamiento comercial, incluyendo simulaciones y pruebas en condiciones controladas. Además, se establece un equipo de monitoreo que sigue de cerca el rendimiento del sistema en tiempo real una vez que está en uso, y se compromete a retirar temporalmente los vehículos si se detectan problemas de seguridad. La empresa también ofrece un plan de compensación para cualquier daño causado por fallos del sistema.

Ejemplo de responsabilidad en la moderación de contenidos en redes sociales, en el siguiente contexto: una plataforma de redes sociales utiliza IA para moderar contenidos, identificando y eliminando publicaciones que violan sus políticas.

La empresa aplica el principio de responsabilidad asegurando la supervisión humana continua de que todas las decisiones automatizadas de moderación de contenidos. También asegurando que los usuarios puedan apelar las decisiones tomadas por la IA, y que un equipo de moderadores revise manualmente estos casos para garantizar que las decisiones sean justas y correctas. La empresa también busca ser transparente sobre los errores cometidos por el sistema de IA y tomar medidas para corregirlos y compensar los posibles daños.

d) Principio de seguridad y privacidad: tiene que ver con la necesidad de garantizar el buen uso y la protección de los datos o de la información asociados a los sistemas de IA, para garantizar que los derechos fundamentales y la seguridad de los individuos, las organizaciones y las comunidades no se vean comprometidas. Tiene que ver con los siguientes aspectos:

Seguridad: la IA debe diseñarse e implementarse de manera robusta y segura, para minimizar los riesgos de ciberataques, el mal uso

o los fallos que podrían comprometer la privacidad y el buen uso de la información.

Privacidad: implica realizar la recopilación y el uso de datos de manera ética y con el consentimiento informado de los individuos. También recopilar y utilizar los datos estrictamente necesarios para el propósito del sistema, usar técnicas de anonimización, y respetar las normas de protección de datos. También supone el control de los usuarios sobre sus datos, incluyendo el derecho a saber qué información se recopila, cómo se usa, y la capacidad de corregir, eliminar o restringir el acceso a sus datos personales.

Ejemplo de seguridad y privacidad en el reconocimiento facial para la seguridad pública en el siguiente contexto: una ciudad implementa un sistema de IA para el reconocimiento facial en cámaras de seguridad con el fin de identificar criminales conocidos en espacios públicos.

En este caso se protege la privacidad de los ciudadanos, si se encriptan todas las imágenes y datos biométricos recopilados, garantizando que solo las autoridades autorizadas puedan acceder a ellos. Además, se establecen límites estrictos sobre cuánto tiempo se almacenan los datos y se prohíbe el uso del sistema para el monitoreo masivo o la vigilancia sin causa justificada. Además, se aplica estos principios si los ciudadanos tienen el derecho a saber si sus datos han sido procesados y a solicitar su eliminación.

Ejemplo de seguridad y privacidad en un sistema de monitoreo y seguridad inteligente para el hogar, en el siguiente contexto: una empresa comercializa un sistema de seguridad para el hogar basado en IA, que monitorea las entradas y salidas del hogar, detectando actividad inusual y alertando a los propietarios. También monitorea actividades de riesgo para los miembros del hogar.

La empresa aplica el principio en la medida que el sistema almacena los datos de monitoreo encriptados y utiliza autenticación multifactor para acceder a la información o realizar cambios en la configuración. Además, configura el sistema para no almacenar imágenes de video o datos de sensores de manera permanente, solo reteniéndolos temporalmente para evitar riesgos innecesarios de la privacidad. Así mismo, asegura a los propietarios la posibilidad de revisar y eliminar registros

específicos, según sea necesario. También declara el tipo de privacidad de los datos generales (no específicos) recopilados, su no comercialización.

e) **Principio de beneficencia:** se refiere al compromiso de utilizar los sistemas de IA de manera que genere beneficios positivos para la sociedad, los individuos y el medio ambiente. Supone que el sentido ético de la IA tiene que ver no solo con evitar daños, sino también con las posibilidades para mejorar la calidad de vida, promover el bienestar y contribuir al bien común.

En un sentido mínimo, la beneficencia supone tomar medidas para evitar daños, especialmente posibles efectos secundarios o impactos negativos no intencionados, pero en un sentido más amplio y positivo este principio tiene que ver con los impactos positivos que esta tecnología puede ofrecer para solucionar problemas, por ejemplo, en relación con la calidad de vida, la reducción de las desigualdades, la eficiencia, la equidad, etc. La aplicación de este principio se puede reconocer en los siguientes ejemplos.

Ejemplo de uso de una IA para la detección de epidemias, en el campo de la salud pública, en el siguiente contexto: una organización de salud global utiliza IA para monitorear y analizar datos de salud pública en tiempo real con el fin de detectar brotes de enfermedades infecciosas.

La organización aplica este principio al analizar con una IA los datos de múltiples fuentes, como informes de hospitales, redes sociales y registros de salud, para identificar patrones que podrían indicar un brote inminente. Esto permite a las autoridades de salud tomar medidas preventivas rápidamente, como la implementación de cuarentenas o la distribución de vacunas, salvando vidas y previniendo la propagación de la posible epidemia.

Ejemplo de uso de una IA para optimizar la producción agrícola de manera sostenible, en el siguiente contexto: una empresa agrícola utiliza IA para optimizar el uso de recursos en la producción de cultivos, como agua, fertilizantes y pesticidas.

La empresa aplica este principio analizando con una IA los datos de sensores en los campos para determinar las necesidades exactas de agua y nutrientes de las plantas, permitiendo un riego y fertilización precisos. Además de mejorar la producción de los cultivos, el sistema

ayuda a reducir el uso excesivo de recursos y minimizar el impacto ambiental, promoviendo una agricultura más sostenible y eficiente.

El espacio mundial más importante donde se promueve actualmente este principio es el programa y la cumbre mundial denominada, “IA for Good”, convocada todos los años por las Naciones Unidas. (<https://aiforgood.itu.int/#/es>).

f) Principio de autonomía: se refiere al respeto por la capacidad de las personas para tomar decisiones informadas y autónomas sobre su propia vida y sobre las formas como desean interactuar con los sistemas de IA. A través de este principio, se busca garantizar que los individuos mantengan el control sobre las decisiones que les afectan y que no sean manipulados o coaccionados por sistemas de IA.

Este principio se desglosa en aspectos como el consentimiento informado, la no coerción, el control sobre los datos personales, el derecho a la autoexclusión. Este principio tiene una relevancia especial en algunas decisiones finales en contextos críticos, por ejemplo, en la atención médica o en cuestiones judiciales, que deben estar bajo el control de humanos y no ser delegadas completamente a una IA. Algunas aplicaciones de este principio se pueden identificar en los siguientes casos:

Ejemplo de usos de una IA en una plataforma de aprendizaje *online* para la educación personalizada en el siguiente contexto: una plataforma de aprendizaje en línea utiliza IA para personalizar los materiales educativos y las rutas de aprendizaje de los estudiantes.

Se podría hablar de una aplicación del principio de autonomía en el caso de una plataforma que permite que los estudiantes puedan elegir sus propios objetivos de aprendizaje y seleccionar el tipo de contenido que prefieren (videos, lecturas, ejercicios interactivos). Aunque la IA sugiera un plan de estudio basado en el progreso del estudiante, los estudiantes pueden tener la libertad de seguir su propio ritmo, repetir lecciones o explorar otros temas según sus intereses.

Ejemplo de uso de una IA para hacer recomendaciones de compras en el comercio electrónico, en el siguiente contexto: un sitio web de comercio electrónico utiliza IA para recomendar productos a los usuarios en función de sus comportamientos de compra anteriores.

El sitio web busca aplicar este principio en la medida que los usuarios puedan optar por recibir recomendaciones o desactivar la función

de personalización de la IA si prefieren explorar productos por su cuenta. También en la medida que tengan la capacidad de modificar sus preferencias y ver cómo sus datos están siendo utilizados para generar estas recomendaciones.

En términos generales, es preciso tener en cuenta las siguientes consideraciones para una buena comprensión y aplicación de estos principios éticos:

- Son principios regulativos y no normas sujetas a un cumplimiento explícito. Es decir, se plantean en un nivel general superior para orientar éticamente las normas inspiradas en ellos en un nivel concreto inferior, sobre los desarrollos, implementaciones y usos de la IA. Estos principios se satisfacen y aplican a través de normas concretas, que deben cumplirse, para evitar las sanciones.
- Son relacionales. Esto significa que el sentido y el cumplimiento de uno usualmente depende del cumplimiento de otros.
- Estos principios y las normas que los concretan son recursos fundamentales especialmente para desarrollar los sentidos prescriptivo, normativo y evaluativo de la ética de la IA, pero son débiles para otros sentidos de la ética como el constructivo, el orientador, el anticipatorio.

Así, los principios éticos en la inteligencia artificial deben entenderse como guías regulativas de alto nivel, diseñadas para orientar la creación de normas concretas que garanticen un uso ético de la IA. Aunque no son directamente aplicables como normas vinculantes, estas se derivan de los principios y deben ser cumplidas para evitar sanciones. Además, estos principios son interdependientes, lo que implica que su correcta aplicación depende del cumplimiento conjunto de todos ellos. Si bien estos principios son esenciales para las dimensiones prescriptiva, normativa y evaluativa de la ética de la IA, son menos eficaces en aspectos como la orientación, anticipación o construcción ética.

La ética de la IA desde un punto de vista más amplio que el normativo

En el mundo actual, la IA generativa ha causado furor por varias razones: 1) porque simula muy bien nuestras capacidades lingüísticas, argumentativas y cognitivas; 2) por el acceso abierto, público y global que tenemos a ella; 3) por su capacidad para organizar, articular y sistematizar información de manera fácil, como si fuera el gran erudito al que podemos preguntarle todo, desde el tema académico más especializado hasta el mejor restaurante de piza en la ciudad; y 4) por la inmensa gama de ámbitos en los que puede utilizarse. Por todas estas razones, es hoy una de las tecnologías más disruptivas.

Las posibilidades generativas del lenguaje ofrecidas por esta tecnología pueden implicar varios problemas morales, como el suministro de información falsa inventada por el sistema, los usos para generar noticias falsas, los usos malintencionados para hacer suplantación de identidades, los usos indebidos de los datos y de la información, el plagio o infracciones de derechos de autor, etc. Desde un punto de vista más amplio, estos desarrollos alimentan el temor en relación con su capacidad para suplantar a los seres humanos en numerosos campos y tareas, incluso en muchas de nuestras actividades cognitivas y creativas.

En términos generales, el manejo de los posibles problemas éticos de los distintos sistemas inteligentes artificiales difícilmente puede realizarse desde una sola perspectiva, por lo cual es necesaria la integración de varias perspectivas. Específicamente, se requieren marcos normativos concretos, precisos, que pongan límites y regulen sus usos, como los que plantean las Naciones Unidas, los distintos estados, la Unión Europea, etc., de acuerdo con algunos principios y valores éticos, como la transparencia, la explicabilidad, la seguridad, la privacidad y la protección de los datos, la justicia, la equidad, la responsabilidad y la rendición de cuentas.

Es necesario realizar lo que se denomina actualmente una “alineación” a valores de cada uno de los sistemas inteligentes artificiales. Incluso, como sucede con las biotecnologías, también se requieren normas que prohíban o restrinjan *a priori* algunos desarrollos y usos posibles, por sus altos riesgos, como el empleo de drones autónomos

capaces de definir a qué personas asesinar en las guerras, de manera similar, en sus proporciones, al *Terminator* de las películas de la ciencia ficción. Desde el punto de vista normativo, sería ideal poder “alinear” estas tecnologías a los valores señalados, desde sus etapas iniciales de diseño y desarrollo, hasta sus etapas de implementación, involucrando a los distintos actores.

