

Big data e inteligencia artificial en la Administración tributaria

Rafael Oliver Cuello

ESERP Business & Law School

Fecha de presentación: enero de 2021

Fecha de aceptación: febrero de 2021

Fecha de publicación: octubre de 2021

Resumen

El empleo de los macrodatos o *big data* por parte de las Administraciones tributarias, junto con el uso de las herramientas de inteligencia artificial, tiene una enorme relevancia en la actual aplicación de los tributos, con indudables consecuencias jurídicas tanto para los derechos y garantías de los obligados tributarios como para la actuación de la Hacienda pública. Las Administraciones tributarias emplean el *big data* y la inteligencia artificial, entre otras cosas, en la selección de contribuyentes para iniciar un procedimiento de inspección tributaria, en los modelos predictivos del riesgo de incumplimiento del obligado tributario, en los asistentes virtuales, en la confección de borradores de declaraciones tributarias, en la actuación administrativa automatizada y en los actos resolutorios automatizados. La utilización de estas herramientas tecnológicas debe caracterizarse por un cumplimiento estricto de los principios jurídicos que rigen la aplicación de los tributos y debe ser respetuosa con los derechos y garantías del contribuyente.

Palabras clave

Big data, inteligencia artificial, Administración tributaria, Hacienda pública, impuestos

Big data and artificial intelligence in tax administration

Abstract

The use of big data by tax administrations, together with that of artificial intelligence tools, has an enormous importance in the current application of taxes, with clear legal consequences both for the rights and guarantees of taxed parties and for public finance practice. Tax administrations employ big data and artificial intelligence, in addition to other cases, in the selection of taxpayers in order to introduce a tax inspection procedure, predictive risk models for taxpayer non-payment, virtual assistants, the drawing up of tax returns, automatic administrative practice and automatic administrative resolutions. The use of these technological tools must be characterised by a strict observance of the legal principles which regulate the application of the taxes, and must respect the rights and guarantees of the taxpayer.

Keywords

Big data, artificial intelligence, tax administration, public finance, taxation

1. Aproximación al *big data* y la inteligencia artificial

En su acepción literal, el concepto de *big data*, como indica la doctrina, hace referencia, en primer término, a la cantidad de datos. No obstante, los macrodatos no se definen únicamente por su volumen, sino por la complejidad aparejada a este crecimiento que conlleva la dificultad de almacenarlos, procesarlos y analizarlos¹. En esta línea, el *big data* se refiere no solo al enorme volumen de los datos acumulados en el tráfico diario de internet, sino a otras características vinculadas a las tecnologías de la información y la comunicación, como son la velocidad, la constante actualización y la variedad de datos, ya sean estructurados, semiestructurados o desestructurados.

En este sentido, es importante destacar que la ingente cantidad de datos que procesan las corporaciones tecnológicas, las empresas privadas y las Administraciones públicas conlleva el abandono de las técnicas de procesamiento de datos anteriores al uso extendido de la red.

En el ámbito tributario, especialmente en la Administración tributaria estatal, desde hace bastantes años se cuenta con gran cantidad de información. Probablemente ello se deba a que, desde la reforma tributaria de 1978, se han establecido en nuestro ordenamiento jurídico unas potentes obligaciones de información, tal como se recoge en el art. 93 de la Ley 58/2003, de 17 de diciembre, General Tributaria (en adelante, LGT), que han facilitado a la Administración tributaria unas eficaces herramientas de obtención de información, ya sea por captación o por suministro.

A todo este enorme conjunto de datos e informaciones obtenidos por la Administración tributaria a través de estas fuentes, hay que añadir otras informaciones o datos facilitados por vía de convenio o por medio del acceso a las bases de datos de otras instituciones públicas, tales como el Catastro, la Dirección General de Tráfico, los organismos

gestores de la Seguridad Social o la Inspección de Trabajo, entre otras.

Este enorme volumen de información disponible por la Administración tributaria es, pues, muy anterior a la irrupción del *big data* y la inteligencia artificial. Incluso es anterior a la aparición de las tecnologías de la información y la comunicación, tal como las conocemos actualmente. El uso intensivo de las herramientas informáticas y, posteriormente, de las telemáticas, durante los años noventa, ha provocado que esa gran cantidad de información haya podido ser tratada de forma eficiente por parte de la Administración tributaria.

