

Reglamento de inteligencia artificial e intervención pública en las relaciones laborales*

Artificial intelligence act and public intervention in labor relations

José María Goerlich Peset

*Catedrático de Derecho del Trabajo y Seguridad Social
Universitat de València*

ORCID ID: 0000-0002-2910-2153

doi: 10.20318/labos.2024.9040

Resumen: Aunque la atención de los juristas del trabajo, y del propio legislador, se ha centrado en el uso de la inteligencia artificial por las empresas, en el marco del contrato de trabajo, lo cierto es que en las dos últimas décadas hemos asistido a un imparable incremento de su utilización por las entidades públicas, administrativas y judiciales. El presente trabajo describe, en primer lugar, la evolución normativa interna en relación con esta última. Se analiza, en segundo lugar, el impacto que sobre la situación actual puede tener la aprobación del Reglamento de Inteligencia Artificial. Finalmente, se valora la situación resultante. En este último terreno, se destacan las insuficiencias existentes, sobre todo en relación con el uso administrativo de la IA, y se ofrecen algunas posibilidades para superarlas.

Palabras clave: Inteligencia artificial, relaciones laborales, administración laboral, proceso laboral.

Abstract: Although the attention of labor jurists, and of the legislator himself, has been focused on the use of artificial intelligence by companies, within the framework of the employment contract, the fact is that in the last two decades we have witnessed an unstoppable increase in its use by public, administrative and judicial entities. This paper describes, firstly, the internal normative evolution in relation to the latter. Secondly, it analyzes the impact that the approval of the Artificial Intelligence Act may have on the current situation. Finally, the resulting situation is assessed. In this last area, the existing inadequacies are highlighted, especially in relation to the administrative use of AI, and some possibilities are offered to overcome them.

Keywords: Artificial intelligence, labor relations, labor administration, labor process.

*Trabajo realizado en el marco del proyecto de investigación, Algoritmos extractivos y neuroderechos. Retos regulatorios de la digitalización del trabajo (PID2022-139967NB-I00).

El título original que tenía esta ponencia en el programa del Congreso era “Las sanciones automatizadas en la Inspección de Trabajo”. En mi intervención oral, ya amplíe la óptica con el ánimo de incluir el conjunto en las decisiones administrativas automatizadas que son, como sabemos, muy frecuentes en el orden social, y analizar cómo impacta en su régimen jurídico el nuevo Reglamento de Inteligencia Artificial (Reglamento [UE] 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial, DOUE 12 julio 2024 –en adelante, RIA–). En esta versión escrita, que, por otra parte, tiene un tono similar al de la presentación oral, añadiré algunas consideraciones adicionales en relación con las posibilidades de utilización de la inteligencia artificial en el marco del proceso laboral.

Mi plan es abordar el tema en tres partes. La primera describe brevemente la situación normativa interna en el momento de la aprobación del Reglamento europeo. La segunda analiza sus disposiciones para determinar cómo podrían afectar a aquella. Finalmente, intento valorar si el conjunto resultante es o no adecuado en atención a los intereses que están en juego.

1. La creciente presencia de la digitalización en los entes públicos

Diría que, en los últimos 30 años, hemos asistido a una creciente incorporación de las nuevas tecnologías, no solo en las empresas a las que se ha dedicado la atención de las ponencias anteriores, sino también en las entidades públicas.

Por lo que se refiere a las Administraciones Públicas en general, podemos fijar convencionalmente el arranque de este proceso en la reforma del procedimiento administrativo de 1992. Según su Exposición de Motivos, la Ley 30/1992, de 26 de noviembre, de Régimen Jurídico de las Administraciones Públicas y del Procedimiento Administrativo Común, estaba llamada a protagonizar la “abierto incorporación de las técnicas informáticas y telemáticas en la relación ciudadano-Administración”. Su art. 45, dedicado a la “incorporación de medios técnicos”, afirmaba que las Administraciones Públicas impulsarían “el empleo y aplicación de las técnicas y medios electrónicos, informáticos y telemáticos, para el desarrollo de su actividad y el ejercicio de sus competencias, con las limitaciones que a la utilización de estos medios establecen la Constitución y las Leyes” (apartado 1); y a continuación preveía que “los programas y aplicaciones electrónicos, informáticos y telemáticos que vayan a ser utilizados por las Administraciones Públicas para el ejercicio de sus potestades, (habrían) de ser previamente aprobados por el órgano competente, quien deberá difundir públicamente sus características”.

Las posteriores reformas siguieron avanzando en este sentido. La regulación de la Ley 30/1992 fue sustituida por la contenida en el art. 39 Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos, que preveía ya la posibilidad de adoptar decisiones automatizadas en el ámbito administrativo. No se trataba solo de la utilización de estos medios sino de que fueran estos los que adoptaran la decisión. Para este caso, se preveía que “deberá establecerse previamente el órgano u órganos

competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado responsable a efectos de impugnación”. Esta regulación se ha incorporado al art. 41.2 de la vigente Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. En su apartado primero, el precepto define, además, lo que debe entenderse por actuación administrativa automatizada, que resulta ser “cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público”. Existen algunas concreciones adicionales de este régimen, tampoco muchas, en el art. 13 del Reglamento de actuación y funcionamiento del sector público por medios electrónicos, aprobado por RD 203/2021, de 30 de marzo.

Es posible advertir cómo en este proceso se ha ido redimensionando el régimen jurídico. Y ello, en un doble sentido. De un lado, en relación con las actuaciones afectadas, se ha pasado de la genérica utilización de “programas y aplicaciones electrónicos, informáticos y telemáticos... por las Administraciones Públicas para el ejercicio de sus potestades” (Ley 30/1992) a la existencia de una actuación “realizada íntegramente a través de medios electrónicos” sin intervención directa de humano (Ley 40/2015). De otro, por lo que se refiere a las obligaciones informativas a cargo de las Administraciones, si hasta 2007 había que comunicar los programas y aplicaciones electrónicos, informáticos y telemáticos que fueran a ser utilizados por las Administraciones Públicas para el ejercicio de sus potestades, desde entonces la información parece más limitada. Volveré al final sobre este tema.