También se requieren análisis y críticas generales frente a los posibles riesgos e impactos sociales y culturales de estas tecnologías, y a las tendencias poco éticas del mercado, que pueden afectar sensiblemente la vida política, el desarrollo del conocimiento, la socialización, etc. Este tipo de posiciones críticas generales provienen de los estudios de riesgos y de impactos, de las iniciativas sociales o gubernamentales, y de los enfoques y reflexiones humanistas interpretativas amplias. Por ejemplo, este tipo de reflexiones se expresó en la declaración que realizaron más de mil expertos en la IA y ejecutivos de empresas tecnológicas, para realizar una pausa de seis meses en el entrenamiento de este tipo de tecnologías, argumentando que su desarrollo acelerado podría desbordar las capacidades de control e implicar grandes riesgos para la humanidad (BBC News Mundo, 2023). Un análisis filosófico más amplio de los posibles riesgos de este tipo de tecnologías para la humanidad se puede identificar en los planteamientos de autores como Sadin (2020), Han (2014) y Harari (2014), entre otros.

Sin embargo, como se vio en el capítulo 3, las perspectivas normativas, humanistas o pragmáticas pueden acusar algunos problemas en relación con sus tendencias al instrumentalismo, al voluntarismo, al sustantivismo y al determinismo. Por ejemplo, los enfoques solo normativos tienen a ser instrumentalistas y voluntaristas al suponer que estos sistemas inteligentes artificiales son neutrales, y que las regulaciones dependen enteramente del control de la voluntad humana, de los estados, de las corporaciones o, en general, de la sociedad. Por otro lado, los enfoques humanistas tienden a asumir puntos de vista sustantivistas y deterministas al considerar que la inteligencia artificial va a derivar o devenir ella misma, de forma cuasi autónoma, en formas de control y de instrumentalización negativa o de alienación de la vida humana. Adicionalmente, los enfoques pragmáticos pueden tender también hacia posiciones voluntaristas, instrumentalistas

y deterministas sociales, al proponer que debe ser la voluntad de las organizaciones o de las movilizaciones sociales, y no la de los individuos, la que debe controlar los riesgos de estas tecnologías, o que es la sociedad la que debe determinar sus posibilidades, sin suponer la capacidad de agenciamiento de estos sistemas.

Además, estas distintas perspectivas, en sus expresiones más radicales, se centran en la evaluación, en los análisis de los riesgos, en el control y en establecer límites, pero poco pueden ayudar a orientar los sentidos positivos de estas tecnologías, los usos con propósitos éticos explícitos, los sentidos constructivos y las alternativas.

Sin ignorar sus riesgos, que son altos, sin dejar de regular normativamente sus usos y sus límites, y sin dejar de realizar críticas amplias a sus posibles impactos e implicaciones culturales, es necesario ir más allá de los análisis tecnofílicos y tecnofóbicos, utópicos y distópicos, que predicen que la inteligencia artificial va a solucionar nuestros principales problemas, o que nos va a dominar, nos va a reemplazar, e incluso que nos puede extinguir.

Este tipo de posturas no van a poder regular y orientar éticamente los desarrollos de la IA al nivel que se requiere. El enfoque que se centra solo en el control no va a ser suficiente para conducir con buenos pasos el caballo brioso de la inteligencia artificial, su potencial disruptivo, su capacidad para servir para casi todo, y su devenir incierto.

Una tecnología tan potente, versátil, polifacética y dinámica como esta requiere también enfoques éticos constructivos (no solo evaluativos) que ayuden a orientar sus desarrollos *in situ*, o en el terreno, de manera concreta, caso a caso, a través de continuas retroalimentaciones y ajustes, de los distintos actores involucrados.

Una ética de la IA necesita abrirse a nuevos enfoques, en los que se desarrollen formas de instrumentalización secundarias o positivas, como lo explica Feenberg (2016), que se pueden reconocer en movimientos como el de *IA for Good* (AI for Good, 2024), que promueve las Naciones Unidas. Parte de la instrumentalización secundaria se logra cuando los sistemas se logran “alinearse” a valores, o se les inserta valores, pero dicha instrumentalización es mucho más compleja de lo que los enfoques regulativos piensan, en la medida que la construcción ética de la IA no es un asunto que compete solo a los diseñadores, los

desarrolladores y las compañías, sino también a los usuarios, y a las máquinas mismas en la configuración de su agenciamiento. Se puede entender como una co-construcción de distintos actores y actantes, humanos y no humanos, colectivos e individuales, como bien lo muestra la teoría actor-red. Más que una alineación a valores, es un ensamblaje o una co-construcción valorativa de los sistemas y todas las redes de actores que participan.

Para entender lo que implica esta co-construcción ética de los sistemas inteligentes artificiales, es fundamental abrir la caja negra, y explicar previamente, tras bambalinas, las distintas entidades humanas y los elementos materiales que lo integran, y que permiten el funcionamiento de lo que parece una máquina autónoma inteligente. Una ética integral debe reconocer e involucrar todos los componentes y todas las fases de lo que hemos denominado “la constelación sinepistémica” de la IA, y no solo lo que se evidencia en las interacciones finales.

Después de abrir la caja negra de los sistemas inteligentes artificiales, se va a discutir cómo se puede entender el estatuto moral o la condición moral de estos sistemas, fuera de los puntos de vista instrumentalistas y voluntaristas acostumbrados.

Estatuto moral de la IA y perspectivas de análisis ético

Un sistema inteligente artificial, como el chat GPT, no tiene un estatus moral por su aparente similitud al ser humano en algunas operaciones. Aunque pueda sorprendernos en relación con el lenguaje que genera, muy similar al de un ser humano, o la capacidad para presentar textos coherentes con una estructura argumentativa sólida, donde se sintetiza el contenido que deseamos, no es correcto asignarle características como conciencia, razonamiento, voluntad, autodeterminación, deliberación moral. Pero tampoco es una máquina neutral, sin estatus moral. De manera más precisa, podríamos definirlo como lo propone Floridi (2023), como un agente artificial con un estatus moral, o un agente moral sin mente. Esto requiere de algunas importantes aclaraciones.

En primer lugar, este sistema inteligente artificial puede desempeñar algunos roles morales significativos. Incluso así lo señala el mismo sistema cuando se le consulta (tabla 13).

Tabla 13. Roles morales que se pueden asumir una IA

Rol moral	Descripción
Asistente ético	Proporciona información y orientación ética en diversas situaciones para ayudar a tomar decisiones morales o resolver dilemas éticos.
Educador moral	Enseña ética y valores morales a través de conversaciones y ejemplos, útil en la educación moral de estudiantes o la capacitación de profesionales.
Moderador en línea	Modera contenido en línea para eliminar o mitigar material inapropiado o inmoral, como el discurso de odio o contenido violento.
Consejero ético	Ofrece apoyo emocional y consejos éticos a personas que enfrentan problemas morales o dilemas éticos en sus vidas.
Generador de políticas éticas	Asiste a organizaciones y gobiernos en la creación de políticas y directrices éticas, proporcionando información y análisis sobre cuestiones morales.
Traductor ético	Traduce documentos o textos éticos de un idioma a otro, o explica conceptos éticos a personas que no están familiarizadas con ellos.
Crítico ético	Analiza y evalúa cuestiones morales en obras literarias, películas, discursos y otros medios, proporcionando perspectivas críticas desde una óptica ética.
Promotor de la empatía y la comprensión	Fomenta la empatía y la comprensión al brindar información y perspectivas éticas sobre cuestiones de diversidad, inclusión y justicia social.

Fuente: OpenAI (2023).

Es importante destacar que, aunque el ChatGPT puede desempeñar estos roles morales, su capacidad para hacerlo depende en gran medida de la calidad de su entrenamiento y de la programación ética que se le

haya proporcionado. Pero también el mismo sistema explica los posibles riesgos morales o las posibles situaciones inmorales que se pueden derivar de su uso (tabla 14).

Tabla 14. Riesgos y desafíos que puede derivar del uso de IA

Riesgo/desafío	Descripción
Desinformación	ChatGPT puede generar contenido falso o engañoso, lo que contribuye a la propagación de desinformación, con graves implicaciones en temas críticos.
Sesgo y discriminación	Los modelos de lenguaje pueden aprender sesgos de los datos con los que fueron entrenados, generando contenido sesgado o discriminatorio.
Manipulación y engaño	ChatGPT puede ser utilizado para crear contenido engañoso, como correos de <i>phishing</i> o noticias falsas, con el fin de manipular o estafar.
Acoso y abuso en línea	El modelo puede ser empleado para acosar o intimidar a personas en línea, con consecuencias emocionales y psicológicas negativas para las víctimas.
Pérdida de empleo	La automatización de tareas como la atención al cliente puede llevar a la pérdida de empleos en sectores como el servicio al cliente y la redacción.
Amenazas a la privacidad	El uso de ChatGPT plantea preocupaciones sobre la privacidad, especialmente cuando genera contenido que contiene información personal o sensible.
Dependencia tecnológica	La dependencia excesiva en modelos de lenguaje puede llevar a confiar ciegamente en ellos, sin cuestionar la precisión o ética de sus respuestas.
Dificultad para verificar la autenticidad	La generación de contenido mediante IA puede dificultar distinguir entre contenido humano y generado por máquinas, afectando la confianza en la autenticidad.

Fuente: OpenAI (2023).

Para abordar estos riesgos, es importante que los desarrolladores, los usuarios y las organizaciones implementen políticas, regulaciones y prácticas éticas sólidas en el uso de modelos de lenguaje, y que promuevan la transparencia y la responsabilidad en su aplicación.

Frente a sus posibles usos y riesgos, es inevitable la pregunta sobre la posible responsabilidad moral de la IA. Como plantea Mark

Coeckelbergh, “si a una IA se le otorga una mayor capacidad de actuación y asume tareas que solían llevar a cabo los humanos, ¿hemos de atribuirle responsabilidad moral?” (2021, p. 98). Por supuesto, hablar de responsabilidad parece inadecuado porque no es un ser humano, y aún falta mucho para que la IA llegue a la singularidad tecnológica anunciada por Kurzweil (2005). Pero, de acuerdo con el nivel de autonomía agenciativa, tampoco parece preciso decir que este tipo de sistemas son neutrales, y que los problemas morales dependen enteramente de los usos y los usuarios.

Antonio Diéguez examina esta cuestión en su ensayo titulado *¿Pueden ser responsables las máquinas?*, y orienta sus planteamientos alrededor de las respuestas a la siguiente pregunta: “¿Podría ocurrir que hubiera también una forma distinta de ser agente moral, y por tanto susceptible de responsabilidad, que fuera muy distinta a la del ser humano y que encajara con el comportamiento de las máquinas?” (Diéguez, 2023).

Siguiendo la línea argumentativa señalada por Diéguez (2023), aunque sin compartir plenamente sus conclusiones, consideramos que la respuesta más adecuada a esta cuestión es la que ofrecen Floridi y Sanders (2004): los sistemas inteligentes artificiales tienen “moralidad sin mente” (*mind-less morality*). Son agentes artificiales que no tienen responsabilidad moral plena (*responsibility*), como los seres humanos, pero sí tienen obligación de rendir cuentas, de responder por las consecuencias (*accountability*). Es decir, no se les adjudica responsabilidad moral plena, sino solo responsabilidad moral causal y legal, pues desempeñan importantes roles morales y pueden generar grandes riesgos. Esta distinción ya la aplicamos con los niños o menores de edad. Frente a una situación moral ellos tienen responsabilidad (*responsibility*), en un grado menor que los adultos, pues son seres humanos conscientes de sus actos, pero no son responsables legales, o no están obligados a rendir cuentas (*accountability*), sino sus padres. Con la IA sucede al revés, aunque no tienen en sentido estricto conciencia de sus actos, estas tecnologías deberían ser transparentes, explicables, y contar con la debida autorización para desempeñar algún rol moral. También, se debería hacer un seguimiento al desarrollo de sus funciones, y ser convocadas como agentes relevantes al respectivo

rendimiento de cuentas. Un punto de vista similar lo plantean Wallach y Allen usando el término “moralidad funcional” (Wallach y Colin, 2009, p. 39) (Coeckelbergh, 2021, p. 51).