Pues bien, en los últimos años, se ha incrementado aún más la capacidad de recopilación de datos derivada del uso de nuevas fuentes y herramientas de obtención de información. Así, las Administraciones tributarias cuentan en su poder con una enorme cantidad de datos provenientes tanto de fuentes internas, ya sean de carácter tradicional (como las que hemos comentado) o de nuevas fuentes de carácter digital, así como datos transferidos de otros Estados, que se encuentran almacenados en distintos formatos estructurados y sin estructurar. Como apunta la doctrina², esta ingente cantidad de datos plantea una serie de retos de carácter técnico relacionados, por un lado, con la disparidad de formatos y fuentes y, por otro lado, con su almacenamiento.

Muy relacionado con este tema se encuentra la denominada «inteligencia artificial», cuyo objetivo es desarrollar técnicas que permitan que los ordenadores aprendan (*machine learning* o aprendizaje automático), es decir, que su desempeño mejore con la experiencia. Este método emplea algoritmos para convertir muestras de datos en programas de ordenador que permiten generalizar comportamientos e inferencias para un conjunto de datos.

Como apunta la doctrina³, durante los últimos años, el concepto de *machine learning* o aprendizaje automático ha surgido parejamente al auge de los datos a gran escala. Mediante este término se hace referencia a la programa-

1. Véase al respecto Dorado Ferrer (2021, págs. 6-8). Comenta este autor cómo diariamente Google procesa miles de millones de búsquedas; se crean cientos de miles de sitios webs nuevos; se envían centenares de millones de tuits; se visualizan miles de millones de vídeos en YouTube; se suben decenas de millones de fotos en Instagram; y se realizan centenares de millones de llamadas de Skype. Vistos estos datos en conjunto, el tráfico diario de internet es de miles de millones de gigabytes.
2. Véase Serrano Antón (2020, págs. 41-43).
3. Véase Dorado Ferrer (2021, págs. 24-25).

ción de inteligencia artificial que no solo aplica la información para predecir el comportamiento de los datos, sino que es capaz de aprender por sí misma.

Además, como una de las notas características que la definen en determinados supuestos, este tipo de programación algorítmica utiliza modelos de «caja negra» (*black boxes* u otros conceptos asociados, como *deep learning* o *ensemble learning*). Por «caja negra» debemos entender aquella programación que, por su complejidad, utiliza modelos cuyos resultados son prácticamente indescifrables, incluso para el propio programador del algoritmo. Esto puede suponer un problema, especialmente en el ámbito de su uso en la Administración pública, ya que puede entrar en colisión directa con el principio de transparencia⁴.

2. Uso del *big data* y la inteligencia artificial por la Administración tributaria

Uno de los principales usos del *big data* y la inteligencia artificial por las Administraciones tributarias se refiere a la selección de contribuyentes para iniciar un procedimiento de inspección tributaria. Efectivamente, estas tecnologías y herramientas están siendo empleadas como métodos predictivos de análisis de datos para determinar el riesgo fiscal de los contribuyentes. La Agencia Estatal de Administración Tributaria (en adelante, AEAT) viene desarrollando una herramienta denominada Hermes para la gestión de riesgos, que se apoya en la información existente en el sistema Zújar⁵.

A este respecto, hay que subrayar que el establecimiento de una organización y procedimientos para asegurar la gobernanza de la información ha sido durante años el pilar básico en el que se ha asentado el análisis de información en la AEAT. Actualmente, hacen uso de la herramienta de

análisis y visualización de datos corporativa de la Agencia, denominada Zújar, más de quince mil de los aproximadamente veinticuatro mil empleados con que cuenta. Esta herramienta es, además, la fuente de datos para cualquier análisis avanzado de información⁶.

La selección de contribuyentes lleva ya, de hecho, más de veinticinco años alimentándose de modelos analíticos, aunque el salto de las hojas de cálculo a los algoritmos ha revolucionado esta práctica⁷. Un ejemplo de análisis de riesgos apoyado en la inteligencia artificial es Hermes, una herramienta desarrollada por la AEAT para implantar un sistema único para el análisis de riesgos y la asignación de modalidades de intervención.

El sistema aprovecha la riqueza de los datos existentes en las bases de la AEAT para producir informes de riesgo estandarizados, que se usan en los procesos de selección de contribuyentes para iniciar un procedimiento de inspección. La herramienta Hermes también optimiza el uso de nuevas fuentes internacionales de información (incluido el intercambio automático de información financiera, la presentación de informes país por país y el intercambio de resoluciones o *rulings*), porque tiene la flexibilidad de adaptarse a los riesgos emergentes y a la incorporación de nuevos conjuntos de datos.

