**LA CONTRATACIÓN ENTRE SISTEMAS DE INTELIGENCIA ARTIFICIAL**

Francisco Pertíñez Vílchez

Catedrático de Derecho Civil. Universidad de Granada

SUMARIO.

**I. INTRODUCCIÓN. II. INTELIGENCIA ARTIFICIAL Y VOLUNTAD CONTRACTUAL. III. LOS CONTRATOS ALGORÍTMICOS**. III.1 Diferencia entre sistemas determinísticos y autónomos. III.2. La perfección de los contratos algorítmicos. III.2.A. La eficacia de los contratos concluidos entre sistemas de IA autónomos: la necesidad de nuevas ficciones. III.2.B. El recurso a la personalidad jurídica de los sistemas de IA. Una opción descartable. III.2.C. La vinculación contractual derivada de la responsabilidad. **IV. LA INEFICIA DE LOS CONTRATOS ALGORÍTMICOS**. IV.2. La insuficiencia del error vicio. IV. 3. El error obstativo evidente y no imputable. IV.4. El dolo algorítmico

1. **INTRODUCCIÓN**

El desarrollo exponencial de la inteligencia artificial y en particular de los sistemas de aprendizaje automático (machine learning) o aprendizaje profundo (deep learning) está transformando todos los ámbitos del conocimiento y de la actividad humana y lógicamente el contrato no permanece ajeno a dicha revolución tecnológica. La inteligencia artificial está influyendo en la manera en la que se contrata y no es exagerado decir que también en el propio concepto de contrato como acuerdo de voluntades.

Las herramientas fundadas en la tecnología de los grandes modelos lingüísticos (LLM, por sus siglas inglesas Large Language Models), como ChatGPT, permiten generar de manera automatizada en apenas segundos cualquier modelo de contrato, adaptado a cualquier ordenamiento jurídico. Cada vez es más habitual el uso por las empresas de sistemas predictivos de inteligencia artificial que establecen correlaciones entre una cantidad ingente de datos, ya sea para evaluar la solvencia de sus contrapartes, para dirigir eficazmente la publicidad hacia potenciales clientes o para predecir el precio que cada cliente potencial estará dispuesto a pagar por sus bienes o servicios. Las plataformas de mercados en línea, como Amazon o Booking, utilizan algoritmos para priorizar las ofertas de terceros empresarios que ofrecen sus bienes o servicios a través de las mismas, de los que depende el éxito o incluso la superviviencia de estos. Los robo advisors y los chatbots permiten suplir la intervención humana en el suministro de información o asesoramiento precontractual.

La irrupción de la inteligencia artificial en el contrato a través de todas estas utilidades sin ninguna duda requiere una adaptación del régimen normativo existente sobre protección de consumidores, prácticas comerciales desleales o protección de datos y una reflexión sobre las prácticas contractuales discriminatorias adoptadas exclusivamente sobre la base de decisiones automatizadas. Pese a ello, ninguno de los usos anteriores de la inteligencia artificial altera los presupuestos del contrato, como producto del acuerdo entre las partes.

Sin embargo, las bases del contrato se pueden ver profundamente afectadas cuando las declaraciones de oferta y aceptación se efectúan a través de dos sistemas de inteligencia artificial sin intervención humana. No es ciencia ficción, la contratación entre sistemas de inteligencia artificial ya está ocurriendo en los mercados financieros, particularmente en los mercados de negociación algorítmica de alta frecuencia -ya definidos en el art. 3.40 de la Directiva 2014/65, de 15 de mayo de 2014, relativa a los mercados de instrumentos financieros, MIFID II- o en las plataformas de intercambio de criptomonedas y, desde un punto de vista más doméstico, un previsible desarrollo del llamado “internet de las cosas”, permitirá que sistemas de inteligencia artificial sencillos contraten desde los hogares con proveedores bienes y servicios de consumo de manera automatizada.

Si la tecnología permite que se realicen acuerdos en los que no intervengan dos voluntades humanas, los juristas tendremos que plantearnos si, al margen de la evidencia tecnológica y social, negamos la consideración de contrato generador de obligaciones a estos acuerdos entre máquinas o si, por el contrario, construimos un nuevo concepto de contrato que ya no sea necesariamente el acuerdo entre dos voluntades. Necesitamos por lo tanto construir nuevas ficciones, reflexionar sobre el papel que desempeña la voluntad en la existencia misma del contrato y reinterpretar otras instituciones del derecho de contratos íntimamente vinculadas a la voluntad, como los vicios del consentimiento, la interpretación o la ilicitud de la causa.

**II. INTELIGENCIA ARTIFICIAL Y VOLUNTAD CONTRACTUAL**

La inteligencia artificial puede intervenir en la formación de un contrato de muy diversos modos. Cuanto mayor sea la intervención del sistema de inteligencia artificial empleado en la formación del contrato, menor será la incidencia de la voluntad humana y más en entredicho estará la consideración misma del contrato como producto del acuerdo de dos voluntades.

Podemos distinguir distintas hipótesis dependiendo de la menor o mayor influencia de los sistemas de inteligencia artificial en la formación de los contratos:

- Una herramienta de inteligencia artificial fundada en la tecnología de los grandes modelos lingüísticos, como ChatGPT, puede simplemente crear un modelo de contrato redactado en lenguaje natural, sin interferir en la perfección del contrato que posteriormente realizarán las partes. Evidentemente, el uso de esta tecnología no tensiona el concepto de contrato como acuerdo de dos voluntades, pues su empleo se puede equiparar al uso de los tradicionales formularios. No obstante, el mal uso de estas herramientas será causa de no pocas distorsiones en relación a cláusulas contractuales generadas por la inteligencia artificial, sin tener en cuenta por lo tanto las circunstancias, contextos y finalidades perseguidas por las partes en el contrato concreto. Los ordenamientos romano-germánicos que hacen prevalecer la verdadera intención de las partes sobre el sentido literal de las palabras estarán mejor dotados para resolver razonablemente estas distorsiones que los ordenamientos de *common law* apegados a la literalidad de los términos y a la “*parol evidence rule*”, conforme a la cual no cabe admitir en la interpretación del contrato pruebas ajenas al mismo.