El punto de vista planteado por Floridi y Sanders (2004) tiene como punto de partida el tema ampliamente debatido hoy del estatus moral de las tecnologías (Kroes y Verbeek, 2014). Dentro de ese debate, la gran mayoría de los filósofos de las tecnologías actuales están de acuerdo en relación con afirmar que todas las tecnologías en general tienen “relevancia moral”, como lo explican Verbeek (2005, 2009, 2011), Coeckelbergh (2021) y Moreno-Ortiz (2019), y no son meros instrumentos neutrales. Sin embargo, para algunos autores, las tecnologías son relevantes en la medida que son agentes en algún sentido, como se ha explicado en este texto y puede verse en Moreno-Ortiz (2020, 2023). Pero en los textos de Floridi y Sanders (2004), y Floridi (2008) se va más lejos, pues se argumenta que los sistemas inteligentes artificiales pueden ser “agentes autónomos”, en un nivel de abstracción de agencia sin intencionalidad ni mente, en la medida que tienen: 1) interactividad, 2) autonomía y 3) adaptabilidad (Floridi y Sanders, 2004). Y a partir de esa autonomía agenciativa pueden ser “agentes morales artificiales”, con un sentido de responsabilidad moral causal y legal específico (*accountability*).

Explicando lo dicho en palabras más familiares, los sistemas de inteligentes artificiales son relevantes, es decir, cuentan en el análisis moral, en la medida que median nuestras percepciones y nuestras acciones. Por ejemplo, actualmente varias IA filtran nuestras percepciones en las redes sociales, con propósitos comerciales, políticos, etc. Basta con que identifiquen nuestro interés, por ejemplo, en comprar unos nuevos lentes, para llenarnos de publicidad sobre las distintas marcas y diseños en lentes. Según los datos tomados como base para su entrenamiento, y a partir de su capacidad para buscar y presentar la información más relevante, el ChatGPT filtra nuestro acceso a la realidad en todas las consultas que realizamos. En otras IA se identifica claramente no solo su papel para filtrar nuestras percepciones, sino también para influir en el curso de nuestras acciones, como sucede con las distintas IA que se están empleando como jueces inteligentes, para asistir las deliberaciones jurídicas,

por ejemplo, el sistema creado en Argentina denominado *Pretoria*, empleado por la Corte Constitucional Colombiana.

Adicionalmente, las IA pueden operar como “agentes sin mente” con una autonomía funcional, en muchos casos, al generar acciones o aportar opciones no previstas por los seres humanos. Algunos ejemplos de esta capacidad de agencia autónoma son los siguientes: Alphago Zero ha propuesto tipos de estrategias de juego nuevas, no conocidas por los seres humanos. Una IA puede crear un lenguaje artificial nuevo, ininteligible para los entrenadores. La función generativa del ChatGPT no consiste simplemente en reciclar la información generada y proporcionada por los seres humanos. También puede generar nueva información. Incluso, si se lo presiona demasiado, el ChatGPT puede ofrecer respuestas imprevistas, inventar cosas o generar alucinaciones. El desarrollo de estas capacidades agenciativas puede derivar de manera casi autónoma, en muchos efectos e implicaciones morales, implícitas o explícitas, asignadas o imprevistas.

El reconocimiento de estas capacidades agenciativas ha planteado la necesidad de un monitoreo permanente del funcionamiento de estas tecnologías, por sus posibles riesgos. Pero también ha servido de base, precisamente, para delegar funciones morales importantes en ellas, que pueden facilitar el manejo de situaciones morales en las sociedades, o incluso resolver cuestiones morales difíciles. Sirven como ejemplos los jueces artificiales que se han incorporado en varios tribunales y cortes, para descongestionar la gran cantidad de procesos jurídicos, o los sistemas para detectar mentirosos en los controles de las fronteras, o los policías de tránsito artificiales que monitorean permanentemente el tráfico y colocan sanciones a los infractores las 24 horas, o los sistemas que ayudan a hacer selección de personal, para evitar conflictos de intereses y preferencias, etc. En estos casos, la IA no solo tiene relevancia moral, no solo ejerce un papel de agente moral artificial, sino que es capaz de agregar valor moral en sentido positivo o negativo, en situaciones donde los seres humanos usualmente somos menos eficaces o no tenemos las capacidades. Seguramente, en los próximos años seremos testigos de la implementación de sistemas inteligentes artificiales en muchas cuestiones morales, que deberán ser monitoreados o controlados por otros sistemas para que no se salgan del control

previsto o para intentar neutralizar los usos indebidos. Por ejemplo, seguramente se implementarán sistemas para generar *fake news* con propósitos de manipulación política, y también se implementarán otros sistemas que filtrarán y bloquearán las *fake news*.

Muy seguramente, nos orientaremos hacia una sociedad en la que la IA se integrará a la mayoría de las situaciones morales ordinarias, y realizará parte de las tareas que hoy realizan los jueces, los árbitros, los policías, los gobernantes, etc. Frente a ello, más que suponer *a priori* escenarios utópicos o distópicos, con los brazos cruzados, es fundamental trabajar para preparar a las sociedades para este nuevo contexto, comenzando por caer en la cuenta de que si pretendemos mantener al genio dentro de la botella, luego de la explosión actual de la IA, la ética de la tecnología es indispensable y esencial, no solo deseable.

Una consideración importante que se debe realizar en la ética de la IA es la necesidad de realizar un proceso de acompañamiento permanente, de orientación continua, y de co-construcción de las pautas morales, en conjunto con los sistemas, considerados como agentes relevantes. Es decir, en la medida que los sistemas inteligentes artificiales no son neutrales, son agentes, y pueden agregar valor moral, la configuración moral debe centrarse en la interacción entre los humanos y las máquinas, de manera análoga, en parte, a la forma como se realiza la orientación y educación ética de una persona. En este caso, no solo se colocan normas *a priori*, y se evalúan las acciones *a posteriori*, sino que también se realiza un continuo acompañamiento, una interpretación y retroalimentación de las experiencias cotidianas, pues en ellas es donde más se puede caracterizar la complejidad moral, y tomar elementos para hacer ajustes a los distintos aspectos y actores involucrados en el curso de las acciones.

Lo que se debe configurar moralmente es el sistema amplio sociotécnico, o el ensamblaje entre los distintos actores y agentes, en el curso de la acción, desde la interpretación de las experiencias, como si fuera un proceso de calibración, o sincronización continua. Para ello es necesario realizar una interpretación continua de las experiencias “con” estas tecnologías, más que “de” ellas.

La teoría actor-red ofrece los elementos básicos para comprender este proceso de configuración moral como una co-construcción

entre los distintos agentes. Y actualmente, una muy buena perspectiva que ayuda a interpretar las experiencias con las tecnologías es la postfenomenología planteada por Don Ihde (1990, 1993, 2009, 2010), Verbeek (2005, 2011) y Rosenberger y Verbeek (2015). En Verbeek y Tijink (2020), se puede encontrar la explicación del enfoque reciente de una ética de la orientación, y de algunas propuestas metodológicas para desarrollarla.

Teniendo en cuenta los niveles de interacción, autonomía y adaptabilidad de la IA, el uso correcto de estas tecnologías es mucho más complejo que el uso correcto de cualquier artefacto que compramos en una tienda, como un horno o un taladro. En estos casos, simplemente seguimos las normas del manual de instrucciones, y no pensamos en abrir la caja negra de todos los procesos involucrados en su construcción, comercialización y uso. No esperamos con el horno o el taladro seguir principios como los de transparencia y explicabilidad. Simplemente, antes de usarlos leemos el manual de instrucciones y allí encontramos todas las pautas para su instalación y uso correcto, y la advertencia de sus principales riesgos. Si seguimos bien ese manual, realizado por la empresa constructora del artefacto, podemos garantizar en buena medida su uso correcto, dentro de parámetros éticos.

Sin embargo, la situación con la IA es mucho más compleja, porque no tiene un manual de instrucciones fijo, predefinido para su uso correcto, pues en tanto es una tecnología interactiva, adaptativa y autónoma, las posibilidades, capacidades y riesgos de esos sistemas pueden cambiar en cada situación, según las interacciones establecidas con ellos. Es necesario, entonces, intentar abrir sus cajas negras hasta donde nos sea posible para identificar sesgos, gracias a los niveles de transparencia permitidos en el sistema. También es necesario “alinearse” a valores específicos los sistemas, y pedirles cuentas en relación con su adecuación a esos valores a partir de sus niveles de explicabilidad. Pero el problema ético con la IA no termina allí, sino que, en la medida que son agentes que cada día van a formar parte, de manera cada vez más activa, de nuestras formas morales de operar sobre el mundo, es necesario también realizar un seguimiento continuo o un acompañamiento de las formas como articulamos o ensamblamos nuestras experiencias morales con los agentes morales artificiales, con el propósito

de poder co-construir sistemas heterogéneos de agentes morales humanos y no-humanos, orientados éticamente. En pocas palabras, máquinas y humanos tendremos que ajustarnos mutuamente para lograr la sociedad ética que deseemos.

Desafíos éticos de la IA

Con el *boom* de la inteligencia artificial se ha discutido mucho en los últimos años sobre las implicaciones y los impactos que está teniendo esta tecnología en la sociedad. Por un lado, se ha señalado de manera reiterada que estamos experimentando una primavera en el desarrollo de la inteligencia artificial (IA), en la que se pronostica que ella será una gran fuerza transformadora, con impactos y riesgos tan profundos en gran parte de las esferas humanas, que superan incluso algunas de las anticipaciones de la ciencia ficción. Ello se corresponde con las gigantes inversiones realizadas por las grandes compañías tecnológicas en los desarrollos y usos de la IA, y los temores expresados en escenarios distópicos. Por otro lado, algunos estudios y análisis del comportamiento de los mercados están dudando del real desarrollo e impacto de esta tecnología, y están sembrado la sospecha sobre una posible burbuja especulativa de sobrevaloración de sus avances y posibles efectos. En efecto, las curvas sobre los usos y el margen de ganancias que esperan obtener las compañías de las implementaciones de esta tecnología no han crecido como se esperaba e incluso amenazan con decrecer. Y el denominado “IA *washing*” o lavado de imagen con inteligencia artificial se ha convertido en una estrategia de marketing engañosa para exagerar la integración de la IA en productos y servicios con el propósito de hacerlos parecer más avanzados e innovadores de lo que realmente son.

Fuera del entusiasmo o del escepticismo alrededor de lo que se puede especular con esta tecnología, es necesario tener en cuenta lo siguiente: 1) este tipo de tecnología ha generado muchas más expectativas que ninguna otra en las décadas recientes, pero las implementaciones y transformaciones asociadas a ella no están sucediendo tan rápido como se esperaba; 2) fuera de la idea de un determinismo tecnológico,

estas transformaciones suponen muchos cambios sociales y económicos, y no solo avances tecnológicos; 3) fuera del entusiasmo generado, en los aspectos funcionales esta tecnología tiene un gran potencial transformador disruptivo, pero ese potencial no necesariamente se corresponde con implementaciones viables y deseables de manera rápida en el mercado y en las sociedades; 4) sin embargo, aunque todavía no tenga el impacto esperado, esta tecnología, especialmente la de más amplio uso, que es la IA generativa, ha llegado para quedarse e instalarse en nuestras vidas, y generar un conjunto de disrupciones que apenas empiezan.

