

## **Anne Foerst. El encuentro entre teología e inteligencia artificial**

### ***Anne Foerst. The encounter between theology and artificial intelligence***

**Francisco J. Génova**

Centro Regional de Estudios Teológicos de Aragón (C.R.E.T.A.)

Recibido: 28 agosto 2017

Aceptado: 17 octubre 2017

*Resumen:* La inteligencia artificial (IA) ha llegado para quedarse, nos rodea y se hace presente en nuestros trabajos, nuestros hogares y en nuestras relaciones personales. El futuro de la humanidad va a ser modelado por las tecnologías emergentes, entre las cuales la IA es un punto de convergencia privilegiado de todos los desafíos que la tecnología de nuestro mundo le presenta a la teología. Una teóloga que ha aceptado este desafío es Anne Foerst. Ella ha planteado la cuestión de si en algún momento los robots podrían llegar a ser considerados personas y formar parte de nuestras comunidades. Aunque su punto de vista no sea compartido debemos reconocerle su importante contribución a una teología de frontera. Una teología encarnada en las fronteras tecnológicas necesita ser valiente, necesita aceptar el marco que ofrecen la ciencia y la tecnología para explorar nuevos lenguajes.

*Palabras clave:* Ciencia ficción, inteligencia artificial, robots, singularidad, teología.

*Abstract:* Artificial Intelligence (AI) is coming to stay. AI surrounds us and penetrates our jobs, our homes, and our relationships. The future of humanity is going to be shaped for emergent technologies, among them AI is a privileged point of convergence of all challenges that must face theology in the technological present of our world. A theologian that has accepted the AI challenge is Anne Foerst. She has raised the question whether some robots could be at some point fellow persons and be part of our communities. Even when her point of view is not shared we must recognize her important contribution to a theology at the boundary. A theology embodied in the final technological frontiers needs to be brave; it needs to accept the frame of science and technology for exploring new languages.

*Keywords:* Artificial intelligence, robots, science fiction, singularity, theology.

## 1. INTRODUCCIÓN: EL DESAFÍO DE LA INTELIGENCIA ARTIFICIAL

En 1950 el matemático Alan Turing planteó en un artículo la cuestión ¿pueden pensar las máquinas?<sup>1</sup> Con esta pregunta introdujo de lleno en el campo académico la preocupación por las máquinas inteligentes.<sup>2</sup> Unos años más tarde, en 1957, el también matemático Irving John Good fue el primero en hablar de una *explosión de inteligencia* que se produciría con la primera máquina ultrainteligente que el hombre llegase a crear, y que sería lo último que el ser humano necesitaría inventar.<sup>3</sup> Pero Good reflejaba también la posibilidad de que llegemos a una situación en la que seamos incapaces de controlar nuestras creaciones crecientemente complejas. Una situación de imprevisibles consecuencias que el matemático y escritor de ciencia ficción Vernor Vinge definió con un nombre ya acuñado anteriormente por John von Neumann:<sup>4</sup> Singularidad.<sup>5</sup> Una singularidad tecnológica que daría a luz una nueva e imprevisible realidad.

<sup>1</sup> A. M. Turing, "Computing Machinery and Intelligence", *Mind* 49 (1950) 433-460.

<sup>2</sup> Todavía no se utilizaba la denominación de inteligencia artificial (IA), que fue introducida en el verano de 1956 por John McCarthy en la conferencia de Dartmouth. Y que no lo fue con consenso, ya que para muchos era, y sigue siendo, ambigua. Sobre todo por la palabra artificial, ante la cual se prefiere las denominaciones alternativas de *Machine Intelligence* o *Cognitive Computing*, poniendo el acento en que la inteligencia de las máquinas será inteligencia pero diferente a la humana. Como diferente es un pájaro de un avión, o un pez de un submarino. Una amplia definición de IA nos la ofrecen Keith Frankish y William M. Ramsey: "Artificial intelligence is a cross-disciplinary approach to understanding, modeling, and replicating intelligence and cognitive processes by invoking various computational, mathematical, logical, mechanical, and even biological principles and devices. [...] Historically, its practitioners have come from such disciplines as logic, mathematics, engineering, philosophy, psychology, linguistics, and, of course, computer science." (K. Frankish W. M. Ramsey, "Introduction", en: K. Frankish W. M. Ramsey (dirs.), *The Cambridge Handbook of Artificial Intelligence*, Cambridge 2014, 1-11, p. 1). Margaret Boden nos ofrece una más breve: "Artificial Intelligence seeks to make computers do the shorts of things that minds can do." (M. A. Boden, *AI. Its nature and future*, Oxford 2016, 1).

<sup>3</sup> I. J. Good, "Speculations Concerning the First Ultraintelligent Machine", *Advances in Computers* 6 (1965) 31-88, p. 33.

<sup>4</sup> M. Ford, *The Rise of the Robots. Technology and the Threat of Mass Unemployment*, London 2015, 229.

<sup>5</sup> V. Vinge, "The Coming Technological Singularity: How to survive in the Post Human Era", en: M. More – N. Vita-More (dirs.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester-UK 2013, 365-375.

Esta preocupación es la que llevó a que Stephen Hawking, junto a otros, firmara en 2014 una carta abierta advirtiendo de los riesgos de la inteligencia artificial (IA) para el futuro de la humanidad si permitimos que ella tome el control y decida por nosotros.<sup>6</sup> No se trata de un escenario de ciencia ficción en el que nuestras máquinas inteligentes han adquirido conciencia y se rebelan. Se trata de algo más prosaico y real, de nuestra capacidad para controlar una complejidad creciente contra la que ya en 1996 alertó Blay Whitby, advirtiendo de cómo los nuevos caminos recorridos por la IA, tales como redes neuronales y algoritmos genéticos, provocaban un efecto “caja negra”, con el que obtenemos resultados sin tener acceso al proceso completo por el que se han producido.<sup>7</sup>

Las actuales y exitosas líneas de *Machine Learning*,<sup>8</sup> un verdadero renacer de la IA, continúan por esa senda.<sup>9</sup> Ejemplo de

<sup>6</sup> <<http://futureoflife.org/ai-open-letter/>> [Consulta: 14 ago. 2017]. Bill Gates también se manifestó en ese sentido en 2015 (<<http://observer.com/2015/08/stephen-hawking-elon-musk-and-bill-gates-warn-about-artificial-intelligence/>> [Consulta: 14 ago. 2017]). En enero de 2017 la Conferencia de Asilomar aprobó veintitrés principios para garantizar una IA que no perjudique a la humanidad. Entre otros muchos Stephen Hawking y Elon Musk los han suscrito (< <https://futureoflife.org/ai-principles/>> [Consulta: 14 ago. 2017]).

<sup>7</sup> B. Whitby, *Reflections on Artificial Intelligence: The legal, moral and ethical dimensions*, Bristol 1996, 29-30.

<sup>8</sup> En los últimos veinte años el resurgir de la IA se ha debido al *Machine Learning*, tras las limitaciones mostradas por la IA clásica basada en algoritmos que especificaban las instrucciones precisas de lo que debía hacer el ordenador, se ha producido una eclosión de algoritmos que aprenden de los datos que se les proporcionan a cumplir sus metas. Para una introducción general E. Alpaydin, *Machine Learning. The New AI*, Cambridge 2016.

<sup>9</sup> Un ejemplo es cómo en marzo de 2016 el programa de IA AlphaGo, de DeepMind (empresa británica adquirida por Google en 2014 por 600 millones de dólares) venció en el juego oriental Go, al campeón mundial Lee Sedol, el cual ostentaba el mayor nivel profesional, el nueve, del juego. Este juego es de complejidad y nivel de estrategia mucho mayor que el ajedrez, las combinaciones exceden el número de átomos en el universo, y para vencer el programa tuvo que utilizar algo similar a la intuición humana. Philip Larrey, de la Universidad Pontificia Lateranense, no duda en afirmar que con AlphaGo “we have seen, perhaps for the first time, a form of intelligence that lacks consciousness” (P. Larrey, *Connected World. From Automated Work to Virtual Wars: The Future, by Those Who are Shaping It*, London 2017, 6). En junio de 2017 se ha conocido otro logro de la IA, en este caso por parte de Microsoft, que ha desarrollado un sistema de IA que ha conseguido en el juego clásico de Atari “Ms. Pac-Man”, lo que hasta ahora ningún ser humano o IA había podido, un marcador que alcanzó los 999,990 puntos. <<https://>

lo cual es el incidente conocido en julio de 2017 sobre dos bots conversacionales programados por personal del *Facebook Artificial Intelligence Research* (FAIR) para ser buenos negociadores.<sup>10</sup> Se trata de algoritmos que debían aprender con cada negociación. Sus creadores decidieron dejar que hablaran el uno con el otro, y cuando más tarde comprobaron su estado descubrieron que para comunicarse entre ellos habían desarrollado su propio lenguaje. Fueron desconectados. Los algoritmos desarrollaron aquello para lo que estaban diseñados, pero los científicos responsables no fueron capaces de prever las consecuencias. Es lo que Stuart Armstrong explica diciendo que el riesgo no es que la IA pueda acabar siendo malvada, sino que los fines de la misma, inicialmente inocuos, acaben siendo peligrosos al aumentar su poder.<sup>11</sup>