- Los chatbot o los robo advisors proporcionan a los potenciales contratantes información precontractual automatizada sobre bienes y servicios que puede tener influencia en la decisión contractual del usuario. Estos sistemas deben considerarse como meros instrumentos de la parte oferente que será, por lo tanto, responsable de los defectos de la información así transmitida, siendo cuestionable que la advertencia a los usuarios de estar interactuando con sistemas de inteligencia artificial, en el sentido del proyectado art. 52 de la Propuesta de Reglamento del parlamento europeo y del Consejo de 21 de abril de 2021 por el que se establecen normas armonizadas en materia de Inteligencia Artificial, sirva para desplazar sobre los usuarios una carga de desconfiar de la fiabilidad de la información transmitida por un agente no humano.

- Un empresario puede poner en marcha un sistema de inteligencia artificial que genere ofertas de manera automatizada, de modo que se pueda producir la aceptación por un usuario humano antes de que el empresario que va a quedar vinculado conozca los términos de la oferta. Nuestro ordenamiento asume desde la introducción del párrafo tercero del art. 1262 CC por mor de la Disp. Adic. 4.1 de la Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y del comercio electrónico que estos contratos quedan perfeccionados desde que se manifiesta la aceptación.

- Los empresarios pueden utilizar, además, sistemas predictivos de inteligencia artificial para generar de manera automatizada ofertas personalizadas sobre el precio para cada usuario. Esta cuestión debe analizarse en el marco del derecho de todo interesado a no ser objeto de una decisión basada únicamente en el tratamiento automatizado de datos del art. 22 del Reglamento 2016/679 de Protección de Datos y de la excepción a esta prohibición cuando una decisión de tal tipo esté autorizada por el Derecho de la Unión (art. 22.2.b del Reglamento 2016/679). Desde esta perspectiva, el “pricing”, sería una práctica lícita, en cuanto que autorizada por el 6.1.e. bis) de la Directiva 2011/83/UE sobre los derechos de los consumidores (Introducido por la Directiva 2019/2161, por la que se modifican las Directivas 93/13/CEE 98/6/CE, 2005/29/CE y 2011/83/UE), que simplemente obliga a informar al consumidor de que el precio ha sido personalizado basándose en la toma de decisiones automatizada. El único límite a esta práctica es que la decisión automatizada sobre precios no esté fundada en categorías de datos que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas o la afiliación sindical (art. 22.4 y 9.1 del RPD), lo que, no obstante, no evita discriminaciones encubiertas fundadas en otras categorías de datos en apariencia no discriminatorias (código postal, preferencias de compra, etc), pero respecto de las que un sistema de inteligencia artificial pudiera trazar una conectividad con aquellas otras categorías discriminatorias.

- Dos sistemas de inteligencia artificial pueden contratar entre sí de manera automatizada, sin intervención humana, a partir de unas instrucciones previamente definidas por las partes, de manera que la oferta y aceptación generada por los sistemas de ambas partes sería simplemente un código de ejecución de tales instrucciones.

- Por último, dos sistemas de inteligencia artificial pueden contratar no sólo de manera automatizada, sino plenamente autónoma, es decir, sin obedecer a unas instrucciones previamente definidas por las partes, diseñando sus propias estrategias, en aras a la consecución de un resultado: maximizar el beneficio de la persona que puso en servicio el sistema.

La manera en la que la inteligencia artificial puede desplazar a la voluntad humana en la formación de un contrato depende no sólo de para qué se utilice (redacción, información precontractual, generación de ofertas o perfección), sino también de qué tipo de inteligencia artificial se utilice para cada una de estas tareas. Así, conviene distinguir entre dos modelos de inteligencia artificial: los que emplean algoritmos determinísticos que hacen exactamente aquello para lo que han sido programados y no tienen capacidad de desarrollar sus propias estrategias y respuestas y los modelos de inteligencia artificial fundados en el aprendizaje autónomo que son capaces de construir respuestas autónomas que no han sido predeterminadas y que resultan impredecibles para sus programadores.

**III. LOS CONTRATOS ALGORÍTMICOS**

**III.1 Diferencia entre sistemas determinísticos y autónomos**

El término contratos algorítmicos se emplea para definir aquellos que se perfeccionan por sistemas de inteligencia artificial, sin intervención humana, ni por el lado de la oferta, ni de la aceptación. Son los algoritmos los que determinan qué contratos celebrar, con quién y a qué precio. Las personas que ponen en marcha sistemas de inteligencia artificial para realizar contratos algorítmicos delegan el uso de su libertad contractual en tales sistemas porque los algoritmos pueden tomar en consideración una serie de datos, números y circunstancias que un humano jamás podría llegar a tener en cuenta y hacerlo además en milisegundos.

La cuestión entonces es si dos programas de software pueden concluir un contrato plenamente vinculante, en ausencia de un acto de intervención humana no solamente desde el lado de la oferta -algo ya plenamente asumido en el art. 1.262. III CC- sino también de la aceptación.

Para enfocar correctamente esta cuestión se hace preciso distinguir dentro de la categoría de contratos algorítmicos los que se perfeccionan mediante el uso de sistemas de inteligencia artificial predeterminada y los que se perfeccionan mediante el uso de sistemas de inteligencia artificial autónomos.