Los anteriores planteamientos parecen contradecirse, pero no es así. Algunos tipos de tecnologías han experimentado grandes avances, que permiten un gran potencial de transformación, pero su aplicación no obedece a un imperativo tecnológico, sino que se ha frenado o se ha producido de manera más gradual de lo esperado por los riesgos que pueden generar, o por las condiciones o los costos asociados a su aplicación. Por ejemplo, los cambios tecnológicos derivados de la genética y la bioingeniería tienen un gran potencial transformador, pero se han limitado y regulado fuertemente su aplicación, debido a los grandes riesgos que supone. O los automóviles eléctricos se dejaron de implementar en las décadas pasadas por motivos políticos, sociales y económicos, independientemente de sus desarrollos funcionales, y ahora se han vuelto relevantes y se están implementando de manera acelerada debido a las nuevas condiciones derivadas del cambio climático. La inteligencia artificial comparte algunas características con los ejemplos señalados: 1) su implementación puede generar grandes riesgos en algunos casos, que apenas se están comenzando a regular y aún no se tiene claro qué tan restrictivas van a ser esas regulaciones; 2) su relevancia y sus alcances no parecen ser muy claros por ahora para justificar su acelerada aplicación en relación con las fuertes condiciones sociales y económicas que requieren su desarrollo e implementación; y 3) los usos más importantes de la IA suponen la integración con otros procesos tecnológicos o suponen el desarrollo de ecosistemas tecnológicos especializados. En efecto, por estas y otras razones, a pesar de las grandes expectativas puestas en este tipo de tecnologías, del gran efecto publicitario generado y de su potencial, aún es prematuro

anticipar las formas específicas como la inteligencia artificial va a introducirse en nuestras vidas y los impactos que generará.

En todo caso, el mejor escenario de desarrollo de esta tecnología es irónicamente el de una apropiación “inteligente” y ética de la inteligencia artificial, es decir, gradual, prudente, y regulada, en contraste con una apropiación “no inteligente” y riesgosa, es decir, desde perspectivas tecnófilas, y de manera poco regulada, orientada principalmente por los efectos especulativos publicitarios y las metas económicas de las empresas privadas que las desarrollan. Precisamente, una aplicación inteligente de esta tecnología puede corresponder con transformaciones más profundas, consistentes y sostenibles, tanto social como ambientalmente, en contraste con aplicaciones superficiales, con efectos poco previsibles y problemas éticos.

Jugando un poco con la contraposición entre las palabras “inteligente” y “no inteligente”, una apropiación no inteligente de la IA tiene que ver con implementaciones precipitadas, impulsada por afanes innovadores muy especulativos, con beneficios económicos claros para el sector que las lidera, pero muy volátiles e inciertos a largo plazo para el mercado y la sociedad. A nivel social, se relaciona con procesos de automatización que prometen agilizar y aumentar la eficiencia, sin un cálculo responsable de las implicaciones en el empleo, en la calidad de los procesos, en las implicaciones éticas en temas como la privacidad, en los impactos ambientales y en la satisfacción de los usuarios. En la medida que esa implementación no inteligente no tiene en cuenta el fortalecimiento de las infraestructuras, de las capacidades y de las condiciones que se requieren para el desarrollo de nichos y contextos tecnológicos inteligentes, puede generar desarrollos parciales donde se fomenta el IA *washing*, y puede conducir a la ampliación de las brechas, desigualdades y las exclusiones sociales. Sin análisis adecuado, los riesgos sociales pueden ser mayores, como la pérdida de numerosos empleos debido a procesos de automatización, y la afectación a la seguridad y la privacidad de la información en el contexto de un capitalismo de la vigilancia (Zuboff, 2020).

Pero esta implementación inteligente tiene como condición importante una sociedad informada, dispuesta a comprender el funcionamiento de este tipo de tecnología, fuera de escenarios imaginarios

de ciencia ficción, y comprometida con analizar sus condiciones sociales, sus posibles problemas éticos, y sus posibilidades de orientación.

Preguntas

1. ¿Puede la inteligencia artificial emular verdaderamente la conciencia humana o su capacidad de razonamiento está limitada a la mera simulación de procesos cognitivos?
2. Si la inteligencia artificial no tiene mente ni conciencia, ¿puede considerarse un agente moral capaz de tomar decisiones éticas y asumir responsabilidades?
3. ¿Qué implicaciones tiene el desarrollo de IA fuerte sobre nuestra concepción del libre albedrío y la autonomía humana?
4. ¿Cómo afecta el uso de sistemas de IA en la toma de decisiones humanas en ámbitos éticos, políticos o sociales a la percepción de la responsabilidad moral de los humanos que los programan?
5. ¿Es éticamente aceptable permitir que sistemas de IA asuman roles que involucren juicios morales, como en la moderación de contenido en línea, sin la intervención de un agente humano?

Para saber más

- *Artificial Intelligence: A Modern Approach*, de Stuart Russell y Peter Norvig (2021), considerado un libro de referencia en el campo de la inteligencia artificial, ofrece una visión integral de los conceptos, teorías y aplicaciones de la IA. Aborda desde los fundamentos algorítmicos hasta las aplicaciones prácticas y los desafíos contemporáneos de la IA.

- *Formalizing Common Sense: Papers by John McCarthy*, de John McCarthy (1990). Este libro recopila algunos de los trabajos más influyentes de McCarthy, uno de los fundadores de la inteligencia artificial. A través de sus artículos, se exploran conceptos clave como la representación del conocimiento y el razonamiento en máquinas, ofreciendo una buena visión sobre la formalización de la lógica y el sentido común en la IA.
- *Moral Machines: Teaching Robots Right from Wrong*, de Wendell Wallach y Colin Allen (2009). Este texto se centra en los aspectos éticos de la IA, discutiendo cómo programar y diseñar robots para que puedan tomar decisiones morales. Es una obra esencial para quienes desean explorar los dilemas éticos que la inteligencia artificial plantea y cómo se pueden abordar mediante la programación y la ética aplicada.

Consideraciones finales

Este libro se propuso ofrecer una introducción sistemática a los asuntos que, con mayor frecuencia, organizan el debate contemporáneo en filosofía de la tecnología. La intención no fue reunir una serie dispersa de temas, sino presentar un marco de comprensión que permita leer los fenómenos tecnológicos con criterios conceptuales estables, evitando dos extremos habituales: reducir la tecnología a objetos aislados o, en sentido contrario, diluirla en explicaciones generales sobre “lo social” sin atender a sus mediaciones materiales, operativas y organizacionales. En ese sentido, el recorrido del libro defendió que pensar filosóficamente la tecnología implica articular tres planos: la pregunta por su naturaleza, la pregunta por el tipo de conocimiento que la hace posible y la pregunta por los criterios con los cuales se la evalúa y orienta en la vida colectiva.

En el primer capítulo se ofreció una definición que funciona como hilo conductor de todo el texto: la tecnología como una forma particular de acción técnica contemporánea mediada por procesos de diseño, caracterizada por su carácter colectivo y organizacional, y consolidada históricamente tras la primera Revolución Industrial. Entendida de este modo, la tecnología no se deja capturar por una lista de artefactos ni por la idea de un “medio” subordinado a fines previamente dados. Se trata, antes bien, de una modalidad de acción que reordena posibilidades: configura entornos, estabiliza prácticas, distribuye competencias y responsabilidades, y se apoya en esquemas de representación y evaluación que orientan la intervención sobre el mundo.

Esta definición también permite evitar la identificación simple entre tecnología y ciencia aplicada. Aunque los vínculos entre ciencia y tecnología son múltiples y, en muchos dominios, estrechos, la tecnología expresa un tipo de racionalidad orientada a la producción, la operación y la transformación, que incorpora criterios normativos internos como funcionalidad, seguridad, confiabilidad, eficiencia, compatibilidad e integración en contextos de uso.

A partir de este punto de partida, el segundo capítulo hizo explícita la necesidad de tratar el conocimiento tecnológico como un objeto filosófico específico. Allí se mostró que el “saber tecnológico” no se limita a proposiciones verdaderas o explicaciones teóricas, sino que incluye modelos, representaciones, procedimientos, habilidades y prácticas de ensayo, ajuste y validación. La comprensión de este tipo de conocimiento exige reconocer tanto su dimensión instrumental como su dimensión normativa, pues los criterios de éxito tecnológico no se reducen a la verdad, sino que involucran estándares de desempeño, restricciones de diseño y compromisos prácticos con condiciones materiales, económicas y organizacionales. Esta perspectiva permite, además, comprender por qué el aprendizaje tecnológico y el diseño no pueden describirse adecuadamente como una simple derivación de la investigación científica, aun cuando se apoyen en ella en muchos casos.

El tercer capítulo trasladó estas clarificaciones al plano ético, mostrando que la pregunta moral por la tecnología no puede resolverse solo mediante evaluaciones posteriores de “impactos”. Si la tecnología es una forma de acción mediada por diseño y sostenida por arreglos institucionales, entonces también organiza el campo de lo posible y lo permisible: habilita cursos de acción, dificulta otros, distribuye cargas y beneficios, y redefine expectativas sobre lo que cuenta como una práctica aceptable. En este marco, la ética de la tecnología requiere prestar atención a las mediaciones: a cómo se incorporan criterios y valores en arquitecturas, interfaces, infraestructuras, métricas, protocolos y procedimientos de mantenimiento y gobernanza. Por ello, el libro insistió en que la evaluación ética debe dialogar con el diseño, con la regulación y con la responsabilidad institucional, y no limitarse a la crítica externa o al juicio retrospectivo.

El cuarto capítulo, dedicado a la inteligencia artificial y, de manera especial, a la IA generativa, permitió poner a prueba el marco general desarrollado en el resto del libro. La IA hace visibles, con particular intensidad, cuestiones clásicas de la filosofía de la tecnología: problemas de agencia y atribución de responsabilidad, límites y alcances del conocimiento automatizado, opacidad de sistemas complejos, y tensiones entre eficiencia, control y rendición de cuentas. Al mismo tiempo, la IA expone de forma nítida el carácter organizacional de la tecnología contemporánea: su dependencia de infraestructuras materiales, disponibilidad energética, cadenas de trabajo, economías de datos, plataformas y marcos normativos. La discusión sobre IA, en consecuencia, no debe entenderse como un apéndice temático, sino como un caso que ayuda a comprender por qué las preguntas ontológicas, epistemológicas y éticas no son compartimentos aislados, sino dimensiones que se implican y se corrigen mutuamente.

Si este libro pretende ser un manual, lo es en un sentido preciso: no como un repertorio exhaustivo, sino como un conjunto de orientaciones para leer y discutir la tecnología con mayor claridad conceptual. El lector puede volver a cada capítulo con una pregunta guía distinta: qué es la tecnología cuando se la toma como forma de acción colectiva mediada por diseño; qué se conoce y cómo se valida en las prácticas tecnológicas; cómo se evalúan y orientan las mediaciones que reconfiguran prácticas y valores; y por qué la IA intensifica dilemas ya presentes en otros dominios técnicos. En esa misma dirección, las preguntas de discusión y las lecturas recomendadas no son anexos ornamentales, sino invitaciones a continuar el trabajo filosófico allí donde este texto necesariamente se detiene: en el análisis de casos, en el contraste de enfoques y en la deliberación crítica sobre la dirección de los desarrollos tecnológicos. Con ello se busca que la filosofía de la tecnología no quede confinada a una reflexión general sobre “la técnica”, sino que se convierta en una herramienta de comprensión y orientación para un mundo en el que lo tecnológico estructura, cada vez más, las formas de vida compartidas.