Las Administraciones tributarias también utilizan el *big data* y la inteligencia artificial con modelos predictivos del riesgo de incumplimiento del obligado tributario. Como apunta la doctrina, con las nuevas técnicas de almacenamiento, tratamiento y análisis de datos que permiten el acceso a la información de forma inmediata o casi inmediata, las Administraciones tributarias están en disposición de conocer la realización del hecho imponible en tiempo real y adelantar el momento de liquidación del impuesto o de las actuaciones de control. Así, las tradicionales actuaciones de control efectuadas con posterioridad a la realización de las operaciones gravadas pueden ser sustituidas por

4. Véase García-Torres Fernández (2021, págs. 459-476).

5. Véase Serrano Antón (2020, pág. 44).

6. La integración organizativa de la función de análisis de información en el Departamento de Informática Tributaria ha facilitado la integración de la seguridad del sistema Zújar y otros sistemas de análisis con la de los sistemas operacionales, asegurando que las infraestructuras técnicas y organizativas dedicadas a la seguridad pueden tener una visión holística de los accesos a la información. Véase Borja Tomé (2020, págs. 211-212).

7. El 53% de las Administraciones tributarias de la OCDE utilizan modelos predictivos que anticipan posibles problemas para que los servicios adopten las medidas correctoras necesarias. Véase González de Frutos (2020, pág. 147).

comprobaciones simultáneas y por predicciones *ex ante* del incumplimiento del obligado tributario⁸.

Actualmente, el ejemplo más relevante en este ámbito viene constituido por el Suministro Inmediato de Información (en adelante, SII). El SII comporta la llevanza de los libros registro (los libros registro de facturas expedidas, de facturas recibidas, de bienes de inversión y de determinadas operaciones intracomunitarias) a través de la sede electrónica de la AEAT, mediante el suministro prácticamente inmediato de los registros de facturación. Los contribuyentes deben remitir a la Administración tributaria de forma continuada los detalles sobre la facturación por vía electrónica, y con esta información actualizada se van configurando, casi en tiempo real, los libros registro del tributo⁹.

Y no hay que olvidar que la Administración cuenta en su poder con una cuantiosa información sobre la actividad económica de los contribuyentes a los que se aplica el SII, una información superior a la que figura en los libros registro que deben llevar el resto de empresarios y profesionales. Mediante el empleo de modelos predictivos, toda esta ingente y valiosa información puede permitir a la Administración tributaria gestionar el riesgo de un incumplimiento del obligado tributario.

Otro de los usos del *big data* y la inteligencia artificial por parte de la Administración tributaria tiene que ver con la asistencia al contribuyente. Coincidimos con la doctrina que afirma que las Administraciones tributarias tienen el deber de utilizar el *big data* y la inteligencia artificial no solo para controlar mejor a los contribuyentes y luchar contra el fraude, sino también al servicio de los mismos, para mejorar la asistencia y, en particular, para ofrecerles servicios más personalizados que faciliten el cumplimiento voluntario¹⁰.

Los asistentes virtuales, mediante técnicas de inteligencia artificial, permiten entender la expresión que utiliza un usuario para preguntar una cuestión o, también, para realizar un trámite. Además, permiten requerir, vía conversa-

ción, las precisiones que sean necesarias para responder a la pregunta u ofrecer el servicio solicitado. El conocimiento en el que se basa un asistente virtual, para comprender lo que se solicita, consta de una serie de intenciones (cuestiones o servicios) y una serie de expresiones para cada intención. La introducción de estas expresiones es el denominado «entrenamiento del asistente virtual»¹¹.

Así, si un usuario utiliza alguna de estas expresiones, el asistente virtual podrá deducir qué cuestión o servicio está solicitando. Pero, además, será capaz de entender otras expresiones que se parezcan a las que se han entrenado. Este proceso de ampliación de la capacidad de comprensión se denomina «expansión».

Algún autor considera acertadamente que los asistentes virtuales mejoran la calidad de la información y permiten a la Administración tributaria ahorrar, al no tener que contratar personal eventual para las campañas de presentación de declaraciones. Se argumenta, en este sentido, que, a diferencia de los servicios de información tradicionales, que ofrecen la información en repositorios estáticos (preguntas frecuentes o notas divulgativas), los asistentes virtuales son dinámicos y, dado que son capaces de interpretar el lenguaje natural y comprender la pregunta del contribuyente, ofrecen la respuesta pertinente de forma rápida y segura¹².