Los sistemas de inteligencia artificial predeterminada han sido programados para ofrecer siempre el mismo resultado (output) ante unos mismos datos (inputs). Estos programas no tienen capacidad de desarrollar sus propias respuestas, sino sólo de ejecutar respuestas que han sido pre-programadas. Sus algoritmos obedecen, por lo tanto, al esquema “if” (si acontece un determinado supuesto o la suma de un número determinado de supuestos) “then” (se produce un determinado resultado). Esto no quiere decir que su funcionamiento sea simple, antes bien, la variedad de resultados que pueden ofrecer estos algoritmos dependerá de la cantidad de inputs o datos que puedan manejar y de la variedad de hipótesis y de correlaciones entre estos datos que hayan previsto sus programadores. Cuando se utilizan estos sistemas para emitir declaraciones negociales concretas no se puede decir que tales decisiones sean totalmente ajenas a la voluntad de la parte que puso en marcha tal sistema de inteligencia artificial, porque el resultado que emite el sistema (comprar o vender a un precio dado tal tipo y tal cantidad de activos, con unas condiciones de pago y ofreciendo o exigiendo tal o cual garantía) ha sido previsto por el programador para el caso de que el sistema reciba información sobre unas determinadas circunstancias. En tales casos, puede decirse que la declaración negocial emitida por el sistema de inteligencia artificial obedece a una voluntad pre programada. Esto no quiere decir que el sistema no pueda producir resultados indeseados que se deban a errores de diseño, de programación o de funcionamiento del sistema o a una alteración de las circunstancias respecto de las que se tuvieron en el momento del diseño de las distintas hipótesis.

Por el contrario, los sistemas de inteligencia artificial autónomos, son capaces de aprender, esto es, de construir estrategias o trazar por sí mismos correlaciones entre datos (inputs) no previamente programadas y ofrecer como resultado soluciones (outputs) en forma de declaraciones negociales no previstas por el programador, en aras a la consecución de un fin previamente definido, que tratándose de un contrato, es conseguir la transacción que maximice el beneficio de quien puso en marcha el sistema. Además, el aprendizaje mecánico permite que estos sistemas se adapten a las nuevas circunstancias rápidamente. Por eso, en determinados mercados, como los mercados financieros de negociación algorítmica de alta frecuencia, en los que el éxito de las transacciones depende de una capacidad de reacción de milisegundos se confía cada vez más en sistemas de inteligencia artificial autónomos fundados en correlaciones instantáneas de datos, extraídos no sólo de los mercados, sino de agencias de noticias financieras o generales o incluso de redes sociales, que en decisiones humanas fundadas en análisis lógicos de previsión y causalidad.

**III.2. La perfección de los contratos algorítmicos**

**III.2.A La eficacia de los contratos concluidos entre sistemas de IA autónomos: la necesidad de nuevas ficciones**

A partir de las consideraciones anteriormente expuestas se colige que el dogma del contrato como acto humano, en cuanto que acuerdo de voluntades sobre el objeto y la causa, está más claramente en entredicho en los contratos algorítmicos surgidos de sistemas de inteligencia artificial autónomos, que en los contratos algorítmicos en los que las declaraciones negociales sean el producto de un sistema de inteligencia artificial predeterminado.

Cuando las declaraciones negociales son la ejecución de un sistema predeterminado, por muy complejo que este sea, por más que contemple miles o millones de hipótesis, la emisión de la declaración de oferta o de aceptación no será más que la ejecución de una voluntad previamente programada. El contrato existirá aunque la parte que puso en marcha en sistema desconozca la existencia misma del contrato, la identidad de la contraparte o incluso los términos del acuerdo alcanzado, porque tal acuerdo obedece a una de las hipótesis (miles o millones) que se concretaron en las instrucciones dadas al sistema para ejecutar la orden de contratar. El supuesto no difiere sustancialmente del de la contratación mediante dispositivos automáticos contemplado en el art. 1.262.III del Código Civil, por más que la “mes legislatoris” de la Ley 34/2002 de Servicios de la Sociedad de la Información, por la que se introdujo tal norma, tuviera presente un acto de aceptación humana y no una aceptación efectuada también de manera automatizada.

Sin embargo, cuando en la perfección de un contrato intervienen sistemas de inteligencia artificial autónomos, difícilmente se puede considerar que una declaración negocial sobre un contrato concreto emitida por uno de estos sistemas sea expresión de la voluntad de la persona que lo puso en marcha. Evidentemente, esa persona puso en marcha el sistema con la intención de que concluyera contratos vinculantes para ella, pero sobre los contratos concretos que el sistema realiza, tal persona no ha expresado una voluntad negocial, ni tan siquiera de manera pre programada.

Aun cuando entonces sea más que cuestionable que el contrato sea el producto del consentimiento, entendido como acto puramente libre y consciente y, en cuanto tal, humano, la seguridad del tráfico reclama atribuir carácter vinculante a tales transacciones. Atribuir carácter vinculante a las declaraciones negociales efectuadas por sistemas de inteligencia artificial autónomos, requiere construir necesariamente nuevas ficciones.

**III.2.B. El recurso a la personalidad jurídica de los sistemas de IA. Una opción descartable**

La primera ficción al respecto es la de atribuir personalidad jurídica a los sistemas de inteligencia artificial, sobre la base de su autonomía de decisión, de modo que, dotados de personalidad, tales sistemas podrían emitir una declaración de voluntad que resultase vinculante para quien puso en marcha el sistema por la vía de la representación. En definitiva, la propuesta consiste en considerar a los sistemas de inteligencia artificial autónomos como una suerte de factor notorio, con capacidad para vincular a su principal frente a terceros en el giro propio de su actividad, en virtud de la notoriedad de su pertenencia al principal (art. 286 C.Com).