Un ejemplo que ofrece Nick Bostrom<sup>12</sup> es el “maximizador de clips para hojas de papel”,<sup>13</sup> una hipotética máquina superinteligente programada para que el único fin de su existencia sea producir clips, la cual sin llevar a cabo ningún acto consciente de rebelión, se limitaría a transformar todo lo que encuentre, incluido el ser humano, en clips. Eliezer Yudkowsky,<sup>14</sup> defensor de garantizar la creación de una IA amistosa (*friendly AI*), lo resume brevemente en una frase: “The AI neither hates you, nor loves

[blogs.microsoft.com/ai/2017/06/14/divide-conquer-microsoft-researchers-used-ai-master-ms-pac-man/#sm.0001sbtudpsiadurspw1mfju2qc2g](https://blogs.microsoft.com/ai/2017/06/14/divide-conquer-microsoft-researchers-used-ai-master-ms-pac-man/#sm.0001sbtudpsiadurspw1mfju2qc2g) [Consulta: 14 ago. 2017].

<sup>10</sup> <<http://www.telegraph.co.uk/technology/2017/08/01/facebook-shuts-robots-invent-language/>> [Consulta: 14 ago. 2017]; <<http://www.techly.com.au/2017/07/31/facebooks-ai-bots-are-communicating-in-a-language-we-dont-understand/>> [Consulta: 14 ago. 2017].

<sup>11</sup> S. Armstrong, “Introduction to the Technological Singularity”, en: V. Callaghan – J. Miller – R. Yampolskiy – S. Armstrong (dirs.), *The Technological Singularity. Managing the Journey*, Berlín 2017, 1-8, 5-6.

<sup>12</sup> Filósofo transhumanista. Profesor de la Universidad de Oxford. Fundador y director del *Future of Humanity Institut* de dicha universidad. Él advierte del peligro de una explosión de inteligencia para la que nos estamos preparando: “we humans are like small children playing with a bomb. Such is the mismatch between the power of our plaything and the immaturity of our conduct. [...] here is not one child but many, each with access to an independent trigger mechanism. [...] Some little idiot is bound to press the ignite button just to see what happens” (N. Bostrom, *Superintelligence. Paths, Dangers, Strategies*, Oxford 2014, 259).

<sup>13</sup> N. Bostrom, *Superintelligence...*, 122-125. <<http://www.nickbostrom.com/ethics/ai.html>> [Consulta: 14 ago. 2017].

<sup>14</sup> Fundador del *Machine Intelligence Research Institut* (MIRI).

you, but you are made out of atoms that it can use for something else".<sup>15</sup>

Sin embargo también hay quienes como Ray Kurzweil,<sup>16</sup> contratado en 2012 por Google como director de ingeniería en el área de IA, defienden el camino de una singularidad tecnológica en la que nuestras máquinas nos superen y nosotros mismos nos fundamos con ellas para vencer así todas nuestras limitaciones, incluida finalmente la muerte. Para Kurzweil "the Singularity will represent the culmination of the merger of our biological thinking and existence with our technology, resulting in a world that is still human but that transcends our biological roots."<sup>17</sup> Un escenario que el antropólogo norteamericano Robert Geraci ha denominado *inteligencia artificial apocalíptica*, y en el cual él ha descubierto una dimensión claramente religiosa que le lleva a volver la mirada hacia la teología y plantearle una pregunta y una tarea.<sup>18</sup> La pregunta es ¿qué significado tiene para la teología cristiana los esfuerzos por construir robots inteligentes? Y la tarea, impulsar una teología de la robótica; aún cuando los teólogos rechacen las pretensiones y planteamientos de esa IA apocalíptica.

Si comenzábamos con la pregunta de Turing, ¿pueden pensar las máquinas?, podemos ahora plantearnos la pregunta de si realmente todo esto tiene algo que ver con la teología. ¿Tiene interés la IA para la teología? No somos los primeros en preguntárnoslo. George Gilder y Jay Richards creen que sí: "although Artificial Intelligence may seem like an esoteric topic with little relevance to anything else, in fact, many of the most important questions we face from technology to theology converge on this single subject".<sup>19</sup> Fraser Watts es de la misma opinión. Cree él que

<sup>15</sup> E. Yudkowsky, "Artificial Intelligence as a positive and negative factor in global risk", en: N. Bostrom – M. M. Cirkovic (dirs.), *Global Catastrophic Risks*, Oxford 2008, 308-345, p. 333.

<sup>16</sup> Un prolífico inventor que en 2008 fundó junto a Peter Diamandis y el NASA Research Park de California, y el apoyo de Google, la Universidad de la Singularidad (*Singularity University*) en Silicon Valley. Esta universidad, que no ofrece grados oficiales, pretende formar a las élites que han de impulsar y desarrollar los cambios tecnológicos que nos han de dirigir hacia la Singularidad. < <https://su.org/> [Consulta: 14 ago. 2017].

<sup>17</sup> R. Kurzweil, *The Singularity is Near. When Humans Transcend Biology*, New York 2006, 9.

<sup>18</sup> R. Geraci, *Apocalyptic AI. Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*, New York 2010, 5.

<sup>19</sup> G. Gilder – J. Richards, "Are We Spiritual Machines? The Beginning of a Debate", en: J. W. Richards (dir.), *Are We Spiritual Machines? Ray Kurzweil vs. The Critics of Strong AI*, Seattle 2002, 1-11, p. 3.

pocos temas científicos plantean cuestiones religiosas tan importantes como la IA, y sin embargo son tan poco abordados.<sup>20</sup> Para Juan L. Ruiz de la Peña la IA es un desafío antropológico convertido en verdadero reto teológico.<sup>21</sup>

La IA es la frontera tecnológica donde las demás convergen y se yerguen desafiantes sobre la idea misma de lo que es el ser humano y lo que puede llegar a ser en el futuro. Es por tanto una frontera que reclama la atención de una teología que se atreva a enfrentarse a la pregunta y a la tarea que marcaba Robert Gera-ci. Una teóloga que ha asumido el reto es Anne Foerst.<sup>22</sup>

## 2. ANNE FOERST. UNA TEÓLOGA EN EL CORAZÓN DE LA INTELIGENCIA ARTIFICIAL

Anne Foerst es una teóloga luterana alemana actualmente profesora asociada<sup>23</sup> en la Universidad de San Buenaventura (*St. Bonaventure University*) en Olean, Nueva York. A la que se trasladó tras verse obligada a abandonar el laboratorio de IA del Instituto Tecnológico de Massachusetts (MIT) en 2001, donde desde 1995 trabajaba como investigadora asociada<sup>24</sup> gracias al apoyo de Rodney Brooks.<sup>25</sup> Foerst es una teóloga que se ha situado en

<sup>20</sup> F. Watts, "Artificial Intelligence", en: R. L. Herrmann (dir.), *God, Science and Humility. Ten Scientists Consider Humility Theology*, Philadelphia 2000, 279-300.

<sup>21</sup> J. L. Ruiz de la Peña, *Crisis y apología de la fe. Evangelio y nuevo milenio*, Santander 1995, 155-209.

<sup>22</sup> "As our technical creatures become more like us, they raise fundamental theological questions. As theologians, we have the responsibility not to shrink away from such a challenge." (A. Foerst, *God in the Machine. What Robots Teach Us About Humanity and God*, New York 2004, 6).

<sup>23</sup> En el momento de escribir este artículo se encuentra inmersa, según ella misma nos ha indicado, en el proceso de promoción a profesora titular (*Full Professor*) de dicha universidad.

<sup>24</sup> A su formación teológica añadía sus estudios de informática y filosofía en la Universidad de Bonn. Después de los cuales realizó estudios teológicos en el seminario luterano de Renania (Alemania), y se doctoró en teología en 1996 en la Universidad Ruhr de Bochum (Alemania).