Referencias

- Achterhuis, H. (Ed.). (2001). *American philosophy of technology: The empirical turn* (R. P. Crease, Trans.). Indiana University Press.
- Adorno, T. W., & Horkheimer, M. (2003). *Dialéctica de la Ilustración: fragmentos filosóficos* (J. J. Sánchez, Trad.; obra original publicada en 1947). Trotta.
- Álvarez Osorio, J. D. (2026). ¿Pueden las inteligencias artificiales ser creativas? *Trilogía Ciencia Tecnología Sociedad*, 18(38), e3710. <https://doi.org/10.22430/21457778.3710>
- Antolínez Uribe, D. (2023). Una comprensión (in)material de las técnicas. A propósito de fomentar vínculos entre las humanidades y los estudios de ciencia y tecnología. *Cuadernos de Filosofía Latinoamericana*, 44(128), 121-138. <https://doi.org/10.15332/25005375.8256>
- Aranguren, M. (2010). Análisis crítico del modelo de variación ciega y retención selectiva de la creatividad. *Interdisciplinaria. Revista de Psicología y Ciencias Afines*, (2), 315-334. https://www.redalyc.org/pdf/180/18018446008.pdf?utm_source=chatgpt.com
- Arroyave Montoya, M. (2013). Compromiso ontológico de la música occidental. *Ideas y Valores. Revista Colombiana de Filosofía*, 62(153), 7-30. https://www.redalyc.org/pdf/809/80929838001.pdf?utm_source=chatgpt.com
- Asimov, I. (1950). *I, Robot*. Street & Smith Publications.
- Baird, D. (2004). *Thing knowledge: A philosophy of scientific instruments*. University of California Press.
- Basalla, G. (2011). *La evolución de la tecnología* (J. Vigil, Trad.). Crítica.

- Baudrillard, J. (2010). *El sistema de los objetos*. Siglo XXI Editores.
- BBC News Mundo. (2023, 29 de marzo). *La carta en la que más de 1.000 expertos piden frenar la inteligencia artificial por ser una “amenaza para la humanidad”*. <https://www.bbc.com/mundo/noticias-65117146>
- Beauchamp, T. L., & Childress, J. F. (1979). *Principles of biomedical ethics*. Oxford University Press.
- Begon, M., Townsend, C. R., & Harper, J. L. (2006). *Ecology: From individuals to ecosystems* (4th ed.). Blackwell Publishing.
- Berg, O., Pedersen, S. A., & Hendricks, V. F. (Eds.). (2009). *A companion to the philosophy of technology*. Wiley-Blackwell.
- Bijker, W. E., Hughes, T. P., & Pinch, T. J. (Eds.). (1987). *The social construction of technological systems: New directions in the sociology and history of technology*. MIT Press.
- Boden, M. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). Routledge. (Original work published 1990).
- Boström, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Brey, P. (2010). Philosophy of technology after the empirical turn. *Techné: Research in Philosophy and Technology*, 14(1), 36-48. <https://doi.org/10.5840/techne20101416>
- Bridgelall, R. (2024). Unraveling the mysteries of AI chatbots. *Artificial Intelligence Review*, 57, 89. <https://doi.org/10.1007/s10462-024-10720-7>
- Bringsjord, S., & Govindarajulu, N. S. (2024). Artificial intelligence. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024 ed.). <https://plato.stanford.edu/archives/fall2024/entries/artificial-intelligence/>
- Broens, R. C. J. A. M., & de Vries, M. J. (2003). Classifying technological knowledge for presentation to mechanical engineering designers. *Design Studies*, 24(5), 457-471. [https://doi.org/10.1016/S0142-694X\(03\)00022-X](https://doi.org/10.1016/S0142-694X(03)00022-X)
- Broncano, F. (2001). *Mundos artificiales: filosofía del cambio tecnológico*. Paidós Ibérica.
- Broncano, F. (2005). La agencia técnica. *Revista CTS*, 95(105).
- Broncano, F. (2006). *Entre ingenieros y ciudadanos: filosofía de la técnica para días de democracia*. Montesinos Ensayo.
- Broncano, F. (2009). *La melancolía del ciborg*. Herder.

- Broncano, F. (2012). *La estrategia del simbiote: cultura material para nuevas humanidades*. Delirio.
- Bunge, M. (1960). *La ciencia, su método y filosofía*. Ediciones Siglo Veinte.
- Bunge, M. (1963). Tecnología, ciencia y filosofía. *Anales de la Universidad de Chile*, (126), 329-347.
- Bunge, M. (1966). Technology as applied science. *Technology and Culture*, 7(3), 329-347. <https://doi.org/10.2307/3101932>
- Bush, V. (1945). *Science, the endless frontier*. United States Government Printing Office.
- Byrne, R. W. (2004). The manual skills and cognition that lie behind hominid tool use. In A. E. Russon & D. R. Begun (Eds.), *The evolution of thought: Evolutionary origins of great ape intelligence* (pp. 31-44). Cambridge University Press.
- Byung-Chul Han. (2014). *Psicopolítica*. Herder.
- Callon, M. (1986). Some elements of a sociology of translation: Domestication of the scallops and the fishermen. In J. Law (Ed.), *Power, action and belief: A new sociology of knowledge?* (pp. 196-223). Routledge.
- Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, 67, 380-400.
- Chakrabarty, D. (2018). *Planetary crises and the difficulty of being modern*. *Millennium: Journal of International Studies*, 46(3), 259-282.
- Chomsky, N., Roberts, I., & Watumull, J. (2023, March 8). Noam Chomsky: The false promise of ChatGPT. *The New York Times*. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
- Coeckelbergh, M. (2021). *Ética de la inteligencia artificial*. Cátedra.
- Crevier, D. (1993). *AI: The tumultuous search for artificial intelligence*. Basic Books.
- Darwin, C. (2013). *The descent of man*. Wordsworth Editions.
- David, M. (2001). *Engines of logic: Mathematicians and the origin of the computer*. W. W. Norton.
- Dawkins, R. (2000). *El gen egoísta: las bases biológicas de nuestra conducta*. Salvat.
- De Fusco, R. (2005). *Historia del diseño* (M. Izquierdo, Trad.). Santa & Cole.

- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, 6(4), 485-507. <https://doi.org/10.1007/s11097-007-9076-9>
- De Vries, M. J. (2016). *Teaching about technology: An introduction to the philosophy of technology for non-philosophers* (2nd ed.). Springer.
- De Vries, M., & Meijers, A. (2013). Technological knowledge. In J. K. B. Olsen, S. A. Pedersen, & V. F. Hendricks (Eds.), *A Companion to the Philosophy of Technology* (pp. 70-78). Wiley-Blackwell.
- Deleuze, G. (1994). *Lógica del sentido* (M. Morey, Trad.). Paidós.
- Deleuze, G. (2002). *Diferencia y repetición* (M. S. Delpy & H. Beccacece, Trans.). Amorrortu.
- Deleuze, G., & Guattari, F. (1995). *El antiedipo: capitalismo y esquizofrenia I* (F. Monge, Trad.). Paidós.
- Deleuze, G., & Guattari, F. (1997). *Mil mesetas: capitalismo y esquizofrenia II* (J. Vázquez & U. Larraceleta, Trans.). Pre-Textos.
- Derry, T. K., & Williams, T. I. (1960). *A short history of technology from the earliest times to A. D. 1900*. Courier Corporation.
- Diéguez, A. (2021). *Cuerpos inadecuados: el desafío transhumanista a la filosofía*. Herder.
- Diéguez, A. (2023, July). ¿Pueden ser responsables las máquinas? *Letras Libres*. <https://letraslibres.com/wp-content/uploads/2023/07/dossier-dieguez-esp.pdf>
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429-452. <https://doi.org/10.1007/s11097-005-9002-y>
- Domínguez Rendón, R. A., & Monterroza Ríos, A. D. (2019). La perspectiva humanista como base para una evaluación crítica y pluralista de la tecnología y el diseño. *Revista Universidad Pontificia Bolivariana*, 71(98).
- Domínguez, R., & Monterroza, A. (2015). Revisión crítica de los modelos de cambio tecnológico y su incidencia en el diseño industrial. En *Por un diseño crítico y social. 10 años de la Facultad de Diseño UPB* (pp. 138-143). Editorial Universidad Pontificia Bolivariana.
- Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, 11(3), 147-162.

- Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
- Duarte Arias, D. A. (2025). Conducta agencial, determinismo y artefactos tecnológicos. *Trilogía Ciencia Tecnología Sociedad*, 17(35), e3300. <https://doi.org/10.22430/21457778.3300>
- Duran Allimant, R. (2020). Ética de la tecnología: acerca de la moralidad de los artefactos técnicos. *Filosofía Unisinos*, 21(1), 47-55. <https://doi.org/10.4013/fsu.2020.211.05>
- Durbin, P. (2003). Ética, o cómo tratar democráticamente los problemas tecno-sociales. *Isegoría*, 28, 19-31. <https://doi.org/10.3989/isegoria.2003.i28.504>
- Ebel, S. J., Call, J., & Tomasello, M. (2020). How apes use tools: Flexible tool use and tool selection in great apes. *Journal of Comparative Psychology*, 134(3), 237-248.
- Echeverría Ezponda, J. (1999). *Los señores del aire: telépolis y el tercer mundo*. Destino.
- Echeverría, J. (2001). *Filosofía de la tecnología*. OEI.
- Ellul, J. (1954). *La technique ou l'enjeu du siècle*. Éditions Economica.
- Ellul, J. (1964). *The technological society*. Knopf.
- Elster, J. (1983). *Explaining technical change: A case study in the philosophy of science*. Cambridge University Press.
- European Commission. (2019, April 8). *Ethics guidelines for trustworthy AI*. https://www.europarl.europa.eu/cmsdata/196377/AI%20HLEG_Ethics%20Guidelines%20for%20Trustworthy%20AI.pdf
- Evans, T. A. y Westergaard, G. C. (2006). Self-control and tool use in tufted capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 120(2), 163-166. <https://doi.org/10.1037/0735-7036.120.2.163>
- Fantl, J. (2017). Knowledge how. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/knowledge-how/>
- Farías, R. H., & Cuevas, M. V. (2018). *De la piedra al pixel: breve historia de la materia, las telecomunicaciones y la ciencia*. Kairón.
- Feenberg, A. (1991). *Critical theory of technology*. Oxford University Press.
- Feenberg, A. (1995). *Alternative modernity: The technical turn in philosophy and social theory*. University of California Press.
- Feenberg, A. (2005). Teoría crítica de la tecnología. *Revista CTS*, 5(2), 109-123.
- Feenberg, A. (2016). *La tecnología en cuestión* (C. Scotta, Trad.). Prometeo Libros.