El recurso a una forma de personalidad jurídica algorítmica para justificar el carácter vinculante para el principal de los acuerdos alcanzados por sistemas de inteligencia artificial autónomos vía representación parece artificioso y, además, no resuelve la cuestión inicial de si es posible atribuir carácter vinculante a declaraciones negociales emitidas sin intervención humana, puesto que conferir a un sistema de inteligencia artificial personalidad jurídica no pasaría de ser el resultado de un convencionalismo, que no dota al sistema de facultades volitivas y en última instancia quien emite la declaración negocial, aunque fuera como un representante, seguiría siendo el sistema de inteligencia artificial. Luego se habría dado un rodeo para volver a estar en el punto de partida inicial.

Algunos de quienes han defendido la personalidad jurídica algorítmica en el ámbito del derecho de contratos lo ha hecho con el propósito de atribuir al principal, antes de quedar vinculado por el contrato, una última palabra o necesidad de ratificación, ya fuera expresa o tácita, que le permitiera desdecirse de contratos realizados por sus sistemas de inteligencia artificial, que tras un acto de verificación humana, no le resultasen satisfactorios. Sin embargo, no parece que en el ordenamiento español el recurso a la ficción de la personalidad jurídica del sistema de inteligencia artificial y al instituto de la representación permita atribuir al principal una última palabra o confirmación para dotar de eficacia vinculante al acuerdo alcanzado por el sistema de inteligencia artificial. Aparte de que desde un punto de vista práctico este acto de revisión humana es inconciliable con la celeridad de los mercados en los que operan los sistemas de inteligencia artificial autónomos, el derecho español sólo contempla la ratificación por el representado de los actos realizados por un “falsus procurator” (art. 1.259 CC). Por otra parte, atribuir una última palabra a quien puso en marcha un sistema de inteligencia artificial autónomo, para validar o no los acuerdos alcanzados por éste, según le convengan o no, es difícilmente conciliable con la seguridad del tráfico.

Además, la personalidad jurídica no implica sólo capacidad de decidir, sino también responsabilidad propia y patrimonio independiente. Desde esta perspectiva, se habrían de plantear las relaciones entre el principal y su propio sistema de inteligencia artificial en los términos propios del contrato de mandato, respondiendo el sistema frente a su principal por los contratos que hubiera realizado excediendo los términos del mandado dado (art. 1.714 CC). Esta responsabilidad del sistema de inteligencia artificial, como mandatario, frente a su hipotético principal parece poco más o menos que absurda o quimérica, por la sencilla razón de que un sistema de inteligencia artificial por muy persona jurídica que fuera no estará dotado de un patrimonio propio con el que responder, ni frente a su principal, ni frente a nadie.

En el estado actual de cosas, pretender justificar la eficacia vinculante de los acuerdos alcanzados por sistemas de inteligencia artificial autónomos en su condición de agentes contractuales dotado de personalidad jurídica capaces de vincular a su principal por la vía de la representación es un intento innecesario, artificioso y de efectos no bien calculados. En este sentido, el recurso a la personalidad jurídica algorítmica que estuvo presente en algunos trabajos preparatorios de la Unión Europea para una futura regulación de la Inteligencia Artificial, como la Resolución del Parlamento Europeo de 16 de febrero de 2017 con recomendaciones a la Comisión sobre normas de Derecho Civil sobre robótica, ha sido abandonado, siendo paradigmas de este nuevo enfoque, tanto la Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial de 28 de septiembre de 2022, como la Propuesta de Directiva del parlamento europeo y del consejo sobre responsabilidad por los daños causados por productos defectuosos, que hacen recaer la responsabilidad de los daños ocasionados por sistemas de inteligencia artificial, no en los propios sistemas, por más autónomos que sean, sino en los operadores y en los fabricantes, respectivamente.

**III.2.C. La vinculación contractual derivada de la responsabilidad**

Descartado el recurso a la personalidad jurídica de los sistemas de inteligencia artificial autónomos, quien emite la declaración negocial no es el sistema de inteligencia artificial, sino quien va a quedar vinculado por aquella, esto es, la persona que puso en marcha el sistema de inteligencia artificial para que realizase contratos en su interés. En definitiva, el sistema de inteligencia artificial no es un agente contractual, sino un instrumento, por muy sofisticado y autónomo que sea, de la persona que lo puso en marcha, empleado por ésta para emitir sus propias declaraciones negociales, sean éstas verdaderamente conscientes y, a posteriori, queridas o no por aquella persona.

En este sentido, los contratos así realizados son fruto de una libertad contractual delegada, pero no delegada en otra persona por la vía de la representación, sino en un instrumento con autonomía para tomar decisiones en aras a la consecución de un objetivo: maximizar el beneficio de la persona que lo puso en marcha, con el convencimiento de que este instrumento, según las circunstancias de un mercado dado, es más eficiente que ella misma o que cualquier otra persona que la pudiera representar.

Quien pone en marcha un sistema de inteligencia artificial para que realice contratos en su interés está exteriorizando al tráfico su voluntad de quedar vinculado por los mismos. Sin embargo, no existe un consentimiento, entendido en su consideración estricta como acto puramente consciente y voluntario, sobre todos y cada uno de los contratos que realiza su sistema de inteligencia artificial.

Lo verdaderamente relevante, entonces, es el valor que se atribuye a la voluntad en la declaración negocial y si puede existir una declaración negocial con efectos vinculantes que no esté fundada en la voluntad, sino en la responsabilidad o en la confianza suscitada en el tráfico por haber puesto en funcionamiento un sistema autónomo de contratación algorítmica.