<sup>25</sup> Rodney Brooks es un experto internacionalmente reconocido en robótica, que revolucionó este campo al buscar caminos nuevos para la IA en una aproximación desde formas de vida más sencillas como los insectos. Ha sido director del laboratorio de IA del Instituto Tecnológico de Massachusetts y una de las referencias principales a nivel mundial en IA

las fronteras tecnológicas de nuestro mundo y ha encarnado vitalmente el encuentro entre teología e IA. Para hacerlo se ha colocado en línea con la teología de Paul Tillich,<sup>26</sup> adoptando dos claves extraídas de su pensamiento: *walking the boundary*<sup>27</sup> (caminar en la frontera) y *the courage to doubt* (el coraje para dudar), parafraseando *the courage to be*<sup>28</sup> de Tillich.

Su caminar en la frontera lo reflejan muy bien las dificultades sufridas en su paso por el MIT y por la *Harvard Divinity School*<sup>29</sup> (HDS). Con un pie en la teología y otro en la IA se sintió incomprendida en los dos lados de esa frontera por la que ella caminaba: “People at HDS found my fascination with high-tech somewhat strange. [...] At MIT, on the other hand, people were suspicious of the theologian in their midst.”<sup>30</sup> En el MIT la incompreensión llegó al explícito rechazo. De hecho el motivo de

y robótica. Fundador de empresas como iRobot y Rethink Robotics. La primera ha lanzado y popularizado los ya comunes robots aspiradores, y la segunda ha abierto una nueva forma de entender los robots industriales, de lo que es un claro ejemplo el robot Baxter (<<http://www.rethinkrobotics.com/baxter/>> [Consulta: 14 ago. 2017]), capaz de trabajar junto a los operarios humanos y con un sistema de programación que no necesita ningún conocimiento informático.

<sup>26</sup> Fue precisamente trabajando en los archivos de Tillich en la *Harvard Divinity School* cómo tomó contacto con el MIT, al descubrir que de las seis conferencias que Tillich había dado en el MIT, solo dos habían sido publicadas.

<sup>27</sup> Paul Tillich lo manifiesta así: “I thought that the concept of the boundary might be the fitting symbol for the whole of my personal and intellectual development” (P. Tillich, *On the Boundary. An Autobiographical Sketch*, Eugene-OR 2011 (1966), 13).

<sup>28</sup> Paul Tillich lo definía así: “The courage to be is the ethical act in which man affirms his own being in spite of those elements of his existence which conflict with his essential self-affirmation” (P. Tillich, *The Courage to Be*, New Haven & London 2000 2 ed. (1952), 3). Y en la siguiente cita encontramos la clara línea de continuidad con Foerst: “*The courage to be is rooted in the God who aears when God has disaeared in the anxiety of doubt*” (P. Tillich, *The Courage...*, 190), la cursiva pertenece al original.

<sup>29</sup> En la Harvard Divinity School contó con el apoyo del profesor Harvey Cox, trabajando como afiliada del *Center for the Study of Values in Public Life*. Con Cox publicó Foerst un artículo: H. Cox – A. Foerst, “Religion and Technology: A New Phase”, *Bulletin of Science, Technology and Society* 17 (1997) 53-60. Y participó con otro en el libro homenaje a Cox con motivo de su jubilación: A. Foerst, “Commander Data. A Candidate for Harvard Divinity School”, en: A. Sharma (dir.), *Religion in a Secular City. Essays in Honor of Harvey Cox*, Harrisburg-PA 2001, 263-281.

<sup>30</sup> A. Foerst, *God in the Machine...*, 5. Este libro, *God in the Machine*, es hasta este momento el único que ha publicado.

su salida del MIT fue, según ella, que los poderes fácticos de la institución no estaban preparados para ofrecerle ahí una plaza como profesora.<sup>31</sup>

Pero a pesar de las dificultades su estancia en el MIT le permitió acercarse a un mundo, el de la IA y sus investigadores, que iba a marcar todo su pensamiento. En 1997 puso en marcha en el MIT el curso *God and Computers*, del que ella comenta: “the purpose of the course is to help artificial intelligence pursue its goal by making researchers aware of their unscientific biases”.<sup>32</sup> Pero más importante que el curso<sup>33</sup> en sí mismo es el rechazo que provocó entre amplios sectores de alumnos y profesores del MIT. Un ejemplo es el artículo del alumno J. Ryan Bender:

The absurd babble coming from the Christian mystics has absolutely nothing to do with the problems faced by artificial intelligence researchers. These mystics, possessed by faith and suspicious of reason, have nothing useful to contribute to the AI endeavor except the mindless drivel of religion. *God and Computers* has no place in Course VI and is an insult to MIT. This Institute, one of the last bastions of rational thought, ought to leave the evangelists on the steps of 77 Massachusetts Avenue.<sup>34</sup>

Pero el rechazo no fue la única dificultad a la que Foerst hizo frente. Como ella misma reconoce. La tensión entre la visión teológica del ser humano que ella traía consigo y la visión mecanicista del laboratorio de IA fue algo que le afectó personalmente e incluso cuestionó su propia fe.<sup>35</sup> Pero tal como ella comenta encontró su fuerza precisamente en la duda que le ayudaba a comprender los argumentos de la otra parte:

<sup>31</sup> Así se lo indicó a la periodista Claudia Dreifus de *The New York Times*: “At M.I.T., the powers-that-be didn’t seem ready to offer her a professorship, she told me in a phone interview in March of 2001” (C. Dreifus, “Anne Foerst. Do Androids Dream? M.I.T. working on it”, en: C. Dreifus, *Scientific Conversations. Interviews on Science from The New York Times*, New York 2001, 103-107, p. 107).

<sup>32</sup> A. Foerst, “Bender Column Missed Point of 6.915”, *The Tech* 117 n.42 (1997) 4. El número 6.915 es el identificativo del curso *God and Computers*.

<sup>33</sup> Para detalles sobre el mismo A. Foerst, “Foreword: Meeting God at MIT”, en: D. E. Knuth, *Thinks a Computer Scientist Rarely Talk About*, Stanford 2001, v-ix.

<sup>34</sup> J. R. Bender, “Taking on God in Course VI”, *The Tech* 117 n.40 (1997) 5. 77 *Massachusetts Avenue* es la dirección del MIT.

<sup>35</sup> A. Foerst, *God in the Machine...*, 7.

Because these two worldviews clashed, I was constantly waking the boundary between them. What I learned at this time was that doubt is a positive thing. If you enter a dialogue between theology and another discipline and don't think the partner in dialogue has a valid perspective, you might as well not talk at all. Whatever the other part is saying will never convince you or even influence you in any way. If, on the other hand, you throw yourself wholeheartedly into a discussion, you have to question yourself constantly.<sup>36</sup>

### 3. EL ENCUENTRO ENTRE TEOLOGÍA E INTELIGENCIA ARTIFICIAL

#### a) El diálogo entre la teología y la inteligencia artificial: aproximación simbólica

La búsqueda de encuentro entre teología e IA que caracteriza el trabajo y la vida de Foerst se expresa simbólicamente en un apretón de manos que ella propició. El que tuvo lugar en 1995 en el MIT entre el teólogo Harvey Cox de la HDS y el robot Cog del laboratorio de inteligencia artificial del MIT. No es extraño que Foerst comience su libro *God in the Machine* con las siguientes palabras “this is the story of a handshake”.<sup>37</sup> Puede decir Foerst que su libro es la historia de un apretón de manos porque ese momento condensa lo que ella ha pretendido y sigue pretendiendo, un diálogo entre teología e IA a través de lo que ella ha denominado aproximación simbólica.

Para Foerst la teología necesita tener el “coraje de dudar”. Un coraje lleno de humildad que la ha de liberar de cualquier pretensión de saberlo todo, y la ha de llenar de la necesidad de aprender de otras disciplinas, yendo a su encuentro no como maestra sino como compañera. No con criterios para juzgar, sino con el deseo de aprender y buscar junto a otros la luz. De este modo se ha planteado ella el encuentro entre teología e IA, como un diálogo capaz de tender puentes en el más amplio campo de la relación entre religión y ciencia. Diálogo para el cual ella realiza una propuesta concreta: el método de la “aproximación simbólica” (*Symbolic Approach*).<sup>38</sup>

<sup>36</sup> A. Foerst, *God in the Machine...*, 7.

<sup>37</sup> A. Foerst, *God in the Machine...*, 1.