Un ejemplo de asistente virtual es el que utiliza actualmente la AEAT para dar soporte a los contribuyentes incluidos en el SII. Mediante esta herramienta, se pueden resolver de forma automatizada las dudas técnicas de los sujetos pasivos durante las veinticuatro horas del día, reduciendo los efectivos dedicados a su resolución y mejorando la productividad y eficacia tanto de los contribuyentes como de la propia Agencia.

Como apunta con acierto la doctrina, nos encontramos ante una evolución del programa Informa, cuyo contenido pasaría a integrar las preguntas y respuestas a las que atiende el asistente virtual. A semejanza del programa Informa, cuando el obligado tributario haya ajustado sus

8. Véase Serrano Antón (2020, pág. 45).

9. Véase Delgado García y Oliver Cuello (2017, págs. 101-124).

10. Véase García-Herrera Blanco (2020, pág. 303).

11. Véase Segarra Tormo (2020, pág. 190).

12. Véase González de Frutos (2020), pág. 149.

actuaciones a las respuestas emitidas por el asistente virtual, no incurrirá en responsabilidad por infracción, de acuerdo con lo previsto por el art. 179.2 LGT. Atendiendo a posibles preguntas complejas, el asistente virtual siempre puede remitir a un funcionario de la Agencia para su atención presencial o en línea, o bien puede derivar a la formulación de una consulta tributaria u otro tipo de asistencia al contribuyente¹³.

También en el terreno de la asistencia al contribuyente, el *big data* y la inteligencia artificial desempeñan un papel clave en la confección de borradores de declaraciones tributarias. Efectivamente, en la actualidad, en el caso del impuesto sobre la renta de las personas físicas (en adelante, IRPF), tres cuartas partes de los borradores se convierten en declaraciones de una forma muy sencilla y rápida, lo cual hace concebir fundadas esperanzas de que, en un futuro próximo, los borradores tributarios puedan extenderse a los demás impuestos y ser directamente elaborados por las máquinas de proceso de información de la Administración tributaria¹⁴.

Compartimos la idea expresada por la doctrina de que debe superarse la concepción de la información obtenida por suministro como mero elemento de contraste de las autoliquidaciones presentadas, de forma que dicha información debe servir para integrar los borradores de autoliquidación que faciliten a los obligados tributarios el cumplimiento de sus obligaciones. En este sentido, el borrador del IRPF ha demostrado que reduce errores, limita los incumplimientos y permite que los recursos se dediquen al control del fraude fiscal¹⁵.

En esta línea, es un objetivo del SII la potenciación de la asistencia al contribuyente, al facilitarle una serie de datos fiscales, en una primera fase, y el borrador de la declaración del impuesto, en una segunda fase. A este respecto, el contribuyente dispondrá en la sede electrónica de la AEAT de un libro registro «declarado» y otro «contrastado» con la información de contraste procedente de terceros que pertenezcan al colectivo de este sistema o de la base de datos de la AEAT. Los contribuyentes podrán contrastar

dicha información antes de la finalización del plazo de presentación de su declaración mensual de impuesto sobre el valor añadido. Y el contribuyente tendrá la posibilidad de corregir los errores cometidos en los envíos sin necesidad de ser requerido por la AEAT para ello.

Otro de los usos del *big data* y la inteligencia artificial por las Administraciones tributarias es la actuación administrativa automatizada. Efectivamente, varias Administraciones tributarias están implementando rutinas basadas en reglas para tratar los riesgos predefinidos y así, por ejemplo, denegar automáticamente una devolución, emitir una liquidación o asignar información dudosa a un contribuyente en concreto (*matching*). Estas actuaciones ocurren en tiempo inmediato o casi inmediato y están sustituyendo a actuaciones previamente realizadas por funcionarios, logrando gestionar la información de manera masiva, temporánea y segura, evitando el error humano, de forma que en muchos aspectos suponen actuaciones más efectivas y eficientes que la tradicional comprobación tributaria presencial¹⁶.

En este contexto de implantación y desarrollo de la Administración electrónica, adquieren suma importancia los procesos de toma de decisiones en los que interviene de una forma trascendental la informática. Es lo que se denomina por algunos autores como «informática decisional», o, lo que es lo mismo, la sustitución de la inteligencia humana por la inteligencia artificial (una aplicación informática) en la toma de decisiones administrativas, y que se regula, por primera vez en un texto legal de estas características, en la LGT. Fenómeno que ya viene siendo aplicado desde hace tiempo en otros ámbitos privados, como, por ejemplo, en las operaciones bursátiles de transmisión de valores en los mercados secundarios¹⁷.