- Feibleman, J. K. (1961). Pure science, applied science, technology, engineering: An attempt at definitions. *Technology and Culture*, 2(4), 305-317.
- Ferguson, E. S. (1992). *Engineering and the mind's eye*. MIT Press.
- Fernández Zubieta, A. (2009). *El constructivismo social en ciencia y tecnología: las consecuencias no previstas de la ambivalencia epistemológica*. *Arbor: Ciencia, Pensamiento y Cultura*, (737), 689-703. <https://doi.org/10.3989/arbor.2009.738n1046>
- Ferrater Mora, J. (1978). *Diccionario de filosofía*. Montecasino.
- Ferry, L. (2016). *La revolución transhumanista*. Alianza Editorial.
- Flórez, D. (2023). Semantic changes in the language of technology. En M. Hinton (Ed.), *The Oxford handbook of applied philosophy of language*. Oxford University Press. (Forthcoming)
- Flórez, D., & García, C. (2017). La naturaleza de la tecnología y sus vínculos con la ciencia: una perspectiva realista y analógica. *Discusiones Filosóficas*, 18(30), 63-78. <https://doi.org/10.17151/difil.2017.18.30.4>
- Flórez, D., García-Duque, C. E., & Osorio, J. C. (2019). *Is technology (still) applied science?* *Technology in Society*, 59, 101193. <https://doi.org/10.1016/j.techsoc.2019.101193>
- Flórez Vega, G. A. (2024). Esbozo para una perspectiva integral sobre la ética en el contexto del determinismo tecnológico. *Trilogía Ciencia Tecnología Sociedad*, 16(33), e3128. <https://doi.org/10.22430/21457778.3128>
- Floridi, L. (2008). The method of levels of abstraction. *Minds and Machines*, 18(3), 303-329. <https://doi.org/10.1007/s11023-008-9113-7>
- Floridi, L. (2023). AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy & Technology*, 36, article 15. <https://doi.org/10.1007/s13347-023-00621-y>
- Floridi, L., & Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3), 349-379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- Foster, R. (1986). Working the S-curve: Assessing technological threats. *Research Management*, 17, 20. <https://doi.org/10.1080/00345334.1986.11756976>
- Franssen, M., Lokhorst, G.-J., & van de Poel, I. (2024). Philosophy of technology. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of*

- Philosophy* (Fall 2024 ed.). <https://plato.stanford.edu/archives/fall2024/entries/technology/>
- Freeman, S., & Herron, J. (2002). *Análisis evolutivo*. Prentice Hall.
- Gallón, L. (2007). *Una introducción a la historia de las teorías de la evolución de la tecnología*. Universidad Pontificia Bolivariana.
- García Palacios, E. M., González Galbarte, J., López Cerezo, J. A., Luján, J. L., Gordillo, M. M., Osorio, C., & Valdés, C. (2001). *Ciencia, tecnología y sociedad: una aproximación conceptual*. OEI.
- García Duque, C.E. (1997). *Evolución histórica del pensamiento científico: desde la antigüedad clásica hasta el período moderno*. Universidad de Manizales
- Gay, A., & Samar, L. (2007). *El diseño industrial en la historia*. Ediciones TEC.
- German, T. P., & Barrett, H. C. (2005). Functional fixedness in a technologically sparse culture. *Psychological Science*, 16(1), 1-5.
- German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review*, 7(4), 707-712. <https://doi.org/10.3758/BF03213010>
- Gentile, N. (2023, marzo). *¿Cómo funciona ChatGPT? La revolución de la inteligencia artificial* [Video]. YouTube. <https://youtu.be/FdZ8LKjJBhQ>
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Gille, B. (Ed.). (1978). *Histoire des techniques*. Gallimard.
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173-198. <https://doi.org/10.1007/BF01700692>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Goodwin, K. (2009). *Designing for the digital age: How to create human-centered products and services*. Wiley.
- Gould, S. J. (2006). *El pulgar del panda* (A. Resines, Trad.). Crítica.
- Grech, P. (2013). *An introduction to engineering thinking*. Morgan & Claypool.
- Greif, M. L., & Needham, A. (2011). The development of human tool use early in life. En T. McCormack (Ed.), *Tool use and causal cognition* (pp. 51-68). Oxford University Press.
- Guzmán-Ortiz, S. M. (2022). Ética crítica y política de las tecnologías. Ciudadanías mediadas en el Estado Plataforma. Análisis de Movilidad en

- Línea. *Cuadernos de Filosofía Latinoamericana*, 43(126). <https://doi.org/10.15332/25005375.7606>
- Habermas, J. (2012). *El futuro de la naturaleza humana: ¿hacia un eugenismo liberal?* Paidós.
- Han, B.-C. (2017). *Psychopolitics: Neoliberalism and new technologies of power*. Verso.
- Hansson, S. O. (2017). *The ethics of technology: Methods and approaches*. Rowman & Littlefield.
- Harari, Y. N. (2014). *Sapiens: de animales a dioses*. Debate.
- Haraway, D. (1995). *Ciencia, cyborg y mujeres: la reinvención de la naturaleza*. Tecnos.
- Hegel, G. W. F. (1978). *Fenomenología del espíritu* (W. Roces, Trad.; obra original publicada en 1807). Fondo de Cultura Económica.
- Heidegger, M. (1994). *La pregunta por la técnica*. En *Conferencias y artículos* (E. Barjau, Trad.; pp. 9-37). Ediciones del Serbal. (Trabajo original publicado en 1953)
- Hickman, L. A. (2001). *Philosophical tools for technological culture: Putting pragmatism to work*. Indiana University Press.
- Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, 26, 26-38. <https://doi.org/10.1007/s10676-024-09775-5>
- Hottois, G. (1993). *Simondon et la philosophie de la "culture technique"*. De Boeck Université.
- Hui, Y. (2016). *On the existence of digital objects* (Vol. 48). University of Minnesota Press.
- Ibarra, A., & Olivé, L. (Eds.). (2003). *Cuestiones éticas de la ciencia y la tecnología en el siglo XXI*. Biblioteca Nueva/OEI.
- Ihde, D. (1979). *Technics and praxis: A philosophy of technology*. Reidel.
- Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Indiana University Press.
- Ihde, D. (1993). *Postphenomenology: Essays in the postmodern context*. Northwestern University Press.
- Ihde, D. (2009). *Postphenomenology and technoscience: The Peking University lectures*. SUNY Press.
- Ihde, D. (2010). *Heidegger's technologies: Postphenomenological perspectives*. Fordham University Press.

- Ingold, T. (2000). *The perception of the environment: Essays in livelihood, dwelling and skill*. Routledge.
- Ingold, T. (2012). Toward an ecology of materials. *Annual Review of Anthropology*, 41, 427-442. <https://doi.org/10.1146/annurev-anthro-081309-145920>
- International Telecommunication Union. (2024, June). *AI for Good*. <https://aiforgood.itu.int/>
- Isaacson, W. (2014). *The innovators*. Simon & Schuster.
- Jasanoff, S. (2016). *The ethics of invention: Technology and the human future*. W. W. Norton.
- Jonas, H. (1984). *The imperative of responsibility: In search of an ethics for the technological age*. University of Chicago Press.
- Kapp, E. (1877). *Grundlinien einer Philosophie der Technik*. Westermann.
- Kapp, E. (2018). *Elements of a philosophy of technology: On the evolutionary history of culture* (L. K. Wolfe, Trans.). University of Minnesota Press.
- Kelly, K. (2016). *The inevitable*. Viking.
- Keulartz, J., Korthals, M., Schermer, M., & Swierstra, T. (Eds.). (2002). *Pragmatist ethics for a technological culture*. Kluwer.
- Keulartz, J., Korthals, M., Schermer, M., & Swierstra, T. (2004). Ethics in technological culture: A programmatic proposal for a pragmatist approach. *Science, Technology, & Human Values*, 29(1), 3-29. <https://doi.org/10.1177/0162243903259188>
- Kirman, A. P. (1993). Ants, rationality, and recruitment. *The Quarterly Journal of Economics*, 108(1), 137-156.
- Krippendorff, K. (2006). *The semantic turn: A new foundation for design*. Taylor & Francis.
- Kroes, P., & Verbeek, P.-P. (Eds.). (2014). *The moral status of technical artifacts*. Springer.
- Kuhn, T. S. (2004). *La estructura de las revoluciones científicas* (A. Contín, Trad.). Fondo de Cultura Económica.
- Kurzweil, R. (1999). *The age of spiritual machines: When computers exceed human intelligence*. Viking.
- Kurzweil, R. (2005). *The singularity is near*. Viking.
- Lafitte, J. (1932). *Réflexions sur la science des machines*. Bloud & Gay.
- Lankshear, C., & Knobel, M. (2008). *Digital literacies: Concepts, policies and practices*. Peter Lang.

- Latour, B. (2005). *Reassembling the social: An introduction to Actor-Network-Theory*. Oxford University Press.
- Latour, B. (2008). *Reensamblar lo social: Una introducción a la teoría del actor-red*. Manantial.
- Law, J. (2009). Actor network theory and material semiotics. In B. S. Turner (Ed.), *The new Blackwell companion to social theory* (pp. 141-158). Blackwell.
- Lawler, D., & Vega, J. (2009). *La respuesta a la pregunta: metafísica, técnica y valores*. Biblos.
- Leroi-Gourhan, A. (1971). *El gesto y la palabra*. Universidad Central de Venezuela.
- Lewens, T. (2015). *Cultural evolution: Conceptual challenges*. Oxford University Press.
- Light, A. (2001). The urban blind spot in environmental ethics. *Environmental Politics*, 10(1), 7-35. <https://doi.org/10.1080/714000511>
- Liu, M.-J., Xiong, C.-H., Xiong, L., & Huang, X.-L. (2016). Biomechanical characteristics of hand coordination in grasping activities of daily living. *PLoS ONE*, 11(1), e0146193. <https://doi.org/10.1371/journal.pone.0146193>
- Liz, M. (2002). *Un metafísico en Tecnolandia: realidad, conocimiento y acción bajo nuevos puntos de vista*. Editum.
- Lozano Leyva, M. (2007). *De Arquímedes a Einstein: los diez experimentos más bellos de la física*. Debate.
- Luna Herrera, L. (2026). ¿Es la tecnología en sí misma buena o mala? Un enfoque intencionalista-disposicional. *Trilogía Ciencia Tecnología Sociedad*, 18(38), e3725. <https://doi.org/10.22430/21457778.3725>
- Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. MIT Press.
- Mangalam, M., Frigaszy, D. M., Visalberghi, E., & Iriki, A. (2018). *Unique perceptuomotor control of stone hammers in wild monkeys*. *Biology Letters*, 14(1), 20170587. <https://doi.org/10.1098/rsbl.2017.0587>
- Marcuse, H. (1968). *El hombre unidimensional*. Ariel.
- Marcuse, H. (1972). El individuo en la gran sociedad. En *Ensayos sobre política y cultura*. Ariel.
- Marcuse, H. (1993). *El hombre unidimensional* (A. Elorza, Trad.; obra original publicada en 1964). Planeta.

- Márquez, M. M. (2020). Artefactos y entidades naturales modificadas por medio de la biotecnología. *Trilogía Ciencia Tecnología Sociedad*, 12(22), 197-215. <https://doi.org/10.22430/21457778.1242>
- Marx, K. (1867). *Das Kapital: Kritik der politischen Ökonomie* (Vol. 1). Verlag von Otto Meissner.
- Marx, K. (2004). *Manuscritos económico-filosóficos de 1844* (F. Rubio Llorente, Trad.; obra original publicada en 1844). Alianza Editorial.
- Marx, L., & Roe Smith, M. (Eds.). (1994). *Does technology drive history? The dilemma of technological determinism*. MIT Press.
- Mayor, A. (2018). *Gods and robots: Myths, machines, and ancient dreams of technology*. Princeton University Press.
- Maxwell, J. C. (1865). A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155, 459-512.
- McCarthy, J. (2006, June 25). *The Philosophy of AI and the AI of Philosophy*. <http://jmc.stanford.edu/articles/aiphil2/aiphil2.pdf>
- Meijers, A. (Ed.). (2009). *Philosophy of technology and engineering sciences*. Elsevier.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Gallimard.
- Miller, D. P. (2009). *James Watt, chemist: Understanding the origins of the steam age*. Pickering & Chatto.
- Mitcham, C. (1989). *¿Qué es la filosofía de la tecnología?* Anthropos.
- Mitcham, C. (1994). *Thinking through technology: The path between engineering and philosophy*. University of Chicago Press.
- Mitcham, C., & Waelbers, K. (2009). Technology and ethics: Overview. In J. K. B. Olsen, S. A. Pedersen, & V. F. Hendricks (Eds.), *A Companion to the Philosophy of Technology* (pp. 365-383). Wiley-Blackwell.
- Mithen, S. (1998). *Arqueología de la mente: Orígenes del arte, de la religión y de la ciencia* (M. J. Aubet, Trad.). Crítica. (Obra original publicada en 1996)
- Moky, J. (1990). *The lever of riches*. Oxford University Press.
- Monterroza-Ríos, Á. (2017). Una revisión crítica a la teoría del actor-red para el estudio de los artefactos. *Trilogía Ciencia Tecnología Sociedad*, 9(17), 49-62. <https://doi.org/10.22430/21457778.616>
- Monterroza-Ríos, Á. D. (2019). El papel retroalimentador de la interacción con los artefactos en el desarrollo de las técnicas humanas. *Trilogía Ciencia Tecnología Sociedad*, 11(21), 49-65. <https://doi.org/10.22430/21457778.1286>