<sup>38</sup> La fuente donde nuestra autora detalla con más claridad esta propuesta es en un artículo publicado en la revista *Zygon*: A. Foerst, “Cog, a

El método de la aproximación simbólica plantea un marco epistemológico alternativo al actual caracterizado para Foerst por una aproximación cartesiana a la realidad. Ante la que nuestra autora plantea una descripción de la misma que la entienda como simbólica y permita un verdadero diálogo entre ciencia y religión que sustituya a los modelos actuales. Para Foerst el marco clásico del diálogo entre ciencia y religión, en la línea de Ian G. Barbour<sup>39</sup> y John F. Haught,<sup>40</sup> asume como válida la idea de que es posible acercarse al mundo como una realidad independiente que puede ser observada y comprendida. Ante esto propone su aproximación simbólica como una vía media entre el cartesianismo y el constructivismo,<sup>41</sup> que acepta la validez de un acercamiento científico a la realidad, pero asume también en esta realidad la presencia de metáforas y mitos a los que no es ajena la propia empresa científica.<sup>42</sup> Según Foerst “both scientific and theological descriptions of reality can be seen as partly constructed and dependent on our perceptual apparatus and maybe even our personality”.<sup>43</sup> Para ilustrar esta complementariedad de visiones Foerst recurre a la psicología de la Gestalt<sup>44</sup> con sus representaciones que permiten dos comprensiones diferentes de una misma y única imagen.

Foerst es consciente de que los presupuestos antropológicos de la IA presentan una concepción meramente mecanicista del ser humano, del cual máquina y ordenador se convierten en metáfora. Pero, precisamente, en la búsqueda por comprender al ser humano que la IA asume para poder “reproducirlo”

Humanoid Robot, and the Question of the Image of God”, *Zygon. Journal of Religion and Science* 33 (1998) 91-111.

<sup>39</sup> El propio Ian Barbour cita a Foerst y resume lo que ella pretende así: “She calls for dialogue and mutual respect between theologians and computer scientists and recognition of the biases and limits of each discipline.” (I. Barbour, *Nature, Human Nature, and God*, London 2002, 86).

<sup>40</sup> John F. Haught ve cuatro tipos posibles de relación entre ciencia y religión que enumera así: conflicto, contraste, contacto y confirmación (J. F. Haught, *Science and Religion. From Conflict to Conversation*, Mahwah-New Jersey 1995).

<sup>41</sup> A. Foerst, “Cog, a Humanoid Robot...”, 95.

<sup>42</sup> Foerst lo expresa así: “following Kuhn, I can express my main thesis by saying that researchers in any scientific field include their beliefs and hidden assumptions in their scientific worldview” (A. Foerst, “Artificial Intelligence: Walking the Boundary”, *Zygon. Journal of Religion and Science* 31 (1996) 681-693, p. 683).

<sup>43</sup> A. Foerst, “Cog, a Humanoid Robot...”, 96.

<sup>44</sup> A. Foerst, *God in the Machine...*, 18-22.

artificialmente, encuentra Foerst el punto de encuentro con la teología. Aun partiendo de presupuestos distintos, teología e IA buscan comprender al ser humano, y ese es, para Foerst, su punto de encuentro.<sup>45</sup>

b) *El ser humano en Foerst: homo narrans*

Para responder a la pregunta por el ser humano Foerst vuelve la mirada hacia el relato del Génesis<sup>46</sup> y se apoya en la teología de Paul Tillich.<sup>47</sup> En el Génesis encuentra nuestra autora la fundamentación para afirmar por un lado la unión del ser humano con el resto de la creación y en especial con los animales, que, como él, han participado en el proceso evolutivo; y que para Foerst sitúa al ser humano sin ninguna característica que por sí misma lo haga especial. Y por otro, la elección que Dios hace del ser humano como compañero y amigo. Estableciendo una relación, esta sí especial, con él. Esta clave “relacional” va a ser fundamental en su pensamiento. Es una elección para la relación. Y es para Foerst esa relación, ese interactuar con Dios, lo que nos configura como imagen suya.<sup>48</sup> Dentro de esa imagen que Foerst entiende desde la relación, sitúa nuestra autora la capacidad creativa del ser humano. Siguiendo a Philip Hefner para ella el ser humano es creado co-creador: “when we attempt to re-create ourselves, we do God’s bidding. We help God. We are *created co-creators*, a term coined by the theologian Phil Hefner, who has influenced my thinking in religion and science enormously.”<sup>49</sup>

De modo que no sólo puede afirmarlo como creado para crear, sino que el acto de crear es en el ser humano un verdadero acto de oración y de adoración.<sup>50</sup> Ante esta dimensión creadora llega nuestra autora incluso a preguntarse si no, ahora que la evolución técnica ha tomado el relevo a la biológica, habrá sido

<sup>45</sup> A. Foerst, *God in the Machine...*, 38.

<sup>46</sup> A. Foerst, “Artificial Intelligence and Theology: From Mythos to Logos and Back”; en: S. Franchi – G. Güzeldere (dirs.), *Mechanical Bodies, Computational Minds. Artificial Intelligence from Automata to Cyborgs*, Cambridge 2005, 489-513, 497-499; “Cog, a Humanoid Robot...”, 104-105.

<sup>47</sup> “Paul Tillich, the theologian who has influenced my theology the most” (A. Foerst, *God in the Machine...*, 47, en nota 2 a pie de página). A. Foerst, *God in the Machine...*, 4-5.

<sup>48</sup> A. Foerst, “Cog, a Humanoid Robot...”, 106-107.

<sup>49</sup> A. Foerst, *God in the Machine...*, 40.

<sup>50</sup> A. Foerst, *God in the Machine...*, 36.57.

creado el ser humano para crear él otras formas de vida, como sería para ella el caso de futuros robots humanoides.<sup>51</sup> Para ella, nuestro profundo deseo de que nuestras propias creaciones se conviertan en compañeros de nuestro camino, está enraizado en el deseo mismo de Dios de darnos vida a nosotros.<sup>52</sup>

Bajo todo este marco Foerst asume la metáfora del *homo narrans* (*storytelling animal*<sup>53</sup>) como fundamental en su propio pensamiento, porque para ella el ser humano busca ordenar el caos que le circunda creando historias que le den sentido. Pero la realidad del pecado, que ella expresa, siguiendo a Tillich, como alienación (*estrangement*),<sup>54</sup> lleva al ser humano a estar separado de sí mismo, de los otros, de la creación, y de Dios.<sup>55</sup> Convirtiéndose esta alienación en un obstáculo a la llamada a entrar en relación con los otros, a formar comunidad. Una alienación que le impide desenvolverse entre las ambigüedades y paradojas de la vida, en lenguaje de Tillich que sigue Foerst,<sup>56</sup> y le hace necesario crear mitos, narrativas, que le permitan percibirse y percibir el mundo con sentido y coherencia. De modo que para Foerst el ser humano está llamado en primer lugar a crear relaciones, a crear comunidad; y esto a través de su ser encarnado, de una corporalidad que le sitúa en el mundo. Pero esta realidad se ve afectada por su estado de alienación, un estado que, para manejarse con las paradojas y ambigüedades de la vida, le lleva a crear narrativas, como *homo narrans* que es, e insertarse él mismo en ellas. Así, al crear una narrativa compartida se crean también los vínculos que permiten la relación y dan forma a la comunidad. Pero bajo el efecto de la alienación esa misma narrativa tiene el poder

<sup>51</sup> A. Foerst, *God in the Machine...*, 40.

<sup>52</sup> A. Foerst, *God in the Machine...*, 39.

<sup>53</sup> La definición del ser humano como *storytelling animal* se la debemos a Alasdair MacIntyre, quien lo expresó así: "Man is in his actions and practice, as well in his fictions, essentially a story-telling animal" (A. MacIntyre, *After Virtue. A Study in Moral Theory*, Notre Dame-Indiana 1984, 216).

<sup>54</sup> "Man is estranged from the ground of his being, from other beings, and from himself" (P. Tillich, *Systematic Theology II. Existence and The Christ*, Chicago 1957, 44). Foerst se expresa en términos similares: "The term *sin* addresses our estrangement from God, from others, and from ourselves." (A. Foerst, *God in the Machine...*, 22, la cursiva pertenece al original).

<sup>55</sup> A. Foerst, *God in the Machine...*, 22-27.

<sup>56</sup> "*Sin* describes humans as being estranged, ambiguous, torn between incoherent wishes and desires, torn between polarities and fears; *sin* means a life in paradoxes." (A. Foerst, *God in the Machine...*, 26, la cursiva pertenece al original).

de excluir a los que no forman parte de ella. Y esta idea final va a ser clave para comprender una de las afirmaciones fundamentales de Foerst, que los futuros robots capaces de interactuar afectivamente con nosotros, de relacionarse en nuestro mismo entorno humano, deberán ser reconocidos como personas, y lo deberán ser en opinión de Foerst como garantía de que ningún ser humano queda excluido.<sup>57</sup>

### c) *El ser persona en Foerst*

Para Foerst la IA cuestiona abiertamente lo que significa ser humano y ser persona:

AI is on the verge of being able to represent humans as biological systems that can be defined in terms of their mechanisms and functions, and suggests that soon it will be possible to rebuild the human machinery. [...] AI challenges our basic understandings of personhood in a way much more immediate than either genetics or evolutionary theories have done in the past. AI does not deal with unobservable entities such as genes, nor is it merely a theory; it threatens to build human-like creatures based on its assumptions and thus to reinforce the belief in their accuracy.<sup>58</sup>

En su respuesta a este desafío a lo que significa ser humano y ser persona Foerst responde distinguiendo claramente entre ambos.<sup>59</sup> Porque, según ella, es lo que siempre hemos hecho.<sup>60</sup> Para ella “humano” es una referencia biológica, mientras que “persona” lo es cultural, y refleja respeto y aceptación en una comunidad. Para ella ser persona es un reconocimiento que se da en la cercanía de la relación, de la aceptación; es participar en la estructura narrativa de una comunidad, “personhood can be understood as participating in the narrative process of mutual storytelling about who each of us is”.<sup>61</sup> Así los tres ejes sobre

<sup>57</sup> A. Foerst, *God in the Machine...*, 160-162.