La cuestión nos retrotrae al debate decimonónico entre las teorías voluntaristas y las declaracionistas. Sin pretender, ni tener capacidad para reducir a unas líneas la complejidad de un debate sostenido durante medio siglo por los más brillantes iusprivatistas del siglo XIX, tras los rescoldos de este debate se ha asumido por el moderno derecho de contratos que las declaraciones negociales vinculan por su mera exteriorización o significación objetiva, aunque faltare la voluntad o esta resultare viciada. Como manifestó el profesor De Castro, existen dos fuentes de responsabilidad negocial: la normal derivada de la autonomía de la voluntad y otra complementaria o excepcional, atribuida a una conducta negocialmente significativa, aun cuando se haya comprobado la falta o el vicio de la voluntad, porque quien actúa socialmente ha de hacerlo responsabilizándose de su conducta. Es lo que ocurre cuando una persona decide poner en marcha un sistema de inteligencia artificial autónomo para realizar contratos de manera que maximicen su beneficio, quedando perfectamente exteriorizada en el tráfico por este sólo hecho su intención de quedar vinculadas por tales contratos. Dotar a esta persona de una última palabra o de una voluntad de desvinculación de los acuerdos que, previa revisión humana, no le resulten satisfactorios iría en contra de la responsabilidad negocial que ha de predicarse de la confianza suscitada por aquella conducta.

La regla de la vinculatoriedad de los contratos realizados por medio de agentes electrónicos sin intervención humana ha sido expresamente positivizada en algunos ordenamientos. Así, por ejemplo, la Sección 14 de la Uniform Electronic Transactions Act de 1999 ya establecía que *“a* *contract may be formed by the interaction of electronic agents of the parties, even if no individual was aware of or reviewed the electronic agents' actions or the resulting terms and agreements.”* En derecho español no parece que haya que retorcer demasiado el sentido del art. 1.261 CC para entender que el consentimiento sobre el objeto y la causa de un contrato realizado por medio de un sistema de inteligencia artificial puede derivar de la intención socialmente exteriorizada de quien puso en marcha un sistema de inteligencia artificial para concluir contratos en su interés, de asumir las consecuencias de las decisiones adoptadas por dicho sistema, por más autónomo que este sea.

**IV. LA INEFICIA DE LOS CONTRATOS ALGORÍTMICOS**

**IV.1 Los contratos algorítmicos absurdos o patológicos**

Conforme a lo anteriormente expuesto, los acuerdos alcanzados por medio de sistemas de inteligencia artificial, ya sean preprogramados o autónomos son contratos con plenos efectos vinculantes para las partes que emplearon tales sistemas.

Otra cosa son las patologías de las que puedan adolecer los acuerdos alcanzados, que pueden provocar que tales acuerdos resulten clara y objetivamente absurdos y, en cuanto tales, no habrían sido realizados bajo ningún concepto por un agente humano razonable.

Estos acuerdos patológicos pueden tener distintas causas:

- Fallos en el diseño, por no haberse seleccionado correctamente los inputs o las circunstancias que deberían tener en cuenta por los sistemas de inteligencia artificial para tomar sus decisiones o no haberse definido de manera idónea las distintas hipótesis que hacen decidir al sistema en uno u otro sentido. La fase de diseño tiene evidentemente más relevancia en los sistemas determinísticos que en los sistemas autónomos.

- Fallos en la estrategia, por no haberse previsto un determinado comportamiento de los sistemas de inteligencia artificial de las contrapartes con las que han de interactuar.

- Cambios imprevisibles de las circunstancias respecto de las que fueron tenidas en cuenta en el momento del diseño del sistema, por ejemplo, el acaecimiento de una pandemia o de una guerra, que hacen que el sistema siga considerando como más eficientes aquellos resultados que en el nuevo contexto ya no lo son.

- Fallos en la programación de los algoritmos, por error en la selección del algoritmo, los parámetros utilizados, los pesos atribuidos a cada variable, por la mala calidad de los datos con los que se ha entrenado el sistema y, en general, por errores en el código que se inserta en el sistema para interpretar y procesar los datos y convertirlos en un resultado Son fallos imputables a los ingenieros que desarrollan el sistema que impiden que el mismo produzca los resultados queridos conforme a su diseño previo. Estos fallos son, obviamente, más fáciles de detectar en los sistemas determinísticos que en los sistemas autónomos.

- Fallos en las fuentes de alimentación de datos (inputs) del sistema, por ejemplo, un fallo de mantenimiento que impide que tengan entrada al sistema los precios de las transacciones en un determinado mercado, manejándose entonces el sistema con una información que ya está obsoleta.

- Infiltraciones de noticias falsas, como fue el caso del conocido como “Tweet Crash” de 2013, en el que la inclusión por unos hackers en la cuenta de Twitter de una agencia de noticias estadunidense de la noticia de que el presidente Obama había sufrido un atentado en la Casa Blanca motivó un desplome del Dow Jones Industrial Average.

- Fallos en el hardware o en el sistema de redes, que en operaciones que se cierran en milisegundos pueden ser determinantes.

- Fallos generados por estrategias “*dolosas o maliciosas*” de los sistemas de inteligencia artificial de la contraparte.

Conviene tener en cuenta además que los contratos patológicos transmiten a los sistemas de inteligencia artificial de otros operadores una información distorsionada sobre los precios susceptible de generar una reacción en cadena, especialmente significativa en los mercados financieros, por lo que el art. 48 de la Directiva 2014/65/UE, de 15 de mayo de 2014, relativa a los mercados de instrumentos financieros (MIFID II), establece una serie de medidas a modo de cortafuegos en los mercados regulados para evitar un efecto contagio, como la obligación impuesta a los **mercados regulados** de establecer mecanismos para rechazar las órdenes que sean manifiestamente erróneas, limitar temporalmente la negociación de un instrumento financiero si se produce una fluctuación significativa de su precio o incluso cancelar, alterar o corregir cualquier operación.