- Monterroza-Ríos, A. D., & Novikova, O. (2020). La cultura como andamio de la agencia humana. *Pensando. Revista de Filosofía*, 11(23), 68-78.
- Monterroza, Á. (2018). *La naturaleza heterogénea de los artefactos técnicos: un análisis ontológico*. Fondo Editorial ITM.
- Monterroza-Ríos, A. D., & Escobar-Gómez, V. A. (2021). La educación tecnológica en Colombia: un marco epistémico para repensar un problema conceptual. *Trilogía Ciencia Tecnología Sociedad*, 13(25), 1-30. <https://doi.org/10.22430/21457778.1759>
- Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 21(4), 18-21. <https://ieeexplore.ieee.org/document/1667948>
- Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, 38(8), 114-117.
- Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.
- Moreno, J. C. (2019). Contribuciones al debate sobre la relevancia moral de los artefactos tecnológicos. *Trilogía Ciencia Tecnología Sociedad*, 11(21), 91-117. <https://doi.org/10.22430/21457778.1327>
- Moreno-Ortiz, J. C., & Vinck, D. (2021). Encuentros entre filosofía de la ciencia, filosofía de la tecnología y CTS. *Trilogía Ciencia Tecnología Sociedad*, 13(25), e1776. <https://doi.org/10.22430/21457778.1776>
- Moreno Ortiz, J. C. (2022). La ética de la técnica en Gilbert Simondon. En S. N. Osorio (Coord.), *Individuación y bioética global*. Implicaciones para la sostenibilidad humana y planetaria (pp. 159-185). Editorial Aula de Humanidades.
- Moreno Ortiz, J. C., Díaz Guzmán, O. D., Escobar Mejía, J. E., & Pérez Peña, N. C. (2024). *Ética de las tecnologías: análisis crítico de perspectivas*. Ediciones USTA. <https://doi.org/10.15332/li.lib.2022.00381>
- Morozov, E. (2015). *La locura del solucionismo tecnológico*. Katz.
- Mosterín, J. (1993). *Filosofía de la cultura*. Alianza.
- Mumford, L. (1934). *Technics and civilization*. Harcourt, Brace.
- Mumford, L. (1964). Authoritarian and democratic technics. *Technology and Culture*, 5(1), 1-8. <https://doi.org/10.2307/3101118>
- Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Harvard University Press.

- Newell, K. M. (1986). Constraints on the development of coordination. En M. G. Wade & H. T. A. Whiting (Eds.), *Motor development in children: Aspects of coordination and control* (pp. 341-360). Martinus Nijhoff.
- Neyrat, F. (2019). *The unconstructable Earth: An ecology of separation*. Fordham University Press.
- Niiniluoto, I. (1997). *Ciencia frente a tecnología: ¿diferencia o identidad?* *Arbor*, 157(620), 285-299. <https://doi.org/10.3989/arbor.1997.i620.1818>
- Norman, D. (2005). *El diseño emocional: por qué nos gustan (o no) los objetos cotidianos*. Paidós.
- Norman, D. A. (2010). *La psicología de los objetos cotidianos*. Nerea.
- Olivé, L. (2000). *El bien, el mal y la razón: facetas de la ciencia y de la tecnología*. Paidós.
- Olsen, J. K. B. (2009). Becoming through technology. En J. K. B. Olsen, E. Selinger, & S. Riis (Eds.), *New waves in philosophy of technology* (pp. 37-55). Palgrave Macmillan.
- ONU. (2015). *Objetivos de Desarrollo Sostenible*. <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>
- OpenAI. (2023). *GPT-4 technical report*. <https://arxiv.org/abs/2303.08774>
- Ortega y Gasset, J. (1914). *Meditaciones del Quijote*. Publicaciones de la Residencia de Estudiantes.
- Ortega y Gasset, J. (1982). *Meditación de la técnica y otros ensayos sobre la ciencia y la filosofía*. Alianza.
- Papanek, V. (1977). *Diseñar para el mundo real: ecología humana y cambio social*. H. Blume.
- Peterson, M. (2017). *The ethics of technology: A geometric analysis of five moral principles*. Oxford University Press.
- Platón. (1992). *Timeo; Critias*. Introducción, traducción y notas de Francisco Lisi. Editorial Gredos.
- Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. University of Chicago Press.
- Polanyi, M. (1967). *The tacit dimension*. Anchor Books.
- Popper, K. R. (1979). *Objective knowledge: An evolutionary approach*. Clarendon Press.
- Preston, B. (2009). Philosophical theories of artifact functions. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 213-233). Elsevier.

- Prigogine, I., & Stengers, I. (1990). *La nueva alianza* (M. García Valverde, Trad.). Alianza.
- Quintanilla, M. (2017). *Tecnología: un enfoque filosófico y otros ensayos de filosofía de la tecnología*. Fondo de Cultura Económica.
- Quintanilla, M. Á. (2002). La tecnología como paradigma de acción racional. *Isegoría*, (26), 67-88.
- Quintanilla, M. Á. (2012). Tecnologías entrañables. En M. Á. Quintanilla & A. Álvarez (Eds.), *Tecnología y sociedad*. OEI.
- Quintanilla, M. Á., Parselis, M., Sandrone, D., & Lawler, D. (2017). *Tecnologías entrañables: ¿es posible un modelo de desarrollo tecnológico democrático?* Los Libros de la Catarata.
- Rabardel, P. (2011). *Los hombres y las tecnologías: visión cognitiva de los instrumentos contemporáneos* (M. Acosta Gempeler, Trad.). Ediciones UIS.
- Radder, H. (2009). Why technologies are inherently normative. En A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 887-921). Elsevier.
- Raynaud, D. (2018). *¿Qué es la tecnología?* Laetoli.
- Ravelo Franco, N. A. (2023). Teoría crítica de la tecnología en Latinoamérica: educar sobre el acceso a electricidad. *Cuadernos de Filosofía Latinoamericana*, 44(128), 62-89. <https://doi.org/10.15332/25005375.8253>
- Redeker, R. (2018). *El eclipse de la muerte: La filosofía ante el desafío del transhumanismo*. Ediciones Encuentro.
- Rivera-Novoa, A. (2024). La tesis de la mente extendida y el ideal transhumanista de mejoramiento cognitivo. *Trilogía Ciencia Tecnología Sociedad*, 16(33), e3142. <https://doi.org/10.22430/21457778.3142>
- Robson, J., & Tsou, J. (Eds.). (2023). *Technology ethics: A philosophical introduction and readings*. Routledge.
- Rueda-Pardo, S. J., & Monterroza-Rios, A. D. (2025). Una crítica al transhumanismo cibernético-computacional de Kurzweil desde el enfoque enactivo de la cognición. *Cuadernos De Filosofía Latinoamericana*, 46(132), 101-115. <https://doi.org/10.15332/25005375.10773>
- Russell, S., & Norvig, P. (2002). *Artificial intelligence: A modern approach* (2nd ed.). Prentice Hall.
- Russell, S., & Norvig, P. (2002). *Artificial Intelligence: A Modern Approach 2nd edition*. Prentice Hall.

- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Ryle, G. (1949). *The concept of mind*. Hutchinson.
- Sadin, E. (2020). *La inteligencia artificial o el desafío del siglo* (M. Martínez, Trad.). Caja Negra.
- Schlosser, M. (2019). Agency. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/agency/>
- Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. Harper & Brothers.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457. <https://doi.org/10.1017/S0140525X00005756>
- Smith, M. R., & Marx, L. (Eds.). (1994). *Does technology drive history? The dilemma of technological determinism*. MIT Press.
- Smitsman, A. W., & Cox, R. F. A. (2008). Perseveration in tool use: A window for understanding the dynamics of the action-selection process. *Infancy*, 13(3), 249-269. <https://doi.org/10.1080/15250000802004379>
- Simondon, G. (2005). *L'individuation à la lumière des notions de forme et d'information*. Millon.
- Simondon, G. (2007). *El modo de existencia de los objetos técnicos* (M. Martínez & P. Rodríguez, Trans.). Prometeo.
- Simondon, G. (2014). *Sur la technique*. Presses Universitaires de France.
- Simondon, G. (2017). *On the mode of existence of technical objects* (C. Malaspina & J. Rogove, Trans.). Univocal Publishing.
- Simonton, D. K. (1999). *Origins of genius: Darwinian perspectives on creativity*. Oxford University Press.
- Skolimowski, H. (1966). The structure of thinking in technology. *Technology and Culture*, 7(3), 371-383.
- Sloterdijk, P. (2009). *You must change your life*. Suhrkamp.
- Sousa, C., Okamoto, S., Matsuzawa, T., & Biro, D. (2009). Behavioral development of nut-cracking in wild chimpanzees (*Pan troglodytes*). *Developmental Science*, 12(5), 703-711. <https://doi.org/10.1111/j.1467-7687.2009.00824.x>
- Staudenmaier, J. M. (1985). *Technology's storytellers: Reweaving the human fabric*. MIT Press.
- Stiegler, B. (1998). *Technics and time, 1: The fault of Epimetheus*. Stanford University Press.
- Takeuchi, N. (2016). *Nanotecnología*. Fondo de Cultura Económica.

- Taton, R. (1972). *Historia general de las ciencias* (Vol. II). Ediciones Destino.
- Tavani, H. T. (2011). *Ethics and technology: Controversies, questions, and strategies for ethical computing* (3rd ed.). Wiley.
- Tirado, F. J., & Domènech, M. (2005). Asociaciones heterogéneas y actantes: el giro postsocial de la teoría del actor-red. *AIBR. Revista de Antropología Iberoamericana*, (Ed. Electrónica).
- Toffler, A. (1970). *Future shock*. Random House.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>
- Turkle, S. (2017). *Alone together*. Basic Books.
- Van de Poel, I., & Royakkers, L. (2023). *Ethics, technology and engineering: An introduction* (2nd ed.). Wiley-Blackwell.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., & Polosukhin, I. (2017). *Attention is all you need*. In Advances in neural information processing systems (pp. 5998-6008). <https://doi.org/10.48550/arXiv.1706.03762>
- Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency, and design*. Penn State University Press.
- Verbeek, P.-P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.
- Verbeek, P.-P. (2015). Beyond interaction: A short introduction to mediation theory. *Interactions*, 22(3), 26-31. <https://dl.acm.org/doi/10.1145/2751314>
- Verbeek, P.-P., & Tijink, D. (2020). *Guidance ethics approach: An ethical dialogue about technology*. ECP Platform for the Information Society.
- Vermaas, P., Kroes, P., van de Poel, I., Franssen, M., & Houkes, W. (2011). *A philosophy of technology: From technical artefacts to sociotechnical systems*. Morgan & Claypool.
- Vincenti, W. G. (1993). *What engineers know and how they know it*. Johns Hopkins University Press.
- Vita-More, N. (2020). Transhumanism and design. En N. Lee (Ed.), *The transhumanism handbook*. Springer.
- Wallach, W., & Allen, C. (2009). *Moral machines: Teaching robots right from wrong*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195374049.001.0001>
- Weisbuch, G. (1991). *Complex systems dynamics: An introduction to automata networks*. Addison-Wesley.