<sup>58</sup> A. Foerst, “Artificial Sociability: From Embodied AI Toward New Understanding of Personhood”, en: G. Bugliarello – A. G. Schillinger – C. A. Mitcham – M. Richardson (dirs.), *Technology in Society. An International Journal*, 21 (1999) 373-386, p. 382.

<sup>59</sup> A. Foerst, *God in the Machine...*, 158-160.

<sup>60</sup> A. Foerst, *God in the Machine...*, 160.

<sup>61</sup> A. Foerst, *God in the Machine...*, 186.

los que gira en Foerst el concepto de persona son: el primero la caracterización del ser humano como *homo narrans*, el segundo una consecuencia de esa caracterización, el formar comunidades donde compartir la narración; y en tercer lugar una condición, la necesidad de ser aceptado como parte de la estructura narrativa que configura la comunidad.

A la luz de esto es cómo podemos comprender su respuesta a la pregunta de si los robots son o pueden llegar a ser personas. Su respuesta es sí.<sup>62</sup> Y afirma que según avance su grado de complejidad y de capacidad de interactuar con los seres humanos, se irán integrando en nuestras estructuras narrativas, en nuestras comunidades; y no será ya una cuestión de lo que sean en sí, sino de lo que sean para la comunidad. Más aún, para ella es necesario caminar hacia una comunidad de personas humanas y no humanas que permita redefinir el propio concepto de persona, y garantice que ningún ser humano pueda nunca ser excluido de él. Tal como a lo largo de la historia humana desgraciadamente ha ocurrido. Para ella es claramente preferible un concepto amplio de persona que incluya a no humanos, a un concepto restringido que excluya a otros seres humanos: “I would prefer to adopt a very broad, non-empirical concept of personhood, which might include robots and some apes, rather than the risk of excluding some humans from community.”<sup>63</sup> Para Foerst lo que realmente nos hará mejores serán nuestros robots: “We should work towards the goal of technical counterparts of ourselves that might teach us tolerance and the appreciation of otherness, capabilities that have been curtailed by our evolutionary development.”<sup>64</sup>

#### d) *Los robots del MIT: Cog y Kismet*

El proyecto de Cog se inició en 1993 con el objetivo de imitar el desarrollo cognitivo infantil. No tenía un cuerpo humanoide

<sup>62</sup> En las escenas iniciales de la película *Surrogates* (Jonathan Mostow 2009) aparece una entrevista a Anne Foerst en un medio audiovisual en la que expresa claramente su opinión al respecto: “There’s no question that at some point they’ll be persons. They will not be human, but will they be part of our community? There’s no question in my mind” (desde 1m. 54s. hasta 2m. 01s) .

<sup>63</sup> A. Foerst, “Artificial Sociability...”, 385.

<sup>64</sup> A. Foerst, “Keeping Homo Sapiens 1.0”, *Artifact* 3-4 (2008) 140-144, p. 144. Las cursivas pertenecen al original.

completo, carecía de piernas y no podía desplazarse, ya que solo estaba formado por el torso con la cabeza y los dos brazos.<sup>65</sup> El robot Kismet fue una iniciativa de la entonces estudiante de doctorado en el laboratorio de IA del MIT y colaboradora en el proyecto del robot Cog, Cynthia Breazeal<sup>66</sup>; se trataba de una cabeza robótica que balbuceaba como los niños. Ambos robots tuvieron una gran influencia en Foerst, que se refleja en sus artículos y en su libro. Incluso los coloca en la base de su pensamiento sobre el ser persona: “Because Cog and Kismet inspired social treatment of themselves, we are encouraged to rediscover the personhood of other humans and even animals.”<sup>67</sup> Foerst está convencida de que “Cog *can* tell us something about who we are”.<sup>68</sup> Pero es la experiencia de Kismet, compartida con Breazeal, la que lleva a las dos a escribir: “if a significant number of people were to accept a socially intelligent robot into their group, and treat the robot as an equal member of it, could that robot then become part of the human community?”<sup>69</sup> Foerst siempre hablará de Kismet como de su gran amor entre los robots y se referirá a su relación con él como algo espiritual: “I personally think it is spiritual when I interact with Kismet and have emotional reactions.”<sup>70</sup>

### e) *La ciencia ficción*

#### La ciencia ficción en Anne Foerst

Lo primero para hablar del papel de la ciencia ficción en el pensamiento de Foerst es remitirnos al artículo que ella y Harvey Cox publicaron titulado “Religion and Technology: A New Phase”. La idea clave del artículo es que la teología ha de aprender de la ciencia ficción a plantear preguntas para afrontar los

<sup>65</sup> “Cog is a huge robot, approximately seven feet tall, and somewhat intimidating. It has a massive steel frame” (A. Foerst, *God in the Machine...*, 1).

<sup>66</sup> Para un acercamiento a Cog y Kismet por parte de la propia Cynthia Breazeal, J. D. Brown, *Robo World, the story of robot designer Cynthia Breazeal*, Whashington, D.C. 2005, 45-54.

<sup>67</sup> A. Foerst, “Artificial Sociability...”, 385.

<sup>68</sup> A. Foerst, “Embodied AI, Creation, and Cog”, *Zygon. Journal of Religion and Science* 33 (1998) 455-461, p. 460.

<sup>69</sup> C. Breazeal – A. Foerst, “Schmoozing with Robots, Exploring the Boundary of the Original Wireless Network”; en: *Proceedings of the 1999 Conference on Cognitive Technology*, San Francisco 1999, 375-390, p. 388.

<sup>70</sup> A. Foerst, *God in the Machine...*, 150.

retos del futuro y presentárselos sin miedo a la sociedad.<sup>71</sup> Para Foerst la ciencia ficción es importante porque no solo se nutre de las expectativas de los expertos sino que influye en ellos impulsando la consecución de lo que ella misma presenta. Foerst lo expresa así:

SciFi authors project from current understandings of the world and the speed of technological development into the future and use current findings in science to do so. Their visions about the future, in return, then often inspire scientists and engineers to do certain research projects and to explore certain phenomena in the world.<sup>72</sup>

Dentro del campo de la ciencia ficción Foerst se remite de un modo especial a la película *Blade Runner* (Ridley Scott 1982), la cual usa en las clases con sus alumnos.<sup>73</sup> Y al robot humanoide *Commander Data* de la serie *Star Trek. The Next Generation*, que para ella presenta las grandes cuestiones a la que nos pueden enfrentar en el futuro nuestros robots: la de su propia libertad respecto a nosotros y la de querer ser como nosotros. Nuestra autora remite especialmente al episodio número nueve de la segunda temporada, *The Measure of a Man*, donde se pone en cuestión si este robot humanoide es un objeto con dueño o un ser libre como sus compañeros humanos. Foerst desde ahí plantea lo que ya hemos visto es su reflexión sobre el ser persona:

The reflections about Commander Data as a child of God might help us to remember in humility that each and every person's value is grounded not in her abilities but in God's

<sup>71</sup> H. Cox – A. Foerst, “Religion and Technology...”, 57-59.

<sup>72</sup> A. Foerst, “Commander Data...”, 264. Esta afirmación de Foerst es confirmada por los propios investigadores. Así lo comentan Cynthia Breazeal y Illah Reza Nourbakhsh respecto a *Star Wars* (C. L. Breazeal, *Designing Sociable Robots*, Cambridge 2002, xi-xiii; I. R. Nourbakhsh, *Robot Futures*, Cambridge 2013, xiii); Rodney Brooks respecto a *2001: A Space Odyssey* (Stanley Kubrick 1968) (R. Brooks, *Flesh and Machines: How Robots Will Change Us*, New York 2002, 63); y Murray Shanahan con las novelas de Isaac Asimov (M. Shanahan, *The Technological Singularity*, Cambridge 2015, xi).