**IV.2. La insuficiencia del error vicio**

Es obvio que quien pone en marcha un sistema de inteligencia artificial para realizar contratos algorítmicos que le vinculen asume que alguno de estos contratos no hayan sido beneficiosos, porque la eficiencia de su uso se contempla de manera general, asumiendo que las operaciones concretas que no hayan sido eficientes, se compensen con las que sí lo son. Igualmente, se ha de asumir que en un contexto de evolución tecnológica constante los sistemas de los que se dota cada persona no siempre van a ser los más eficientes, en contraste con los que puedan utilizar las contrapartes y que de ese avance tecnológico se han de derivar beneficios para quien legítimamente utilice una tecnología más puntera. Todos estos casos deberían quedar fuera sin más de la órbita de los vicios negociales y asumirse que son el resultado de la aleatoriedad intrínseca a una determinada forma de contratar. Otra cosa es que un contrato algorítmico, a consecuencia de una de las patologías anteriormente descritas, sea desde un punto de vista objetivo absurdo o ilógico. Estos contratos pueden llegar a causar, además, daños sistémicos por transmitir a su vez una información incorrecta sobre el mercado a otros agentes humanos o a otros sistemas de inteligencia artificial.

En relación a estos contratos algorítmicos que han producido unos resultados objetivamente absurdos es necesario plantearse si su posible ineficacia debe enjuiciarse desde la perspectiva tradicional de los vicios del consentimiento, fundada en una incorrecta formación de la voluntad, o si por el contrario, dado que los acuerdos no son el resultado de una voluntad humana sino de una decisión mecanizada, los requisitos de su ineficacia deben valorarse desde una perspectiva distinta. No existen apenas precedentes jurisprudenciales al respecto, ni en el derecho español, ni en derecho comparado, aunque es de presumir que una vez que la contratación algorítmica se extienda a ámbitos distintos de los mercados financieros los litigios sobre ineficacia de contratos algorítmicos indeseados proliferen.

Uno de los pocos casos conocidos es el denominado caso Quoine, resuelto por la Corte de apelación de Singapore en sentencia de 24 de febrero de 2020 (Civil Appeal Nº 81/2019), que constituye un buen ejemplo de los resultados insatisfactorios que puede generar trasladar sin más el esquema de los vicios del consentimiento a los contratos algorítmicos:

La compañía BB2C había realizado a través de la plataforma de intercambio de criptomonedas Quoine un contrato de cambio de una determinada criptomoneda, Ethereum (ETH) por Bitcoins (BTC), siendo contrapartes la compañía Pulsar y un particular, Mr. Tomita. El tipo de cambio al que se hizo la transacción fue 10 BTC por cada ETH, lo que supuso una valoración de los ETH 250 veces superior a su valor de mercado en ese momento, que era de 1 EHT por cada 0,04 BTC. Obviamente, tal valoración de los ETH, que supuso un negocio redondo para BB2C, resultaba ilógica y fuera de cualquier parámetro razonable del mercado.

La patología de la transacción fue causada por un fallo en el sistema de la propia plataforma Quaine por labores de mantenimiento, que impidió que pudiera recibir información de otros mercados de intercambios de criptomonedas. En cuanto que la plataforma Quaine actuaba también como agente en su propio mercado, realizando operaciones en su propia plataforma para generar un mercado continuo (*market making*), transmitió a los sistemas de inteligencia artificial de las partes concernidas en el intercambio unos inputs erróneos sobre el valor de los BTC. Con arreglo a esta infravaloración de los BTCs, los sistemas de inteligencia artificial de las partes perjudicaras interpretaron, en primer lugar, que debían vender sus reservas de BTCs para saldar una posición negativa que en realidad no era tal, sino el resultado de una infravaloración errónea de esa criptomoneda y, por otro lado, que debían aceptar ofertas de intercambio de BTCs por un valor que, por la misma razón, estaba profundamente distorsionado.

Es decir, en el caso concreto la causa del contrato absurdo ni tan siquiera era imputable a los sistemas de inteligencia artificial de las perjudicadas, que funcionaron correctamente y adoptaron sus decisiones de acuerdo con la información que erróneamente les proporcionó el sistema empleado por la plataforma de intercambio.

Pues bien, a pesar de lo anterior, la Corte de Apelación de Singapore resolvió que el contrato era válido, tras pasarlo por el tamiz del error unilateral o unilateral mistake, que en los sistemas de *common law* es una suerte de dolo por omisión, que requiere, además de que el error sea relevante, que el mismo haya sido conocido por la contraparte. Desde esta óptica, la Corte de Apelación de Singapore entiende que el momento en que debe valorarse el estado mental de las partes: la incorrecta representación de la realidad por parte de quien sufre el error y el conocimiento de tal error por la contraparte, no es el momento en que el contrato se perfecciona, sino el momento en que actúa la voluntad de las partes, que no es otro sino el de la programación del sistema de inteligencia artificial. En consecuencia, para que este error hubiera sido relevante, la parte que se benefició del mismo, BB2C, en el momento de la programación de su sistema debería haber sido consciente de la debilidad del sistema de la contraparte y haber programado su sistema para obtener una ventaja de ello.

Retrotraer el momento de la valoración del error al momento de la programación de los sistemas de inteligencia artificial supone considerar que las transacciones que se produzcan por fallos ulteriores al momento de la programación, son en realidad irrelevantes, debiendo entonces asumir las partes contratantes los riesgos de tales fallos que, imputables o no a las partes, pasarían a formar parte del riesgo o la aleatoriedad intrínseca a esta forma de contratar.