Por lo tanto, aquí no se trata simplemente de tramitar un procedimiento administrativo por medios telemáticos, sino que se da un paso más allá, ya que las posibilidades de las tecnologías de la información y la comunicación son amplias, permitiendo que la resolución de dicho procedimiento sea adoptada por una aplicación informá-

13. Véase Bilbao Estrada (2019, págs. 134-135).

14. Véase Hurtado Puerta (2020, pág. 178).

15. Véase Delgado García (2020, págs. 214-223).

16. Véase González de Frutos (2020, págs. 147-148).

17. Véase Delgado García y Oliver Cuello (2007, págs. 1-12).

tica. Lo cual merece un juicio positivo, ya que se han de aprovechar todas las posibilidades que ofrecen este tipo de tecnologías y, si puede tramitarse de forma prácticamente total un procedimiento por medios telemáticos, no debe haber inconveniente en admitir que la resolución del mismo la efectúe una aplicación informática. De esta manera, la celeridad en la toma de decisiones es considerable, por lo que, en nuestra opinión, es previsible que en un futuro inmediato aumenten los supuestos de aplicación de la informática decisional en la Administración tributaria. Ahora bien, es importante establecer las garantías y los límites que han de rodear a este tipo de actuaciones.

El ámbito tributario ha sido pionero en el reconocimiento legal de la informática decisional. En efecto, el art. 100.2 LGT declara que «tendrá la consideración de resolución la contestación efectuada de forma automatizada por la Administración tributaria en aquellos procedimientos en que esté prevista esta forma de terminación».

A nuestro juicio, es positivo que en la LGT se realice una mención específica a la informática decisional. Este art. 100.2 LGT debe ponerse en relación con el art. 96.3 de la propia LGT, que dispone que «los procedimientos y actuaciones en los que se utilicen técnicas y medios electrónicos, informáticos y telemáticos garantizarán la identificación de la Administración tributaria actuante y el ejercicio de su competencia. Además, cuando la Administración tributaria actúe de forma automatizada se garantizará la identificación de los órganos competentes para la programación y supervisión del sistema de información y de los órganos competentes para resolver los recursos que puedan interponerse».

Por lo tanto, a nuestro juicio, de acuerdo con lo previsto en los arts. 96.3 y 100.2 LGT, conviene delimitar claramente cuándo estamos ante «actuaciones automatizadas» (a las que la Ley denomina, en el primer inciso del art. 96.3 LGT, «procedimientos y actuaciones en los que se utilicen técnicas y medios electrónicos, informáticos y telemáticos») y cuándo nos hallamos ante «actos resolutorios automatizados» (a los que la Ley se refiere, en el segundo inciso del art. 96.3 LGT, aludiendo a que «la Administración tributaria actúe de forma automatizada», y a los que el art. 100.2 LGT se refiere como «contestación efectuada de forma automatizada»).

Como hemos visto, las consecuencias jurídicas de la ubicación de un determinado acto o actuación administrativa en una categoría u otra son diferentes. En el primer caso de las actuaciones automatizadas, el art. 96.3 LGT dispone que se garantizará únicamente la identificación de la Administración tributaria actuante y el ejercicio de su competencia. Mientras que en el segundo caso de los actos resolutorios automatizados, se establece una garantía adicional en el art. 96.3 LGT consistente en que se identificarán, además, los órganos competentes para la programación y supervisión del sistema de información y de los órganos competentes para resolver los recursos que puedan interponerse. Asimismo, en este último caso de los actos resolutorios automatizados, el art. 100.2 LGT prevé que dicha terminación del procedimiento de forma automatizada debe estar expresamente prevista por la Ley.