Las autoridades que ejercitan competencias en el ámbito social no han sido ajenas a esta evolución, como he analizado con más detalle en otro lugar (Goerlich Peset, 2021). Con el soporte que les proporcionaban las citadas normas han ido modernizando sus estructuras y sumándose a la imparable digitalización. Las actuaciones automatizadas se han ido incorporando, primero, a la gestión de la Seguridad Social y, luego, a la Inspección de Trabajo y Seguridad Social. En el primer terreno, desde mediados de la década de los 90, informática y telemática se han puesto al servicio de las actividades instrumentales que desarrolla la Seguridad Social: actos de encuadramiento y recaudación han sido afectados por sucesivas normas que implican su completo tránsito a la esfera digital. Esto proceso se corona a mediados de la siguiente década, cuando el sistema tradicional de recaudación basado en la autoliquidación se sustituye por el actualmente vigente: la aprobación de la Ley 34/2014, de 26 de diciembre, de medidas en materia de liquidación e ingreso de cuotas de la Seguridad Social, presupone, de un lado, una formidable capacidad de gestión de la información por parte de las Entidades Gestoras y garantiza, de otro, un igualmente formidable caudal de datos a disposición del uso administrativo de la inteligencia artificial.

La incorporación normativa de la posibilidad de adoptar decisiones automatizadas se produce en 2009, para las prestaciones de desempleo (disp. adic. 46ª LGSS-1994, redacción aprobada por RDL 10/2009, de 13 de agosto) y se amplía después, por la

ya citada Ley 30/2014, al conjunto de las prestaciones, excluidas las pensiones en su modalidad no contributiva. Incorporadas estas reglas al art. 130 del texto refundido aprobado en 2015, este ha vuelto a ser objeto de modificación por el RDL 2/2021, de 26 de enero, de refuerzo y consolidación de medidas sociales en defensa del empleo. En su nueva redacción, actualmente vigente, extiende la posibilidad de adoptar decisiones automatizadas al conjunto de la gestión del sistema de Seguridad Social, con la única salvedad de las prestaciones no contributivas. Por lo demás, aunque separado del régimen administrativo común por razones vinculadas a la tradición, las previsiones de su segundo párrafo para las decisiones automatizadas en el ámbito de la Seguridad Social reproducen prácticamente al pie de la letra las previsiones del art. 45 Ley 40/2015.

Algo parecido ha venido ocurriendo en el ámbito de la Inspección de Trabajo. Aunque aparentemente tiene poco que ver con el tema, llama mucho la atención la propia evolución de la definición legal del sistema de Inspección en la ley reguladora de 1997 y en la que la sustituyó en 2015. En la primera, Ley 42/1997, de 14 de noviembre, Ordenadora de la Inspección de Trabajo y Seguridad Social, el art. 1.1 nos decía que venía constituido por “el conjunto de principios legales, normas, órganos, funcionarios y medios materiales que contribuyen al adecuado cumplimiento de las normas” que entran en su ámbito de competencia. Veinte años después, el art. 1.1 Ley 23/2015, de 21 de julio, Ordenadora del Sistema de Inspección de Trabajo y Seguridad Social, indica que lo está por “el conjunto de principios legales, normas, órganos, personal y medios materiales, incluidos los informáticos, que contribuyen al adecuado cumplimiento de la misión que tiene encomendada”. Encontramos una diferencia significativa entre los dos preceptos: la relevante referencia a “los medios informáticos” que no eran aludidos en la norma de 1997 y aparecen en la de 2015. Se resalta de este modo su papel central en la actuación inspectora. Por supuesto, no puede dejar de traerse a colación la reforma de la LISOS por el ya citado RDL 2/2021. En su redacción vigente, el 53.1.a) LISOS prevé, nada más y nada menos, la posibilidad de que un acta de inspección sea emitida a través de una decisión administrativa automatizada generada por vía algorítmica. Con posterioridad, el RD 688/2021, de 3 de agosto, ha modificado el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social (RD 928/1998) y establecido una serie de reglas específicas para estas actas automatizadas –que, por cierto, hasta donde yo alcanzo, aún no se han puesto en marcha–.

Me he centrado, hasta ahora, en los usos administrativos de la IA que han cristalizado en concretas normas jurídicas, legales o reglamentarias. Pero las entidades que hemos considerado hacen uso también de la IA en terrenos diferentes que no dejan trazas formales –o las dejan de otra naturaleza–. Este uso no da lugar a decisiones administrativas automatizadas en el sentido expuesto sino que se mueve en el terreno de la inteligencia artificial predictiva que viene siendo utilizada, tanto por la Seguridad Social como por la Inspección de Trabajo, para la detección de posibles espacios de «oscuridad». Sabemos, en este sentido, que la Seguridad Social emplea este tipo de IA para detectar supuestos o bolsas de fraude, en el terreno de la recaudación y también en el del control de prestacio-

nes –en relación, por ejemplo, con la incapacidad temporal–. Sin embargo, las noticias que tenemos de este uso son menores. En los casos de decisiones automatizadas formales la sede virtual (<https://sede.seg-social.gob.es/wps/portal/sede/sede/Inicio/Normativa-yLegislacion>) incluye un catálogo de resoluciones administrativas que informan sobre la utilización de determinadas herramientas algorítmicas y cuáles son sus características generales. Por el contrario, en el caso de la IA predictiva, la información que tenemos es indirecta a través de la literatura especializada (Aibar Bernard, 2020; Redondo Rincón, 2020). Es seguro, sin embargo, que se está generalizando como ha puesto de manifiesto la AISS en fecha relativamente reciente (2019, p. 34).