<sup>73</sup> “One of our favorite examples. The 1982 classic, *Blade Runner*, deals with the problem of the relationship between thought, memory and embodiment exactly what practitioners of Embodied Artificial Intelligence (Embodied AI) are working on the late 1990's” (H. Cox – A. Foerst, “Religion and Technology...”, 58-59). C. Dreifus, “Anne Foerst...”, 107.

promise and in that alone. The fictional Data might thus serve as a thinking tool to prepare us for the AI machines to come.<sup>74</sup>

### Ciencia ficción y teología

Este uso de la ciencia ficción por parte de Foerst es sin duda una de sus aportaciones a la teología. Porque nos está recordando que ese es un campo que también necesita de la mirada reflexiva del teólogo que se acerque al campo de la IA. Una mirada completa al campo de la IA no puede prescindir de las expresiones de los mitos de nuestro tiempo, que tienen en la ciencia ficción su vehículo privilegiado de expresión. En la literatura pero sobre todo y fundamentalmente en las películas y series.<sup>75</sup> En nuestra época tecno-científica no se ha renunciado al mito, más bien son la ciencia y la técnica las fuentes de una mitología renovada. Jean Brun lo expresa claramente:

Il convient de rappeler, ensuite, l'importance considérable des romans d'anticipation et des *science-fictions*. Voici que la science, que certains nous présentaient comme la grande libératrice des mythes, comme l'expression du *Logos* ayant succédé au *Mythos*, est devenue, avec l'Histoire, le plus puissant facteur mythogène de notre temps. Tous ces récits mythologico-scientifiques nous brossent un tableau d'une présence du futur déjà perceptible.<sup>76</sup>

Dos aportaciones importantes en este campo son la del sociólogo Douglas E. Cowan con su obra<sup>77</sup> *Sacred Space, The Quest for Transcendence in Science Fiction Film and Television*. Este autor defiende que tanto las películas como las series de ciencia ficción desempeñan un papel importante en la búsqueda de trascendencia del ser humano actual, una búsqueda que para él no ha quedado arrinconada en nuestra época, sino que sigue

<sup>74</sup> A. Foerst, "Robot: Child of God", en: R. Stannard (dir.), *God for the 21<sup>st</sup> Century*, Philadelphia 2000, 137-140, p. 139.

<sup>75</sup> En nuestra época ya no se corresponde con la realidad unir series y televisión. Hoy día son un género de más amplias fronteras que la clásica televisión.

<sup>76</sup> J. Brun, *Les masques du désir*, París 1981, 220.

<sup>77</sup> En la que cita a Anne Foerst a propósito del apretón de manos entre Harvey Cox y el robot Cog ( D. E. Cowan, *Sacred Space. The Quest for Transcendence in Science Fiction Film and Television*, Waco-TX 2010, 50).

presente.<sup>78</sup> Y la del teólogo James F. McGrath,<sup>79</sup> que ha dirigido la edición de varios volúmenes,<sup>80</sup> y escrito una obra que condensa su pensamiento, *Theology and Science Fiction*, donde establece paralelismos y encuentros entre ciencia ficción y teología, defendiendo una reflexión teológica capaz de integrar las aproximaciones a la realidad del ser humano que muestran ciencia ficción y religión. Cree también que la teología ha de participar junto a la ciencia en las reflexiones acerca del futuro y del papel de la humanidad en el universo.<sup>81</sup>

#### 4. DEBILIDADES DEL PENSAMIENTO DE FOERST

##### a) *La crítica de Mary Gerhart y Allan Melvin Russell*

Los planteamientos de Foerst son valientes y han abierto un camino importante para la teología en el encuentro con la IA, pero en el método de la aproximación simbólica para el diálogo entre teología e IA no ha profundizado, en nuestra opinión, lo necesario; y a su planteamiento de la concesión del estatuto de persona a nuestros futuros robots, lo creemos problemático en el marco de una antropología cristiana. En este sentido Foerst sufrió una crítica importante por parte de Mary Gerhart y Allan Melvin Russell en un artículo en la revista *Zygon* titulado “Cog Is to Us as We Are to God: A Response to Anne Foerst” y que era una respuesta al que previamente había publicado Foerst: “Cog, a Humanoid Robot, and the Question of the Image of God”. En su artículo Gerhart y Russell realizan una crítica profunda de los planteamientos de Foerst que cuestiona la fundamentación de los mismos. Y que ellos resumen en una frase: “The paradoxical result is that Foerst offers us a bleached theology and a romantized robotology”.<sup>82</sup> Y es una crítica en nuestra opinión acertada

<sup>78</sup> D. Cowan, *Sacred Space...*, x-xi.

<sup>79</sup> Profesor de Nuevo Testamento en la Universidad Butler de Indianapolis.

<sup>80</sup> J. F. McGrath (dir.), *Religion and Science Fiction*, Eugene-OR 2011; A. Crome – J. McGrath (dirs.), *Time and Relative Dimensions in Faith. Religion and Doctor Who*, London 2013.

<sup>81</sup> J. F. McGrath, *Theology and Science Fiction*, Eugene-OR 2016.

<sup>82</sup> M. Gerhart – A. M. Russell, “Cog Is to Us as We Are to God: A Response to Anne Foerst (Response)”, *Zygon. Journal of Religion and Science* 33 (1998) 263-269, p. 267.

pero que deja de lado el aspecto más importante a tener en cuenta en Foerst, su carácter de pionera en un compromiso teológico en las fronteras tecnológicas de nuestro mundo, en su caso en el campo de la IA. Ser conscientes de las limitaciones de su teología, una teología diluida en palabra de Gerhart y Russell, no evita reconocerle el mérito de su valentía para empezar a recorrer un camino al que ya no puede renunciar la teología. Estamos de acuerdo también con Gerhart y Russell en hablar en Foerst de una robotología romantizada, la cual unida a su planteamiento sobre el ser persona de los futuros robots suponen los puntos más críticos para nosotros de su pensamiento y que exigen por tanto una respuesta.

#### b) *Ser humano y ser persona*

Foerst partía de *ser humano* y *ser persona* como dos conceptos distintos, el primero biológico y el segundo cultural, como demostraría, según ella, la negación de la condición de ser persona que durante la historia han recibido y reciben muchos seres humanos. Porque para nuestra autora la condición de persona es una concesión al ser admitidos en una narrativa a la que somos llamados a formar parte: “any entity that is assigned personhood by some human beings, and that is potentially capable of being estranged, is a person.”<sup>83</sup> En Foerst el ser persona queda condicionado a la relacionalidad y a la concesión. Como Foerst sí que defiende que a ningún ser humano se le niegue la condición de persona, plantea como solución a esa división, lo que ella denomina una concepción inclusiva del ser persona que al extenderse más allá de los límites humanos hasta los robots garantice que ningún ser humano sea excluido.

Robots, therefore, can be used as thinking tools to analyze the concept of personhood afresh. They make us aware of our own bonding limitations and enable us to work towards a world in which we assign personhood to all human beings –despite our prejudices and judgments– even if this might mean that we have to assign personhood to the children of our technological future as well.<sup>84</sup>

<sup>83</sup> A. Foerst, “Robots and Theology”, *Erwägen Wissen Ethik* 20 (2009) 181-193, p. 192.

<sup>84</sup> A. Foerst, “Robots and Theology”, 192.

Sin embargo, en nuestra opinión, el argumento de Foerst consigue lo contrario a lo que pretende. Porque al asumir la separación entre ser humano y ser persona, y condicionar esta última a la relacionalidad y la concesión, la misma puerta que abre para dejar entrar a robots es la que puede servir para privar a algunos seres humanos de su condición de personas. Tratándose no de una cuestión nominal, porque al depositar sobre el reconocimiento como persona la base del valor y dignidad del ser humano, privándole de ese reconocimiento le negamos su valor. Y la teología, sea cual sea la evolución del concepto de persona, ha de defender siempre al ser humano como sujeto de un valor y dignidad inalienables, evitando que por la vía de redefinir el concepto de persona se vacíe de contenido lo humano. Foerst, ante este tipo de crítica, enfatiza que ella cree que todos los seres humanos se merecen el título de persona y lo intenta mostrar con lo que ella denomina un lenguaje inclusivo, el que le permite incluir a los robots.<sup>85</sup> Pero el ser persona no es un título que se *merezca*. Para la teología cristiana todos los seres humanos *son* personas porque, siguiendo a J. L. Ruiz de la Peña, podemos afirmar que esa es la categoría que utilizamos para expresar al ser humano como un ser relacional, libre y responsable, irreductible a lo infrahumano y abierto constitutivamente a Dios.<sup>86</sup> Porque para la teología la categoría de persona aplicada al ser humano no es un título, es una forma de expresar quién es realmente el ser humano.<sup>87</sup> No es el ser persona lo que confiere dignidad y valor al ser humano. El valor y dignidad del ser humano, toda la profundidad y trascendencia de la existencia humana enraizada en Dios, los expresamos con la categoría de persona. Es lo humano lo que da contenido a la categoría, no ésta a lo humano.