Esta conclusión no se puede aceptar sin más como satisfactoria. Ciertamente, la utilización de sistemas de inteligencia artificial comporta riesgos, pero es cuestionable si debe comportar el riesgo de estar vinculado por un contrato absurdo, que sólo puede ser explicado como el resultado de un error en el funcionamiento de los sistemas, máxime cuando este error, como en el caso de Quoine, no resulta imputable a quien lo sufre.

Sin embargo, los vicios del consentimiento, tal y como los conocemos, esto es, fundados en la quiebra de la voluntad, más que en el valor objetivo de la declaración, no resultan útiles a tales efectos.

El error vicio, como incorrecta representación de la realidad, directamente no es predicable de los contratos realizados mediante sistemas algorítmicos autónomos, que adoptan sus propias decisiones y estrategias, al margen de la voluntad de la persona que los puso en marcha. Respecto de los sistemas algorítmicos pre-programados, sólo se podría hablar de una incorrecta representación de la realidad en relación a las circunstancias que no fueron tenidas en cuenta en el diseño del sistema, pero en tal caso es más que dudoso que el error superase los filtros de la esencialidad, por no recaer sobre las condiciones de la cosa objeto del contrato sino sobre motivos o circunstancias externas al mismo, y de la excusabilidad.

**IV. 3. El error obstativo evidente y no imputable**

Los fallos en las declaraciones de los sistemas de inteligencia artificial en nuestro marco mental secular, fundado en la distinción entre error vicio y error obstativo, de manera intuitiva se aproximan a los contornos del error obstativo. Sin embargo, la solución tradicionalmente asumida como válida en relación al error obstativo, conforme a la cual el contrato es inexistente en caso de que el destinatario, con arreglo a la buena fe, no pudiera desconocer el carácter errado de la declaración, no se puede extender sin más a los contratos realizados entre dos sistemas de inteligencia artificial, por la sencilla razón de que el destinatario de la declaración errada no es una persona física de quien se pueda predicar su mala fe en virtud de la apariencia indubitadamente errónea de la declaración, sino otro sistema de inteligencia artificial que actúa al margen de cualquier control humano y movido por la finalidad de maximizar el beneficio en las transacciones de la persona que lo puso en marcha y en cuyo interés actúa.

Desde esta perspectiva, no se puede hacer descansar la ineficacia del contrato solamente en la evidencia, desde una perspectiva objetiva (para cualquier observador imparcial con conocimiento y experiencia en el mercado en cuestión), de que la declaración de la contraparte es la consecuencia de un fallo del sistema, sino que ha de tenerse en cuenta también la imputabilidad del fallo a quien emite la declaración errónea, pues quien pone en marcha un sistema de inteligencia artificial para que realice contratos en su nombre debe asumir una responsabilidad en el diseño, programación, funcionamiento y mantenimiento del mismo.

En el ámbito concreto de los mercados financieros, art. 17 de la Directiva 2014/65/UE, de 15 de mayo de 2014, relativa a los mercados de instrumentos financieros (MIFID II) impone a las empresas de servicios de inversión que se dedique a la negociación algorítmica el deber de implantar sistemas y controles de riesgo adecuados a sus actividades y efectivos para garantizar que sus sistemas de negociación sean resistentes, tengan suficiente capacidad, se ajusten a los umbrales y límites apropiados y, sobre todo, por lo que ahora importa, que limiten o impidan el envío de órdenes erróneas. Igualmente, deben implantar mecanismos efectivos que garanticen la continuidad de las actividades en caso de disfunción de sus sistemas de negociación y asegurare de que sus sistemas sean íntegramente probados y convenientemente supervisados.

Este deber de diligencia prácticamente imposibilita en este ámbito el recurso a la ineficacia –salvo que lo sea por un engaño generado por el sistema de la contraparte, cuestión que trataremos seguidamente- pues entre los deberes de diligencia se impone el de establecer sistemas de control efectivos que impidan el envío de órdenes erróneas, siendo entonces indiferente que esta orden errónea traiga su causa en un fallo del sistema que sea imputable o no a quien la emite; simplemente, las empresas que prestan estos servicios de emisión deben evitar que tales órdenes se emitan, pues recordemos que la emisión de una orden errónea puede tener un efecto sistémico en el mercado financiero.

No obstante, teniendo en cuenta que la finalidad de estas cautelas de la Directiva MIFID II es la de proteger los mercados financieros, evitando que las órdenes erróneas provoquen una distorsión de precios que tenga un rápido efecto propagatorio, en otros ámbitos distintos al de los mercados financieros, la regla de la imputabilidad de la declaración errónea podría ser menos severa. Así, podrían considerarse como no imputables las declaraciones negociales erróneas que se deban a fallos ajenos al propio sistema, como los causados por la toma en consideración de datos externos (inputs) erróneos, no detectables razonablemente por el propio sistema (como aconteció en el caso Quoine).

En la medida en que los consumidores pudieran hacer uso de sistemas de inteligencia artificial diseñados y desarrollados por terceros para realizar contratos, el criterio de imputabilidad podría ser todavía más laxo, no siendo imputables al consumidor que utiliza el sistema los fallos que se debieran a defectos en el diseño o en la programación del sistema. No obstante, el principal recurso a disposición del consumidor para conseguir ineficacia de los contratos algorítmicos manifiestamente erróneos no será el de la nulidad por error obstativo, sino el ejercicio del derecho de desistimiento, en cuanto que se tratará de una venta a distancia (art. 102 TR-LGDCU).