En cuanto a su uso en el ámbito de la Inspección, sabemos que existe una “herramienta de lucha contra el fraude”, o sea un conjunto de aplicaciones que permiten hacer minería de datos y, tras su tratamiento, detectar sectores o espacios en los que concentrar la actuación inspectora. Tenemos bastante información sobre sus utilidades prácticas. A través del buscador del BOE, detectamos su presencia en los sucesivos planes que la Inspección de Trabajo ha ido publicando en los últimos años. Se alude a ella, por ejemplo, en el Plan estratégico 2018/2020 (BOE 19 abril 2018) y en el inmediato Plan director por un trabajo digno 2018-2020 (BOE 28 julio 2108). El último Plan estratégico conocido, para el período 2021/2023 (BOE 3 diciembre 2021), contiene más de 30 referencias, tanto en los objetivos como en las actuaciones. Es más, tiene sustantividad orgánica como “unidad del OEITSS” (actuación 32.2), con funciones de planificación (actuaciones 30.1 y 30.2) y en el terreno de las políticas de personal (22.4) y las reformas orgánicas (23.3). Por lo que se refiere a los objetivos pueden destacarse los núms. 30 –“Mejorar la planificación de las actuaciones inspectoras, utilizando las herramientas más avanzadas de inteligencia artificial”– y 32 –“Adecuar las infraestructuras, medios materiales, los sistemas, redes, aplicaciones y equipamientos informáticos, así como fortalecer la Herramienta de Lucha contra el Fraude”–. En cuanto a las actuaciones, sale en varias de las recogidas en los diferentes Ejes –salarios (1.1); tiempo de trabajo (1.2); contratación (1.3), convenios (1.4), falsos autónomos (1.7); diferentes aspectos de PRL (2.1 y 2.3); trabajo no declarado (3.1); discriminación por razón de sexo (5.1) u otras causas (6.1); recaudación (8.1, 8.5); inaplicación (19.1)–.

Por último, también la actividad jurisdiccional se ha visto afectada por la digitalización (para un análisis de conjunto, Nores Torres, 2023 con detalladas referencias adicionales). La evolución de la redacción del art. 230 LOPJ es muy significativa del proceso. Si originalmente el precepto se refería al uso de “medios técnicos de documentación y reproducción, siempre que ofrezcan las debidas garantías de autenticidad”, desde 1994 (LO 16/1994) se ha dado entrada al de “cualesquiera medios técnicos, electrónicos, informáticos y telemáticos, para el desarrollo de su actividad y ejercicio de sus funciones”; y, además, se ha posibilitado la relación entre interesados y órganos judiciales y garantizado la autenticidad de la documentación y la identificación de sus autores. Si inicialmente el uso de estos medios era voluntario, con posterioridad se ha hecho obligatorio para los órganos judiciales (LO 7/2015) y, después, para las personas que recurran a ellos (LO 4/2018).

Estos cambios en la norma orgánica han ido acompañados de otros más detallados a lo largo del período que se considera (entre otras disposiciones, leyes 18/2011 y 42/2015). Recientemente, el extensísimo RDL 6/2023, de 19 de diciembre, por el que se aprueban medidas urgentes para la ejecución del Plan de Recuperación, Transformación y Resiliencia en materia de servicio público de justicia, función pública, régimen local y mecenazgo, dedica la friolera de ¡100 artículos! a las diferentes cuestiones implicadas, estableciendo una regulación aplicable a todos los órdenes jurisdiccionales. Heredero en este punto del proyecto de ley de Medidas de Eficiencia Digital del Servicio Público de Justicia decaído como consecuencia de la disolución de las Cortes en primavera de 2023, esta parte del RDL 6/2023 se presenta como “una herramienta normativa completa, útil, transversal y con la capacidad suficiente para dotar a la Administración de Justicia de un marco legal, coherente y lógico en el que la relación digital se descubra como una relación ordinaria y habitual, siendo la tutela judicial efectiva en cualquier caso la prioridad absoluta, pero hallando bajo esta cobertura de normas y reglas un nuevo cauce, más veloz y eficaz, que coadyuvará a una mejor satisfacción de los derechos de la ciudadanía”.

Entre otras cuestiones, la nueva regulación dedica su atención a la utilización de la inteligencia artificial en el marco de las actuaciones jurisdiccionales; y lo hace estableciendo importantes condicionantes (cfr. Nores Torres, 2024a, p. 232 ss. y 2024b, p. 101 ss.). De un lado, la posibilidad de actuaciones algorítmicas autónomas queda confinada a los aspectos instrumentales: conforme al art. 56.2 RDL 6/2023, únicamente se admite “la automatización de las actuaciones de trámite o resolutorias simples, que no requieren interpretación jurídica”; junto a ellas se admiten las actuaciones proactivas, esto es aquellas que se inician de forma autónoma “por los sistemas de información sin intervención humana, que aprovechan la información incorporada en un expediente o procedimiento de una Administración Pública con un fin determinado, para generar avisos o efectos directos a otros fines distintos, en el mismo o en otros expedientes, de la misma o de otra Administración Pública” (art. 56.3). Fuera de estos casos, el uso de la inteligencia artificial en el ámbito judicial queda limitado a las denominadas “actuaciones asistidas” (art. 57), que implican la posible generación de borradores de documentos complejos. Queda explícitamente prohibido que estos borradores constituyan, “sin validación de la autoridad competente”, una resolución judicial o procesal; y se impone que solo puedan generarse “a voluntad del usuario” que además ha de poder modificarlos “libre y enteramente”.