### c) *Una robotología romantizada*

Ciertamente creemos acertada la crítica de Gerhart y Russell afirmando que Foerst realiza una robotología romantizada

<sup>85</sup> A. Foerst, "Prólogo", en: F. J. Génova, *Anne Foerst. Dimensión religiosa de la búsqueda de inteligencia artificial*, Barcelona 2016, 11-12.

<sup>86</sup> J. L. Ruiz de la Peña, *Imagen de Dios. Antropología teológica fundamental*, Santander 1988, 153-154.

<sup>87</sup> "En el debate antropológico hoy en curso, la cuestión crucial es la que se pregunta si –y en qué medida– el hombre está frente a su mundo como sujeto frente a objetos, si es ser personal o, por el contrario, es una entidad natural *más*" (J. L. Ruiz de la Peña, *Imagen de Dios...*, 166. La cursiva pertenece al original)

(*romanticized robotology*). Porque Foerst en todas sus exposiciones no profundiza en la realidad de la IA y los robots que la encarnan así como en sus desafíos. Ella torna el mito del buen salvaje en el mito del buen robot o la buena IA, dejando completamente de lado la realidad del mal a la que no escapan las creaciones humanas, dando por supuesto que nuestras creaciones serán mejores que nosotros y que en ellas hemos de poner las esperanzas de un futuro más humano. Ella es rotunda en esto poniendo al robot Kismet como ejemplo: “Kismet is our destiny, because it can help us to turn this world into a better and more accepting place.”<sup>88</sup>

Pero ¿cuáles son realmente los perfiles de esa comunidad de personas humanas y no humanas por la que aboga Foerst<sup>89</sup>? ¿Cuáles los retos reales que en nuestra opinión Foerst no llega a alcanzar? Los perfiles se están ya delineando y no encajan con la imagen ideal que Foerst presenta. Los robots dotados de IA que Foerst espera nos hagan mejores seres humanos se despliegan hoy ya en tres grandes campos: industrial-profesional, social y militar. En el área industrial-profesional la idea clásica del robot aislado en una cadena de montaje programado por especialistas da paso a un nuevo tipo de robot, del que es un buen ejemplo Baxter de *Rethink Robotics*.<sup>90</sup> Un robot que puede trabajar junto a los operarios humanos sin peligro para ellos y que no requiere especialistas que lo programen.

En este campo queda igualmente superada por la realidad la idea de que con la automatización las únicas profesiones en peligro son las menos cualificadas. No estamos hablando ya solo de robots, sino de sistemas de IA que son capaces de sustituir y desplazar a los actuales profesionales humanos, desde médicos, pasando por ingenieros y llegando hasta abogados y periodistas, muchas profesiones se van a ver profundamente afectadas o incluso desaparecer en opinión de Richard y Daniel Susskind.<sup>91</sup> Entre otros autores que han profundizado en este tema destacamos a Nicholas Carr, que en su obra *The Glass Cage*,<sup>92</sup> nos alerta no solo de una creciente automatización que va a suponer la

<sup>88</sup> A. Foerst, *God in the Machine...*, 189.

<sup>89</sup> En el último capítulo de su libro titulado *The Community of Human and Nonhuman Persons*. A. Foerst, *God in the Machine...*, 153-190.

<sup>90</sup> <<http://www.rethinkrobotics.com/baxter/>> [Consulta: 14 ago. 2017].

<sup>91</sup> D. Susskind – R. Susskind, *The Future of the Professions. How Technology Will Transform the Work of Human Experts*, Oxford 2015. Estos autores incluyen también a los clérigos.

<sup>92</sup> N. Carr, *The Glass Cage. Automation and Us*, New York 2014.

desaparición de trabajos y profesiones, sino del peligro de ir relegando todas nuestras decisiones y nuestra confianza en esos sistemas, apartando a los seres humanos de la cadena de decisiones sobre su propia vida.<sup>93</sup> En la misma línea Martin Ford nos advierte en *The Rise of the Robots* de las profundas transformaciones del sistema económico y de sus consecuencias sociales. Que para Mark O'Connell supondrían el logro final de la lógica del tecno-capitalismo: el capital no solo dueño ya de los medios de producción sino de la propia fuerza de trabajo, esto es, los sistemas de IA y robots.<sup>94</sup> Pero no todas las visiones son negativas, es el caso de Ruchir Sharma, que defiende la automatización como necesidad y oportunidad para el futuro de la humanidad, mostrándose especialmente crítico con Ford.<sup>95</sup>

El área de los robots sociables y personales es la de aquellos destinados a interactuar directamente con el ser humano. Ya sea cuidando de enfermos, ancianos<sup>96</sup> o niños; realizando tareas en el hogar, las empresas, oficinas, comercios; o como compañeros personales que mitigan la soledad, sirviendo de objetos sexuales,<sup>97</sup> o

<sup>93</sup> El 26 de septiembre de 1983 el teniente coronel del ejército soviético Stanislav Petrov evitó que la información de los ordenadores soviéticos y sus operadores, afirmando que los Estados Unidos había lanzado un ataque nuclear contra la Unión Soviética, supusiera un contraataque nuclear que hubiese desencadenado la Tercera Guerra Mundial. El tenía la responsabilidad de dar por buena la información, pero su conciencia de lo que había en juego le hizo no tomar ninguna determinación sin confirmación del ataque por otros medios. El error informático fue corregido por la decisión de un ser humano frente a las presiones para que confirmara el ataque. El hecho ha quedado recogido en el documental *The Man Who Saved The World* (Peter Anthony 2014).

<sup>94</sup> M. O'Connell, *To Be a Machine. Adventures among Cyborgs, Utopians, Hackers, and the Futurists Solving the Modest Problem of Death*, London 2017, 120.

<sup>95</sup> R. Sharma, *The Rise and Fall of Nations. Ten Rules of Change in the Post-Crisis World*, New York 2016, 24-27, 54-56.

<sup>96</sup> Interesante resulta el proyecto Romeo. En cuya página Web se describe así: "Romeo is a 140 cm tall humanoid robot, designed to explore and further research into assisting elderly people and those who are losing their autonomy. Romeo is the fruit of collaboration between numerous French and European laboratories and institutions. His size was determined so as to enable him to open doors, climb stairs and reach objects on a table (< <https://www.ald.softbankrobotics.com/en/robots/romeo>> [Consulta: 15 ago. 2017]).

<sup>97</sup> D. Levy, *Love and Sex with Robots: The Evolution of Human-Robot Relationships*, New York 2007. «Many experts in robotics believe that sex will drive many of commercial advances in robotics, because, well, sex sells» (P. W. Singer, *Wired for War. The Robotics Revolution and Conflict in the 21st*

colaborando en las mismas tareas de los seres humanos. En esta área es clave el desarrollo de lo que se ha venido en denominar computación afectiva<sup>98</sup> y nos permite hablar de inteligencia artificial emocional (*Artificial Emotional Intelligence*).<sup>99</sup>

La inteligencia artificial emocional permite al robot, o a la IA en general, reconocer las emociones humanas y dar una respuesta acorde con ella. Ya sea un bot conversacional que atiende una llamada telefónica. O un robot personal en el hogar interactuando con un ser humano. En el área militar se trata del paso de drones controlados a distancia a drones y robots autónomos que actúan como un ejército de autómatas.<sup>100</sup> No es casual que sea el ejército de Estados Unidos, a través del programa DARPA (*Defense Advanced Research Projects Agency*), el principal suministrador de fondos en las investigaciones que en el campo de la IA y la robótica han tenido lugar.<sup>101</sup> Como ejemplo podemos citar el *DARPA Robotics Challenge* (DRC),<sup>102</sup> cuyo objetivo declarado era impulsar la investigación de robots capaces de actuar en entornos peligrosos y capaces de utilizar los mismos dispositivos que los seres humanos. Fruto de este desafío es el robot humanoide Atlas de Boston Dynamics.<sup>103</sup> Este último ejemplo, con las posibilidades que se abren a un robot impulsado desde el campo militar para actuar en entornos peligrosos y manejar dispositivos utilizados por seres humanos, nos muestra lo idealizada que es realmente la presentación de Foerst. Pero no menos ilustrativo es el de los robots sociales dedicados a cuidar ancianos y enfer-

*Century*, New York 2010, 419). Para un acercamiento a lo que supone la IA emocional aplicada a robots sexuales R. Yonck, *Heart of the Machine. Our Future in a World of Artificial Emotional Intelligence*, New York 2017, 179-193.