**IV.4. El dolo algorítmico**

A lo largo de este trabajo se ha hecho referencia a comportamientos “*dolosos o maliciosos*” de los sistemas de inteligencia artificial de la contraparte, siempre entre comillas, porque el dolo, en cuanto que engaño, es un acto o comportamiento típicamente humano. La cuestión, por lo tanto, es si se puede predicar la ineficacia de un contrato realizado mediando una actuación de un sistema de inteligencia artificial, que de haber sido realizada por un humano, sería calificada como una maquinación insidiosa en el sentido del art. 1.269 CC. Esto es, si tiene relevancia en el plano de la ineficacia contractual el engaño de una máquina a otra máquina, porque los sistemas de inteligencia artificial, ya sea porque hayan sido preordenados para ello o porque hayan adquirido de manera autónoma la habilidad para hacerlo tienen capacidad para generar indicios o informaciones incorrectas, que induzcan al sistema de inteligencia artificial de la contraparte a actuar de una manera errónea.

La cuestión queda fuera del ámbito del proyectado art. 5.1.a de la Propuesta de Reglamento de Inteligencia Artificial de 21 de abril de 2021, que en la redacción dada por las enmiendas aprobadas por el parlamento europeo el 14 de junio de 2023, prohíbe la introducción en el mercado, la puesta en servicio o la utilización de sistemas de inteligencia artificial que se sirva de técnicas deliberadamente engañosas con el objetivo de alterar el comportamiento de una persona produciendo un perjuicio significativo a esa persona, pues la norma hace referencia expresa a que el destinatario del acto de engaño sea una persona.

No obstante, cuando el engaño sea causado por un sistema de inteligencia artificial determinístico que obedezca a un diseño preordenado a engañar, en cuanto que el engaño es imputable a la persona que crea y utiliza el sistema de inteligencia artificial para tal fin, será esta la responsable del acto doloso y, por lo tanto, debe pechar con las consecuencias de la ineficacia del contrato.

Cuando la estrategia de engaño no obedezca a un diseño, sino que haya sido aprendida de manera autónoma por el sistema de inteligencia artificial para maximizar el beneficio de la persona que lo puso en marcha, consideramos que igualmente el contrato ha de ser susceptible de ser anulado por vicio de dolo, si la actuación del sistema autónomo de haber sido realizada por un humano pudiera considerarse como un engaño o maquinación insidiosa en los términos del art. 1.269 CC. En este caso, es necesario acudir a la ficción de considerar a las máquinas como si fueran agentes humanos. Lo contrario, esto es, considerar legítima la ventaja obtenida por el sistema de inteligencia artificial mejor dotado para general estrategias de engaño a sus contrapartes, implicaría reconocer un ámbito de la contratación inmune al fraude.

**BIBLIOGRAFÍA**

- BÁRCENA SUÁREZ, N, “Inteligencia artificial y personalización. Una aproximación al futuro del Derecho de contratos en *Estudios de Derecho Contractual Europeo: nuevos problemas, nuevas reglas*, (Dir. F. Gómez Pomar e I. Fernández Chacón), Aranzadi 2022.

- CARRASCO PERERA, A. “A propósito de un trabajo de Gunter Teubner sobre la personificación civil de los agentes de inteligencia artificial avanzada”, *Centro de Estudios de Consumo*, 11 de diciembre de 2019.

- CHABOUD, A., CHIQUOINE, B., HJALMARSSON, E., VEGA, C.,“Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market”, *International Finance Discussion Papers*, Nº 980 29 September 2009 <<http://www.federalreserve.gov/pubs/ifdp/2009/980/ifdp980.pdf>>).

- COCA PAYERAS, M., “Las iniciativas de la Unión Europea sobre Inteligencia Artificial: de la persona electrónica, al difícil equilibrio entre la necesidad de impulsarla y evitar sus riesgos”, *Revista de Derecho Civil*, Vol. X, Nº 2, junio 2023.

- DE CASTRO Y BRAVO, F. *El negocio jurídico,* Civitas, Madrid (reimpresión), 1985.

- FARRANDO MIGUEL, I., “Algoritmos en el mercado de valores y protección del inversor: “robo advisors””, *La regulación de los algoritmos* (Dir. Huergo Lora, A.), Aranzadi, Navarra, 2020.

- GIBBONS, K., GLEESON, S., SCOTT, K., “Legal considerations around smarts contracts: contracts between computer programs”, *Clifford Chance*, September, 2021.

- HUERGO LORA, A. “Una aproximación a los algoritmos desde el derecho administrativo”, *La regulación de los algoritmos* (Dir. Huergo Lora, A.), Aranzadi, Navarra, 2020.

- KORSMO, C. R, “High-Frequency Trading: A Regulatory Strategy” . *RICH. L. REV*. 523, 2014.

- NAVAS NAVARRO, S., “Derecho e inteligencia artificial desde el diseño. Aproximaciones”, Inteligencia Artificial. Tecnología Derecho, Tirant lo Blanch, Valencia, 2017.

- NAVAS NAVARRO, S., *Daños ocasionados por sistemas de inteligencia artificial. Especial atención a su futura regulación*, Comares, Granada, 2022.

- PALMA ORTIGOSA, A., *Decisiones automatizadas y protección de datos. Especial atención a los sistemas de inteligencia artificial*, Dykinson, Madrid, 2022.

- SÁNCHEZ LORENZO, S. *El derecho inglés y los contratos internacionales*, Tirant lo Blanch, Valencia, 2013.

- SCHOLZ, L. H., “Algorithmic Contracts”, *STAN. TECH. L. REV*. 128, 2017.

-TEUBNER, G “Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagentem”, *Archiv für die Civilistische Praxis*, 218, 2018.

- VON UNGERN-STERNBERG, A., “Discriminatory AI and the Law. Legal Standards for Algorithmic Profiling”, *The Cambridge Handbook of Responsible Artificial Intelligence*, Cambridge University Prees, 2022