2. Reglamento de Inteligencia Artificial y actuaciones públicas

Descrito sucintamente el panorama de la situación previa, conviene abordar, como segunda cuestión, el impacto que puede desplegar sobre ella la reciente aprobación del Reglamento de Inteligencia Artificial. Avanzando la conclusión, diría que será limitado. Esto es así, creo, por dos razones: la primera es la propia noción de inteligencia artificial; la segunda, la forma que se aborda su regulación. Vayamos por partes.

Hay una cosa que está clara: el uso de la inteligencia artificial por las entidades públicas está afectada por el RIA. No puede alcanzarse otra conclusión si se miran las definiciones de proveedor de sistemas de IA [art.3.3) RIA] o responsable de su despliegue [art.3.4) RIA], que incluyen a “persona física, jurídica o autoridad pública, órgano u organismo”. Es claro, por tanto, que subjetivamente las entidades públicas están afectadas, tanto si desarrollan este tipo de sistemas como si los utilizan “bajo su propia autoridad”. Ello no obstante, es posible que no todos los usos que hemos considerado en el apartado anterior quepan en la noción de inteligencia artificial que utiliza el art. 3.1) RIA. Este precepto la conceptúa como “un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales”. Entiendo que no existe equivalencia entre esta noción y la de decisión automatizada que hemos considerado en el apartado anterior.

Esta falta de equivalencia es doble. Por un lado, es posible que la noción jurídica de IA no incluya todos los supuestos de decisiones automatizadas concebibles, puesto que cabe pensar que existan algunas de estas, de mayor o menor sencillez, en las que no se detecta el requisito de la autonomía al que se refiere el precepto para considerar que estamos ante un sistema de IA. El extensísimo preámbulo del RIA permite alcanzar la conclusión de que los “distintos niveles de autonomía” y la “capacidad de adaptación tras el despliegue” excluyen de tal noción “los sistemas basados en las normas definidas únicamente por personas físicas para ejecutar automáticamente operaciones”; se hace preciso además que tengan “capacidad de inferencia” lo que “trasciende el tratamiento básico de datos al permitir el aprendizaje, el razonamiento o la modelización” (considerando 12). De este modo, muy probablemente no constituye un sistema de IA la herramienta informática que permite la emisión automatizada de una reclamación de deuda ante el impago de cotizaciones con base en el art. 33 LGSS. Por otro lado, pero ahora en un sentido diverso, la amplitud con la que se conciben los “resultados de salida” hace que los sistemas de IA desborden la estricta noción de decisiones automatizadas que ya hemos examinado. Las “predicciones, contenidos, recomendaciones o decisiones que puedan influir en entornos físicos o virtuales” a las que se refiere el art. 3.1) RIA permiten reconducir al ámbito de aplicación del nuevo reglamento actuaciones que no se toman en el marco de un proceso administrativo en el sentido del artículo 41 Ley 40/2015 –acaso las remisiones masivas relacionadas con vicisitudes de la contratación que hemos visto estos años–, así como la planificación de actividades inspectoras o de control por la seguridad social.

En otro orden de consideraciones, el modelo regulatorio del nuevo Reglamento se basa en una visión positiva de la IA. En esta línea, el preámbulo señala que “puede proporcionar ventajas competitivas esenciales a las empresas y facilitar la obtención de resultados positivos desde el punto de vista social y medioambiental en los ámbitos de la asistencia sanitaria, la agricultura, la seguridad alimentaria, la educación y la formación, los medios de comunicación, el deporte, la cultura, la gestión de infraestructuras, la ener-

gía, el transporte y la logística, los servicios públicos, la seguridad, la justicia, la eficiencia de los recursos y la energía, el seguimiento ambiental, la conservación y restauración de la biodiversidad y los ecosistemas, y la mitigación del cambio climático y la adaptación a él, entre otros, al mejorar la predicción, optimizar las operaciones y la asignación de los recursos, y personalizar las soluciones digitales que se encuentran a disposición de la población y las organizaciones” (§ 4). De este modo, como regla general, su distribución y uso son libres, salvo que se tope con una prohibición (art. 5 RIA) o, en razón de la materia a la que afecta, se esté en presencia de un sistema de alto riesgo, cuya puesta en funcionamiento queda sujeta a unas cautelas especiales (art. 6 RIA).

El repaso de estos condicionantes a distribución y despliegue de sistemas de IA muestra que las cuestiones que nos preocupan en este momento solo están presentes de manera limitada. Esto es bastante claro con las prácticas de IA prohibidas listadas en el art. 5 RIA, que no incluyen ninguna de las que ahora estamos analizando. En cuanto a la enumeración de sistemas de IA de alto riesgo contenida en el Anexo III, al que se remite el art. 6 RIA, sí encontramos algunas, aunque en modo alguno son coextensas con el ámbito que consideramos.

Por lo que se refiere al uso administrativo de la IA, las referencias del Anexo a la materia social no son muchas. Al margen la inclusión en el apartado 7 de los sistemas que afectan a “Migración, asilo y gestión del control fronterizo” –que pudieran tener cierto impacto colateral en las gestiones de la ITSS relacionadas con el trabajo de extranjeros–, cabe destacar la previsión de su apartado 5 que incluye entre los de alto riesgo los “sistemas de IA destinados a ser utilizados por las autoridades públicas o en su nombre para evaluar la admisibilidad de las personas físicas para beneficiarse de servicios y prestaciones esenciales de asistencia pública, incluidos los servicios de asistencia sanitaria, así como para conceder, reducir o retirar dichos servicios y prestaciones o reclamar su devolución”. La razón de la clasificación de estos sistemas de IA entre los de alto riesgo la ofrece el § 58 de los considerandos del preámbulo del RIA. Su utilización en este terreno, se argumenta, “podría tener un efecto considerable en los medios de subsistencia de las personas y vulnerar sus derechos fundamentales, como el derecho a la protección social, a la no discriminación, a la dignidad humana o a la tutela judicial efectiva”. Sin embargo, “el presente Reglamento no debe obstaculizar el desarrollo y el uso de enfoques innovadores en la Administración, que podrían beneficiarse de una mayor utilización de sistemas de IA conformes y seguros”. La forma de equilibrar ambos intereses es su calificación como de alto riesgo.