<sup>98</sup> La pionera y una gran experta en este campo es Rosalind Picard, investigadora del MIT. R. Picard, *Affective Computing*, Cambridge 1997.

<sup>99</sup> R. Yonck, *Heart of the Machine...*

<sup>100</sup> P. W. Singer está convencido de que el deseo de un soldado ideal, efectivo y que cumpla sin dudar las órdenes, se impondrá sobre cualquier otra consideración (P. W. Singer, *Wired for War...*, 126-134).

<sup>101</sup> Braden R. Allenby y Daniel Sarewitz, en *The Techno-Human Condition*, invitan a preguntarse por las intenciones verdaderas de muchas de las investigaciones que se producen en los campos tecnológicos. Especialmente advierten de que son los programas militares los más interesados y con mayores proyectos en ejecución (B. R. Allenby – D. Sarewitz, *The Techno-Human Condition*, Cambridge 2011, 24-29. 152. 159-189).

<sup>102</sup> < <https://www.darpa.mil/program/darpa-robotics-challenge> > [Consulta: 15 ago. 2017].

<sup>103</sup> < <https://www.bostondynamics.com/atlas> > [Consulta: 15 ago. 2017].

mos, que como sofisticados sustitutos de una verdadera relación humana ofrecen una afectividad simulada.

Precisamente siguiendo a Foerst en la aplicación de la ciencia ficción a la teología, podemos volver la mirada sobre la película *Robot & Frank* (Jake Schreier 2012) que enmarca lo que verdaderamente supone el proyecto Romeo, y que cuestiona abiertamente la visión de Foerst de unos robots que hagan el mundo más humano. En esa película se muestra la relación afectiva que se establece entre un anciano y un robot que le regala uno de sus hijos para paliar la poca atención que le prestan. La escena final en una residencia donde cada anciano es acompañado en su soledad por uno de esos pequeños robots humanoides, refleja muy bien el peligro que ha pasado desapercibido para Foerst. Otro riesgo unido al anterior nos lo ilustra un robot que ya ha empezado a comercializarse, Buddy de Blue Frog Robotics.<sup>104</sup> Es un pequeño asistente familiar para el hogar, que se desplaza sobre ruedas, y está pensado para convertirse en uno más de la familia. Sustituyendo a los padres, comunicando a distancia con los abuelos para que estén presentes virtualmente desde su soledad, en definitiva cambiando el marco mismo de las relaciones personales y familiares.<sup>105</sup> Este panorama que someramente presentamos es el que nos permite afirmar que Foerst ha presentado una visión idealizada de la robótica, que es desafiada no por el futuro, sino ya por el presente.

#### d) *La aportación de Foerst*

La verdadera aportación de Anne Foerst es haber abierto un camino nuevo e irrenunciable para la teología. Es haber puesto el foco de la atención teológica sobre una IA que ha llegado para quedarse y que delinea los contornos de un futuro para la humanidad esculpido con unas tecnologías emergentes que forman ya parte de nuestra vida. Y que desafían no solo la idea de ser humano sino su continuidad como tal. A Foerst por tanto le hemos de reconocer el habernos presentado un verdadero reto para la teología. Pero afrontar este reto supone en primer lugar hacernos una pregunta que consideramos fundamental: ¿para qué

<sup>104</sup> < <http://www.bluefrogrobotics.com/en/buddy-your-companion-robot/> > [Consulta: 15 ago. 2017].

<sup>105</sup> El video de la presentación, aunque esté emocionalmente cuidado, despierta suficientes preocupaciones éticas por sí mismo, <<https://youtu.be/51yGC3iytBY>> [Consulta: 16 ago. 2017].

teología? Creemos que no para una teología que no haya superado lo que Iliá Delio denomina el complejo helénico,<sup>106</sup> no para una teología anclada en cosmovisiones antiguas y medievales, estáticas y alejadas de la comprensión dinámico-evolutiva,<sup>107</sup> y del conocimiento que el ser humano tiene hoy sobre sí mismo y sobre el mundo. Porque el primer desafío que ha de afrontar la teología es ella misma. Es necesaria una teología valiente y humilde que se encarne en las fronteras tecnológicas de nuestro mundo y explore nuevos lenguajes para ser fiel a su misión. Una teología que Iliá Delio plantea como integral y que para ella ha de incluir tanto espiritualidad y humanidades, como tecnología y ciencia.<sup>108</sup> En definitiva, una teología nueva que asuma la experiencia del mundo de hoy y desde ahí responda a los retos que éste le presenta.<sup>109</sup>

<sup>106</sup> Delio se expresa en estos términos: “theologians would do well to overcome the inherited hellenic complex or at least translate the best of it into a contemporary theological framework. By *hellenic complex* I mean the Greek architecture of metaphysics” (I. Delio, *The Unbearable Wholeness of Being. God, Evolution, and the Power of Love*, New York 2013, xx). Más aún: “Catholic theology must relinquish the Greek philosophical and cosmological framework in which is embedded. These forms have become inadequate for the continued life and development of the Christian faith” (I. Delio, *Making All Things New. Catholicity, Cosmology, Consciousness*, New York 2015, 140).

<sup>107</sup> En este punto son muy interesantes las aportaciones de John Haught y de Iliá Delio, que hablan de la teoría de la evolución como del *regalo de Darwin a la teología*: J. F. Haught, *Making Sense of Evolution. Darwin, God, and the Drama of Life*, Louisville-KY 2010, 137; *God After Darwin. A Theology of Evolution*, Oxford 2000, 45-56. I. Delio, *Christ in Evolution*, New York 2008, 22.

<sup>108</sup> I. Delio, *The Unbearable Wholeness...*, xvii-xviii; *The Unbearable Wholeness...*, xix. Teilhard lo planteaba así: “S’ils veulent parler dans une langue intelligible, et mieux encore, persuasive, à nos contemporains, il est indispensable, avant toutes choses, que les théoriciens du Christianisme comprennent, acceptent et aiment l’idée nouvelle que l’Homme moderne a été *scientifiquement* amené à se faire de lui-même” (P. Teilhard de Chardin, *Comment je crois*, 164. La cursiva pertenece al original). No algo muy alejado de lo que propone Teilhard es lo que nos presenta el discurso que el libro de los Hechos de los Apóstoles atribuye a Pablo en el Areópago de Atenas (Hch 17,22-31). Un discurso que no parte de la experiencia del pueblo de Israel, sino del propio pueblo griego al que va dirigido.

<sup>109</sup> Pierre Teilhard de Chardin lo expresaba así: “Prenons loyalement le Monde, tel qu’il se présente à nous aujourd’hui à la lumière de notre raison: non pas le Monde de quatre mille ans, encerclé dans ses huit ou neuf sphères, *pour lequel a été écrite la Théologie de nos livres*.” P. Teilhard de Chardin, *Comment je crois*, ..., 97. La cursiva pertenece al original). Porque como decía Ruiz de la Peña “no se puede hacer teología sobre el mundo de espaldas a las cuestiones fronterizas” (J. L. Ruiz de la Peña, *Teología de la Creación*, Santander 1996, 12. La cursiva pertenece al original). Y la IA los es.

## 5. CONCLUSIÓN

En estas líneas hemos pretendido un acercamiento al pensamiento de una teóloga, Anne Foerst, que ha encarnado el compromiso vital en la frontera de la IA. Y lo hemos hecho enmarcándola someramente en la realidad de la IA hoy. Su pensamiento muestra en nuestra opinión debilidades, que podemos resumir en un concepto de persona, que sostenido exclusivamente en el polo de la relacionalidad, le permite plantear la inclusión de robots para garantizar la no exclusión de ningún ser humano. Pero el resultado es abrir la puerta a lo que ella no quiere, la exclusión de seres humanos que no encajen en esa clave de relacionalidad. También parte de una visión idealizada de lo que pueden llegar a ser los robots, sin abordar el profundo desafío teológico que presentan. Pero esta crítica no es óbice para reconocer la importancia del paso que ella ha dado para la teología. Ha abierto una puerta por la que nos impulsa a seguir, aunque nuestros caminos difieran. Porque la teología no puede ser ajena a la realidad del mundo que le rodea, salvo que decida encerrarse en sí misma y renunciar a ser sal que da sabor y levadura que fermenta la masa. Y la realidad del mundo es hoy inseparable de la IA y del resto de tecnologías que se despliegan junto a ella. Un despliegue que no es sino el despliegue de la dimensión creadora de quien es imagen del Dios creador. Un despliegue que pone en juego la posibilidad de consumir aquello a lo que es llamado desde el fondo de su ser, o de frustrarlo empujando su destino al fracaso. De esto ha de ser consciente una nueva teología sin miedo a caminar en la frontera, como ha hecho Anne Foerst.