- Winner, L. (1980). *Do artifacts have politics?* *Daedalus*, 109(1), 121-136.
- Wikipedia contributors. (2024, June 6). *Artificial intelligence*. Wikipedia. https://en.wikipedia.org/wiki/Artificial_intelligence
- World Design Organization. (2022, February). *Definition of industrial design*. <https://wdo.org/>
- Young, M. T. (2018). Intuition and ineffability: Tacit knowledge and engineering design. In A. Fritzsche & S. J. Oks (Eds.), *The future of engineering: Philosophical foundations, ethical problems and application cases* (pp. 53-67). Springer International Publishing.
- Zawidzki, T. W. (2018). *Mindshaping: A new framework for understanding human social cognition*. MIT Press.
- Ziman, J. (2000). *Technological innovation as an evolutionary process*. Cambridge University Press.
- Zubiri, X. (1962). *Sobre la esencia*. Alianza.
- Zubiri, X. (2010). *Cinco lecciones de filosofía*. Alianza.
- Zuboff, S. (2020). *The age of surveillance capitalism*. PublicAffairs.

Sobre los autores

Álvaro David Monterroza Ríos

Profesor del Departamento de Diseño del Instituto Tecnológico Metropolitano (Medellín). Editor de la revista *Trilogía Ciencia Tecnología Sociedad*. Co-coordinador de la Red Colombiana de Filosofía de la Tecnología (*PhiTec*). Doctor en Filosofía de la Universidad de Antioquia. Ingeniero Químico y magíster en ingeniería de la Universidad de Nacional de Colombia. Filósofo de la Universidad de Antioquia.

Correo electrónico: alvaromonterroza@itm.edu.co.

Orcid: <https://orcid.org/0000-0002-4326-5178>

Juan Carlos Moreno Ortizv

Profesor del Doctorado en Filosofía de la Facultad de Filosofía y Letras de la Universidad Santo Tomás (Bogotá). Editor de la revista *Cuadernos de Filosofía Latinoamericana*. Co-coordinador de la Red Colombiana de Filosofía de la Tecnología (*PhiTec*). Doctor en Filosofía de la Pontificia Universidad Javeriana. Especialista en Filosofía de la

Ciencia y la Tecnología de la Universidad del Bosque. Filósofo de la Universidad Javeriana.

Correo electrónico: juancmoreno@usta.edu.co.

Orcid: <http://orcid.org/0000-0003-4759-4398>

Jorge Francisco Maldonado Serrano

Profesor de la Escuela de Filosofía de la Universidad Industrial de Santander (UIS) (Bucaramanga). Doctor en Filosofía de la Universidad Autónoma de Madrid. Magíster Filosofía de la Pontifica Universidad Javeriana. Especialista en Docencia Universitaria de la Universidad El Bosque. Licenciado en Filosofía y Letras de la Universidad Santo Tomás.

Correo electrónico: jmaldona@uis.edu.co.

Orcid: <https://orcid.org/0000-0002-7707-154X>

John Jairo Cardozo Cardona

Director de la Maestría en Filosofía de la Tecnología de la Universidad Nacional Abierta y a Distancia (UNAD) (Bogotá). Doctor en conocimiento y cultura en América Latina del Instituto Pensamiento y Cultura en América Latina (México). Magister en Filosofía Latinoamericana y Filósofo de la Universidad Santo Tomás.

Correo electrónico: john.cardozo@unad.edu.co.

Orcid: <https://orcid.org/0000-0002-4815-5242>

Daian Tatiana Flórez Quintero

Profesora Departamento de Ciencias Humanas de la Universidad Nacional de Colombia (Manizales), y Profesora Asociada de la Universidad de Caldas. Directora del Departamento de Filosofía

de la Universidad de Caldas. Doctora en filosofía de la Universidad Nacional de Colombia. Magister y licenciada en filosofía en filosofía de la Universidad de Caldas.

Correo electrónico: daian.florez@ucaldas.edu.co.

Orcid: <https://orcid.org/0000-0002-9371-1850>

Diana María Muñoz González

Profesora del Departamento de Filosofía de la Universidad Nacional de Colombia (Bogotá). Doctora en Filosofía por la Universidad París 8 – St. Denis. Magíster en Filosofía y Filósofa de la Universidad Nacional de Colombia.

Correo electrónico: mmunozd@unal.edu.co.

Orcid: <https://orcid.org/0000-0003-1298-8805>

Dairon Alfonso Rodríguez Ramírez

Profesor de la Escuela de Filosofía de la Universidad Industrial de Santander (UIS) (Bucaramanga). Doctor en Humanidades (Filosofía) por la Universidad Autónoma Metropolitana. Magíster en Ciencias Cognitivas de la Universidad Autónoma del Estado de Morelos. Filósofo de la Universidad Industrial de Santander.

Correo electrónico: darodri@uis.edu.co.

Orcid: <https://orcid.org/0000-0002-5183-7121>

Índice temático y onomástico

- A**
affordances 63, 96, 97, 98
agencia técnica 226
- B**
Broncano 22, 26, 32, 33, 34, 35,
37, 38, 42, 48, 58, 66, 68, 129,
139, 226, 227
- C**
cambio técnico 17, 19, 51, 53, 54,
55, 56, 58, 60, 61
cambio tecnológico 32, 33, 52,
141, 226, 228
Coeckelbergh 210, 211, 227
concretización 53, 55, 108, 110,
111, 112, 153, 154, 159, 161
conocimiento técnico 74, 75, 76
conocimiento tecnológico 17, 27,
28, 71, 72, 89, 90, 91, 93, 94,
97, 99, 102, 103, 104, 105,
106, 114, 124, 222
constructivismo 30, 31, 149, 151,
230
contexto de diseño 21, 41, 44, 45,
46, 47, 67
contexto de uso 21, 41, 44, 45, 47,
48, 49, 50, 67
- D**
determinismo 31, 32, 33, 34, 35,
36, 67, 134, 137, 155, 169,
205, 215, 229, 230
- E**
Ellul 32, 33, 34, 128, 134, 141, 229
ética 17, 26, 40, 127, 128, 129, 130,
131, 132, 133, 143, 144, 145,
148, 150, 152, 153, 154, 155,
156, 158, 160, 168, 169, 170,
171, 173, 194, 200, 203, 204,
206, 207, 208, 209, 213, 214,
215, 217, 219, 222, 230, 236
ética crítica 232
ética de la IA 194, 203, 204, 206, 213

- F**
 Feenberg 27, 33, 41, 132, 134, 136, 142, 149, 151, 152, 206, 229
 Floridi 139, 140, 187, 195, 196, 207, 210, 211, 230
 Franssen 125, 144, 231, 240
 función 26, 29, 49, 58, 61, 114, 122, 131, 137, 152, 153, 154, 189, 190, 191, 192, 202, 212
 funcionamiento 63, 69, 83, 84, 85, 86, 88, 89, 97, 100, 109, 111, 112, 120, 153, 154, 159, 166, 187, 196, 199, 207, 212, 217
- G**
 georingeniería 161, 162, 164, 165, 166, 167, 168
 giro empírico 31, 148, 149, 150, 155
- H**
 Heidegger 65, 128, 134, 136, 141, 142, 143, 149, 232, 233
 humanismo 167
 humanista 129, 140, 141, 142, 153, 160, 228
- I**
 IA fuerte 218
 IA generativa 18, 182, 184, 204, 216, 223
 Ihde 31, 85, 119, 149, 157, 214, 232, 233
 imaginación inventiva 108, 111, 113
 individuación 53, 55, 107, 108, 110, 111, 152, 153, 154, 161
- instrumentalismo 133, 157, 168, 169, 205
 instrumentalización tecnológica 130, 135, 136, 139
- K**
 Kapp 106, 107, 140, 233
 Kroes 125, 149, 150, 155, 156, 211, 233, 240
 Kurzweil 26, 32, 53, 54, 185, 210, 233, 238
- L**
 Latour 30, 31, 48, 138, 140, 152, 155, 234
 Lawler 34, 234, 238
- M**
 Marcuse 26, 107, 134, 136, 141, 234, 235
 mediación 15, 138, 158
 Mitcham 90, 106, 107, 129, 132, 133, 134, 140, 141, 142, 149, 150, 171, 235
 Mumford 32, 33, 34, 65, 128, 134, 141, 236
- N**
 neutralidad 135, 147, 155
 neutralidad de las tecnologías 135, 147
- O**
 objeto técnico 107, 108, 109, 110, 111, 112, 113, 153

Ortega y Gasset 19, 22, 34, 41, 68,
134, 141, 237

P

pensamiento 17, 20, 35, 44, 65,
86, 101, 106, 107, 108, 113,
141, 175, 176, 177, 179, 231
pensamiento creativo 101
pensamiento técnico 65, 108, 113
posibilismo tecnológico 32
postfenomenología 149, 214
principio ético 194

Q

Quintanilla 24, 27, 34, 46, 47, 64,
65, 66, 68, 100, 101, 238

R

Radder 84, 85, 238
relevancia moral de los artefactos
154, 236

S

Simondon 22, 34, 53, 55, 101, 106,
107, 108, 109, 110, 111, 113,
114, 152, 153, 154, 159, 160,
232, 236, 239
singularidad tecnológica 210
solucionismo tecnológico 160,
161, 165, 168, 170, 236
Stiegler 22, 34, 50, 152, 240
sustantivismo 133, 142, 169, 205

T

técnico 17, 18, 19, 21, 24, 26, 35,
41, 44, 51, 53, 54, 55, 56, 58,

60, 61, 65, 73, 74, 75, 76, 90,
100, 107, 108, 109, 110, 111,
112, 113, 141, 145, 147, 152,
153, 159, 160, 174, 178, 179,
180

tecnologías entrañables 238

tecnológico 16, 17, 20, 25, 26, 27,
28, 31, 32, 33, 34, 35, 36, 37,
38, 39, 43, 44, 45, 47, 51, 52,
53, 54, 62, 63, 66, 67, 71, 72,
89, 90, 91, 93, 94, 97, 98, 99,
102, 103, 104, 105, 106, 109,
114, 124, 128, 132, 133, 134,
135, 136, 137, 139, 141, 142,
143, 145, 146, 148, 151, 155,
159, 160, 161, 165, 166, 167,
168, 170, 215, 216, 222, 223,
226, 228, 230, 236, 238

teoría actor-red 155, 207, 213

teoría crítica de la tecnología 41,
149

transhumanismo 161, 163, 164,
165, 166, 167, 168, 238

V

Van de Poel 129, 171, 240

Verbeek 129, 140, 142, 143, 155,
156, 157, 158, 211, 214, 233,
240

Vermaas 82, 85, 86, 91, 99, 125, 240

voluntarismo 133, 157, 168, 169,
205

W

Winner 27, 241



Esta obra se editó en Ediciones USTA.
2026



La filosofía de la tecnología es indispensable para comprender cómo nuestras prácticas se reconfiguran en entornos mediados por sistemas técnicos. A menudo, el debate público reduce la tecnología a simples objetos o aplicaciones de la ciencia, ignorando su profundo impacto en la sociedad.

Filosofía de la tecnología: una introducción conceptual ofrece un mapa riguroso para superar estos reduccionismos. A través de tres ejes —ontológico, epistemológico y ético-político—, la obra analiza la naturaleza del diseño tecnológico, el conocimiento específico que moviliza y los criterios normativos para su evaluación. Además, aplica este marco a los desafíos de la inteligencia artificial, demostrando que el análisis filosófico de la técnica es inseparable de la comprensión de las sociedades contemporáneas. Una obra imprescindible de Ediciones USTA para estudiantes e investigadores.



UNIVERSIDAD
SANTO TOMÁS