Ello impacta, por supuesto, en la gestión de la seguridad social, si bien no incluye todo el ámbito de decisiones automatizadas en este terreno conforme al art. 130 LGSS pues, al centrarse en el acceso o retirada de servicios o prestaciones, deja al margen “los procedimientos de afiliación, cotización y recaudación”. Hay que entender, por otro lado, que extiende su ámbito de aplicación más allá de lo que venimos considerando gestión de la seguridad social, toda vez que concesión, reducción o retirada de dichos servicios y prestaciones o la reclamación de su devolución son también objeto de actuaciones sancionadoras en los términos de los arts. 17, 24 ss. y 47 LISOS. En fin, la noción

de IA que maneja el Reglamento impone, como se desprende de las consideraciones desarrolladas más arriba, que las herramientas de IA predictiva que se empleen en este terreno, pese a no dar origen a decisiones automatizadas –o a “resoluciones”, conforme a la nomenclatura empleada por el art. 130.1 LGSS– se sujeten a sus prescripciones.

En cuanto a las restantes cuestiones vinculadas al poder administrativo sancionador en materia social no quedan afectadas por el Anexo III. Es verdad que su apartado 6 se dedica a los sistemas de IA relacionados con la “garantía del cumplimiento del Derecho, en la medida en que su uso esté permitido por el Derecho de la Unión o nacional aplicable”. Sin embargo, el Reglamento de Inteligencia Artificial está pensando exclusivamente en las cuestiones penales. Se parte de que “las actuaciones de las autoridades garantes del cumplimiento del Derecho que implican determinados usos de los sistemas de IA se caracterizan por un importante desequilibrio de poder y pueden dar lugar a la vigilancia, la detención o la privación de libertad de una persona física, así como tener otros efectos negativos sobre los derechos fundamentales consagrados en la Carta” (§ 59). Y, sobre esta base, las definiciones normativas se centran en aquellas. En esta línea, el art. 3.45) RIA define la «autoridad garante del cumplimiento del Derecho» como “a) toda autoridad pública competente para la prevención, la investigación, la detección o el enjuiciamiento de delitos o la ejecución de sanciones penales, incluidas la protección frente a amenazas para la seguridad pública y la prevención de dichas amenazas, o b) cualquier otro organismo o entidad a quien el Derecho del Estado miembro haya confiado el ejercicio de la autoridad pública y las competencias públicas a efectos de prevención, investigación, detección o enjuiciamiento de delitos o ejecución de sanciones penales, incluidas la protección frente a amenazas para la seguridad pública y la prevención de dichas amenazas”. En esta misma línea la «garantía del cumplimiento del Derecho» es definida en el apartado 46) como “las actividades realizadas por las autoridades garantes del cumplimiento del Derecho, o en su nombre, para la prevención, la investigación, la detección o el enjuiciamiento de delitos o la ejecución de sanciones penales, incluidas la protección frente a amenazas para la seguridad pública y la prevención de dichas amenazas”. Por lo demás, el apartado 6 del Anexo III ni siquiera contempla todos los aspectos penales: solo ciertos aspectos que despliegan efectos particularmente intensos sobre las personas. A la postre, el uso administrativo de la inteligencia artificial no parece quedar afectado por el nuevo Reglamento.

Diferente es, en fin, el caso del uso judicial de la inteligencia artificial, al que se refiere el apartado 8 del anexo III cuya rúbrica es “Administración de justicia y procesos democráticos”. Se incluyen en su apartado a) los “sistemas de IA destinados a ser utilizados por una autoridad judicial, o en su nombre, para ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la garantía del cumplimiento del Derecho a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios”. El uso del verbo “ayudar” implica que el RIA no está pensando en sistemas que generen decisiones judiciales automáticas pues presupone que son adoptadas por la “autoridad judicial”. Estamos pues pensando en IA generativa, relacionada con la interpretación del derecho, lo que nos sitúa en el terreno

de lo que legalmente se denomina “actuaciones asistidas” (art. 57 RDL 6/2023). Desde este modo, los automatismos procedimentales o las “actuaciones proactivas” (art. 56 RDL 6/2023) no serían considerados sistemas de alto riesgo

Conviene reparar, por último, aunque no sea por ello menos importante, en que la consideración de un sistema de IA como de alto riesgo no impide su utilización; únicamente sujeta su distribución y utilización a determinadas garantías establecidas en los arts. 8 ss. Estas afectan, de un lado, a su configuración e incluyen, aparte los requisitos de “precisión, solidez y ciberseguridad” (art. 15), el establecimiento y mantenimiento de un sistema de gestión de riesgos (art. 9) –del que resulta complementaria la “conservación de registros” (art. 12)–, el control de los datos utilizados en el entrenamiento (art. 10), las obligaciones de documentación técnica (art. 11) y transparencia (art. 13) y, en fin, la “supervisión humana” (art. 14). De otra parte, comprenden específicas obligaciones para distribuidores y responsables del despliegue (arts. 16 ss.). No resulta posible, ni seguramente necesario, entrar al detalle sobre todo ello, sin perjuicio de que pueda ser traído a colación cuando sea necesario.

3. Una valoración de conjunto

Hecha esta reflexión sobre la regulación del RIA y su impacto en el tema que nos ocupa, es posible entrar en el último aspecto: la valoración de si nuestra regulación es adecuada o no lo es. La idea es no centrarse solo, ni tanto, en el posible impacto en ella del Reglamento de Inteligencia Artificial –que ha sido sucintamente descrita en el apartado anterior–. Pretendo más bien verificar si, del conjunto de normas examinadas, se desprende una razonable salvaguarda de los intereses de los posibles sujetos afectados. Parto, al respecto, de que la inteligencia artificial es una nueva realidad a cuya utilización debemos prestar atención, incluso cuando se utiliza por el bien común. Mi impresión general es, sin embargo, que los laboristas hemos volcado nuestro interés fundamentalmente en el uso por las empresas de la inteligencia artificial. Aunque seguramente ello es razonable habida cuenta el objeto y los principios esenciales de nuestra disciplina, no podemos quedarnos solo en eso: hemos de fijarnos también en el uso de la IA por los organismos públicos. Si se acepta este punto de partida, se hace necesaria una revisión crítica de la situación a la que conduce el entramado de las normas europeas e internas en relación con la utilización por las entidades públicas de la inteligencia artificial que, a mi juicio, pasa por tres órdenes de consideraciones.

La primera es una crítica a la fragmentación que se desprende de la normativa aplicable. Acabamos de ver como el Reglamento de Inteligencia Artificial obliga a parcelar el tratamiento de la incorporación de los sistemas de IA a las decisiones públicas. Pero esto no parece razonable no solo por lo que está en juego desde la perspectiva ciudadana sino también por la dificultad de introducir particiones que, en rigor, distan de tener un sentido claro cuando estamos considerando el funcionamiento de instituciones públicas. Las normas del RIA, como anteriormente las del RGPD, piensan fundamentalmente en

la protección de las personas físicas; y por ello, cuando se proyectan sobre el problema que nos ocupa, dejan un amplio margen de actuaciones en el que no son aplicables. No es fácil, sin embargo, aceptar que la misma entidad quede sujeta a determinadas obligaciones solo cuando están aquellas implicadas y no lo esté en caso de que la afectación se refiera a personas jurídicas. Las entidades públicas son únicas y parece razonable que el régimen de sus actuaciones debe ser uniforme, con independencia de la naturaleza de los sujetos que queden afectados por ellas. Esto se advierte de forma bastante clara en el terreno de la utilización de la IA en la justicia, que siempre es calificada como de alto riesgo por el Reglamento, sin fijar la atención en los afectados por la actuación judicial como hemos visto en el apartado anterior. Si esta diversidad es criticable, se impone intentar una reconstrucción del tratamiento del uso de la IA que supere la diversificación normativa y que lo sujete a un marco de garantías unitario.

Esto es relativamente sencillo, en segundo lugar, en relación con algunos aspectos de las garantías a las que hemos hecho referencia. Es el caso del principio de la supervisión humana de la IA y sus derivadas, como el “derecho a de explicación de decisiones tomadas individualmente”. Aunque el RIA los establezca solo para los sistemas de alto riesgo (cfr. arts. 14 y 86, respectivamente), lo cierto es que pueden deducirse derechos análogos de otras fuentes, con independencia de que nos encontremos o no en ante el funcionamiento de uno de ellos. Es verdad que las reglas sobre decisiones automatizadas del art. 22 RGPD, aunque son aplicables con independencia de que la IA que las genere sea o no de alto riesgo, tienen un ámbito de aplicación limitada en la medida en que solo las personas físicas son titulares del derecho de protección de datos (cfr. art. 1.1). Sin embargo, a través de reglas internas del máximo nivel, habrá que entender que se alcanzan resultados similares a los que derivan de una normativa europea que, en principio, no resultaría aplicable. Sin entrar en el debate general sobre si las personas jurídicas son o no titulares de los derechos fundamentales reconocidos por nuestra Constitución, parece que esta titularidad no puede discutirse de aquellos de ellos de carácter más formal o relacional, entre los que sin duda se incluyen los contenidos en el art. 24 CE. Asimismo, las exigencias constitucionales respecto al funcionamiento de los órganos administrativos (cfr. art. 105) y jurisdiccionales (art. 120) imponen específicas características a los procedimientos que conducen a sus decisiones que no pueden sino proyectarse en la configuración infraconstitucional de aquellos que se apoyen en el uso de inteligencia artificial imponiendo la supervisión humana, antes y/o después de la adopción de la decisión. De este modo, la regulación legal vigente conduce a la aplicación, por vía de las reglas sobre audiencia, motivación o recurso, de criterios de análoga significación a los derivados del principio de supervisión humana del RIA o a los establecidos al respecto en el art. 22.3 RGPD.

En esta línea, cabe observar que la incorporación de la inteligencia artificial a decisiones judiciales o administrativas tiende a no sustituir la intervención humana. En la regulación actual, esto es muy claro para las “actuaciones asistidas” en el terreno jurisdiccional que, como hemos visto, siempre quedan sujetas a la “validación” del órgano judicial (art. 57 RDL 6/2023). Pero también lo es para las actuaciones administrativas que aparecen como más polémicas desde el punto de vista de los derechos fundamentales: las

actas automatizadas de la Inspección. Al margen de que, conforme a las reglas generales, las actas implican únicamente la iniciación del procedimiento sancionador, las reglas reglamentarias aplicables imponen la intervención humana del personal de la inspección en caso de que el interesado alegue “hechos o circunstancias distintos a los consignados en el acta, insuficiencia del relato fáctico de dicho acta, o indefensión por cualquier causa” (cfr. art. 47.3 Reglamento de procedimiento sancionador en el orden social, en su redacción por RD 688/2021). Los usos predictivos de la IA, que coadyuvan a la planificación de las actuaciones inspectoras, no plantean, por lo demás, problemas relevantes desde la perspectiva del principio de supervisión humana pues se limitan a fijar el blanco para unas posteriores actuaciones inspectoras se desarrollarán por personas físicas (Todolí Signes, 2020).

Es posible, sin embargo, que ello no sea suficiente. Con carácter general se ha señalado, a propósito de los sesgos discriminatorios, la aparición de discriminaciones de nuevo cuño que, además, son mucho más sutiles que las tradicionales, por las dificultades para detectarlas y por las capacidades de justificación al alcance de la inteligencia artificial. No hay que descartar, por otro lado, el impacto del temor reverencial o la admiración que puede suscitar del funcionamiento de la IA. Es el llamado sesgo de automatización “por el que, en general, los humanos tendemos a confiar en lo que nos dicen los sistemas (de IA) empleados y rara vez lo cuestionamos” (Ponce Solé, 2024, p. 177). En este contexto, acaso la intervención humana resulte por sí sola insuficiente si no va acompañada de específicas acciones de entrenamiento para mejorar; y, sobre todo, resulta necesario completar su acción mediante el establecimiento de nuevos mecanismos de control de las decisiones públicas, quizá estableciendo por ejemplo una presunción de falibilidad (Tahirí Moreno, 2023), o, en todo caso, mejorando los de carácter previo a la incorporación de la tecnología a la actividad administrativa y jurisdiccional garantizando la fiabilidad de los datos que se utilizan y haciendo transparentes las características de los procesos que desarrollan.

Diría que es este, en tercer lugar, el terreno en el que se detectan los mayores déficits, como se ha puesto de manifiesto de forma insistente (Goerlich Peset, 2021, p. 37 ss., con referencias adicionales; con posterioridad, por ejemplo, Jiménez-Castellanos Ballesteros, 2023). Pero no quedan solventados por la entrada en vigor del Reglamento. Es verdad que su art. 26.8 RIA impone también obligaciones de registro a “autoridades públicas o instituciones, órganos y organismos de la Unión” que sean “responsables del despliegue de sistemas de IA de alto riesgo” y les impide desplegar aquellos que no hayan sido registrados previamente por su proveedor o distribuidor. El alcance de la obligación de registro se concreta en los arts. 49 y 71 y supone que, con algunas salvedades que resultan de poco interés en nuestra materia, una información relativamente amplia del sistema (anexo VIII RIA) queda a disposición del público “de manera sencilla” (art. 71.4).

Estas reglas completan las establecidas por el ordenamiento interno, cuando el sistema en cuestión tiene la condición de alto riesgo. Así ocurre en el marco de las actuaciones judiciales, en las que el marco formal establecido por el art. 58 RDL 6/2023 se completa con las previsiones del RIA. De hecho, este precepto anticipa las previsiones europeas, puesto que los sistemas afectados no son solo los que presuponen decisiones automatizadas: las actuaciones proactivas en las que la decisión final es humana están tam-

bién incluidas en “la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, la auditoría del sistema de información y de su código fuente” (apartado 1) y se prevé la publicidad y objetividad de “los criterios de decisión, dejando constancia de las decisiones tomadas en cada momento” (apartado 2).

En el ámbito administrativo, sin embargo, las cosas son más complicadas, habida cuenta que, en su mayor parte, las actuaciones algorítmicas quedan, como sabemos, extramuros de la consideración de sistemas de IA de alto riesgo, con la única salvedad de su uso en la dispensación de servicios o prestaciones vinculados a la protección social. En este contexto, disponemos únicamente de las reglas generales (art. 41 Ley 40/2015) o específicas (art. 130.I LGSS), cuyas insuficiencias son notables por dos razones. La primera se relaciona con la estricta noción de decisión automatizada, que incluye únicamente la “adopción de resoluciones” en el caso de la gestión de la Seguridad Social (art. 130.I LGSS) o, si se amplía la óptica hasta incluir “cualquier acto o actuación realizada íntegramente a través de medios electrónicos” como hace el art. 41.1 40/2015, se requiere “que no haya intervenido de forma directa un empleado público”. De este modo, las actuaciones prospectivas de las entidades administrativas quedan fuera de las obligaciones de transparencia legalmente establecidas. Estas, en segundo lugar, son muy limitadas. Literalmente se concretan en la determinación previa del “órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente” y en la información sobre “el órgano que debe ser considerado responsable a efectos de impugnación” (arts. 41.2 Ley 40/2015 y 130.II LGSS). La transparencia es, pues, estrictamente «orgánica» –quién determina las características y ante quién se pueden recurrir las actuaciones automatizadas– y en modo alguno tiene componentes «funcionales» –relacionadas con la dinámica del sistema algorítmico–.

Seguramente este estado de cosas debe ser revisado. Textos posteriores adoptan, en este sentido, una perspectiva más amplia. El art. 23 Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, se sitúa en ella al referirse de forma genérica a “los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas”, expresión notablemente más amplia que las empleadas en las normas anteriores. Asimismo, hace específicas alusiones a la transparencia. Más allá de su consideración en el art. 23.1, se prevé que las administraciones públicas “priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos”. Poco antes, la Carta de Derechos Digitales, aprobada por el Gobierno en julio de 2021, había insistido en esta línea. En su apartado XVII.6, en el marco de la promoción de “los derechos de la ciudadanía en relación con la inteligencia artificial... en el marco de la actuación administrativa” se reconocieron, entre otros, los derechos a “que las decisiones y actividades en el entorno digital respeten los principios de buen gobierno y el derecho a una buena Administración digital, así como los principios éticos que guían el diseño y los usos de la inteligencia artificial” (letra a) y, sobre todo, a “la transparencia sobre el uso de instrumentos de inteligencia artificial y sobre su funcionamiento y alcance en cada procedimiento concreto y, en particular,

acerca de los datos utilizados, su margen de error, su ámbito de aplicación y su carácter decisorio o no decisorio” (letra b)).

Sin embargo, desde una perspectiva práctica estos textos tienen una virtualidad limitada. La Carta de Derechos Digitales “no tiene carácter normativo”. Así se reconoce de forma taxativa en sus “consideraciones previas”. Por ello, y en relación con los derechos de transparencia aludidos se prevé que “la ley podrá regular las condiciones de transparencia y el acceso al código fuente, especialmente con objeto de verificar que no produce resultados discriminatorios”; y en este nivel jerárquico estamos donde estábamos. Por su parte, el citado art. 23 Ley 15/2022, como muchos de los preceptos de esta norma, tienen un claro componente programático que dificulta extraer concretas prescripciones de él. Por lo demás, una y otro se centran en el terreno de los derechos fundamentales lo que incide en la discutible fragmentación del tratamiento del uso de la IA por las Administraciones, en función de que sean afectadas personas físicas o jurídicas.

A la espera de avances normativos específicos y más concretos, acaso sea posible superar este estado de cosas mediante el recurso a la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. El argumento que suministra esta norma en relación con el problema que nos ocupa se encuentra en su art. 7, que en el marco de la “transparencia de la actividad pública” regula la “información de relevancia jurídica”. En concreto, obliga a las Administraciones Públicas, en el ámbito de sus competencias, a publicar “las directrices, instrucciones, acuerdos, circulares o respuestas a consultas planteadas por los particulares u otros órganos en la medida en que supongan una interpretación del Derecho o tengan efectos jurídicos”.

Existen ya supuestos de aplicación por el Consejo de Transparencia y, en último término, por los órganos judiciales, de esta regla al ámbito que nos interesa. Por lo que se refiere a los primeros, los ha estudiado recientemente Amparo García Rubio (2024, Parte 3.2). A su juicio, las resoluciones del indicado Consejo, que han estimado las reclamaciones frente a las solicitudes de información frente a las entidades gestoras o la ITSS, han de llevar a la obvia conclusión de que “entre la nada y el todo existen niveles intermedios de información que, teniendo en cuenta las particularidades y fines de la actuación inspectora, seguramente proporcionarán ese punto de equilibrio a lograr entre la eficacia administrativa y la transparencia algorítmica”. Cabe indicar que, en otros ámbitos, estas cuestiones han desembocado ya en procesos judiciales en los que se ha buscado ir más allá de la posición del Consejo. Aparte la STSJ Galicia cont. 89/2024, 22 marzo, en relación con el algoritmo de un programa de citación de usuarios en el sistema público sanitario, es necesario traer a colación las resoluciones recaídas en el llamado caso Bosco, en el que, a instancias de una fundación privada, se analiza la información que debe ser suministrada en relación con el algoritmo que controla la asignación del bono social. Sin entrar en detalles sobre las distintas resoluciones recaídas (Res. Consejo Transparencia R/0701/2018 y sentencias JC-8 cont. 143/2021, 30 diciembre y AN cont. 30 abril 2024, rec. 51/2022), este episodio abrirá probablemente la posibilidad de que el Tribunal Supremo se pronuncie sobre su alcance y, en concreto, sobre las posibilidades de acceder al conocimiento del código-fuente. Con toda probabilidad, y con independencia

de la solución que se dé a esta cuestión, parece necesaria una intervención normativa ulterior, que implique la adecuación por parte de las Administraciones a la doctrina que va siendo emanada del Consejo de Transparencia y haga efectivos los principios que se desprenden de las citadas fuentes de *soft law*.

Bibliografía citada

- Aibar Bernard, J. (2020). El Big Data y el análisis de datos aplicados por la Tesorería General de la Seguridad Social como medio de lucha contra el fraude en la Seguridad Social. *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 11.
- AISS. (2019). *10 desafíos mundiales para la Seguridad Social. Evolución e innovación*. <https://www.issa.int/sites/default/files/documents/publications/3-10-challenges-Global-2019-WEB-263632.pdf>.
- García Rubio, M. A. (2024). Retos actuales de la Inspección de Trabajo y Seguridad Social desde la perspectiva de los derechos de los administrados. *Documentación-Laboral*, II(132).
- Goerlich Peset, J. M. (2021). Decisiones administrativas automatizadas en materia social: algoritmos en la gestión de la Seguridad Social y en el procedimiento sancionador. *Labos: Revista de Derecho del Trabajo y Protección Social*, 2(2), 22-42.
- Jiménez-Castellanos Ballesteros, I. (2023). Decisiones automatizadas y transparencia administrativa: Nuevos retos para los derechos fundamentales. *Revista española de la transparencia*, 16, 191-215.
- Nores Torres, L. E. (2023). El proceso de digitalización en la jurisdicción social: algunos avances y perspectivas. *Lex social: revista de los derechos sociales*, 13(2).
- Nores Torres, L. E. (2024a). Digitalización, inteligencia artificial y proceso laboral. *Labos: Revista de Derecho del Trabajo y Protección Social*, 5(1), 222-246.
- Nores Torres, L. E. (2024b). *Proceso laboral y digitalización tras el RDL 6/2023, de 19 de diciembre*. Tirant lo Blanch.
- Ponce Solé, J. (2024). Inteligencia Artificial, decisiones administrativas discrecionales totalmente automatizadas y alcance del control judicial: ¿indiferencia, insuficiencia o deferencia? *Revista de Derecho Público: teoría y método*, 9, 171-220.
- Redondo Rincón, G. (2020). Las nuevas tecnologías en el control de la Incapacidad Temporal: la aplicación de la analítica predictiva. *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 11.
- Tahirí Moreno, J. (2023). El principio de presunción de falibilidad de las decisiones algorítmicas desfavorables: Una nueva garantía jurídica frente a las decisiones automatizadas y el uso de sistemas de inteligencia artificial en la administración pública. *Revista Aragonesa de Administración Pública*, 60, 188-214.
- Todoí Signes, A. (2020). Retos legais do uso do big data na selección de suxeitos a investigar pola Inspección de Traballo e da Seguridade Social. *REGAP: Revista galega de administración pública*, 1(59), 79-102.