En nuestra opinión, dentro del primer grupo de actuaciones automatizadas, cabe incluir, por ejemplo, las actuaciones de comprobación llevadas a cabo en el seno de un procedimiento de verificación de datos, mediante el cruce de datos automatizado; el recibo de la presentación telemática de cualquier declaración tributaria o documento presentado ante un registro electrónico; la identificación de la Administración cuando un ciudadano solicita una conexión segura con la misma, identificándose a través de un certificado de servidor; la identificación telemática ante las entidades de crédito con ocasión de pago de deudas tributarias; la generación automatizada del Número de Referencia Completo, que es un código que se incorpora automáticamente a los distintos documentos, fundamentalmente como justificante del pago electrónico, pero también, por ejemplo, para validar avales o para la constitución de depósitos para participar en subastas de internet; y la programación anticipada de las pujas de los interesados en las subastas realizadas por vía telemática, de forma que tales interesados no tienen por qué intervenir precisamente en el momento de la subasta, obedeciendo el sistema de información las órdenes encomendadas.

Por otra parte, el segundo grupo de actos resolutorios automatizados cuenta, en la actualidad, con menos manifestaciones que en el caso anterior. Entre ellas, pueden citarse los siguientes ejemplos: el levante aduanero en el caso de mercancías clasificadas en el canal verde, algunos supuestos de diligencias de embargos de cuentas bancarias o bien la resolución de los aplazamientos de menor cuantía.

A nuestro juicio, un buen ejemplo de informática decisional se puede encontrar en la introducción de filtros o parámetros de detección de discrepancias o irregularidades en la base de datos consolidada de la AEAT. Dichos filtros o parámetros detectan, por ejemplo, un error aritmético, una incoherencia en la propia declaración o una discrepancia entre los datos consignados y los obrantes en la propia base de datos obtenidos de declaraciones de terceros. El programa informático aconseja entonces realizar, por ejemplo, una verificación de datos o una comprobación limitada. Se trata, pues, de herramientas informáticas de apoyo a la decisión.

El último eslabón en el procedimiento informatizado consiste en el acto resolutorio automatizado. No obstante, resulta imprecisa la frontera entre los actos resolutorios automatizados y la gestión masiva de actos tributarios. Estos últimos son frecuentes en un sistema tributario como el nuestro, caracterizado, precisamente, por esta nota de gestión en masa de los tributos.

Ahora bien, las previsiones contenidas en el art. 96.3 LGT y en el art. 100.2 LGT van referidas, en nuestra opinión, a aquellos supuestos en que el acto resolutorio del procedimiento tiene carácter automatizado, y no simplemente a los casos en que «la Administración Tributaria actúe de forma automatizada». Pues esta última expresión del art. 96.3 LGT es mucho más amplia y, como se ha señalado, es muy frecuente que la Administración tributaria utilice herramientas informáticas de apoyo a la decisión que no implican la emisión automatizada de un acto administrativo.

Ante la ambigüedad de la terminología empleada por la LGT y el previsible aumento del empleo de la informática decisional por parte de la Administración tributaria, sería conveniente, a nuestro juicio, que se produjera un desarrollo reglamentario en el que se clarificara en qué casos nos encontramos ante una actuación automatizada de la Administración tributaria que requiera la identificación de los órganos a los que se refiere el art. 96.3 LGT. Igualmente, debería quedar claro en qué casos está prevista la «contestación efectuada de forma automatizada por la Administración tributaria» como forma de terminación del procedimiento, tal como prevé el art. 100.2 LGT.

Por último, consideramos que la finalidad garantista del art. 96.3 LGT debería, a nuestro juicio, quedar reforzada

mediante la información clara y precisa al obligado tributario en el propio acto administrativo de su condición de acto automatizado. Es decir, sería conveniente que, además de las menciones de los órganos competentes a las que se refiere el art. 96.3 LGT, todo acto automatizado se identificara claramente como tal, para que el obligado tributario tenga conocimiento de dicha circunstancia.

Finalmente, hay que mencionar que, lamentablemente, a nuestro juicio, la regulación que se contiene sobre este tema en los arts. 84 y 85 del Reglamento General de Gestión e Inspección Tributaria no aporta la claridad que sería necesaria en esta materia.

3. Relevancia jurídica en la aplicación de los tributos

Son numerosas las implicaciones jurídicas del uso del *big data* y la inteligencia artificial por la Administración tributaria. A nuestro juicio, las principales tienen que ver, en primer lugar, con el principio de transparencia; en segundo lugar, con el derecho a la privacidad; y, por último, con el principio de igualdad o no discriminación.

3.1. Principio de transparencia

En relación con el principio de transparencia, ya se ha hecho referencia a la necesidad, a nuestro juicio, de una información clara y precisa al obligado tributario en el propio acto administrativo de su condición de acto automatizado, a los efectos de su posible impugnación.

Además, hay que destacar que en el art. 96.4 LGT se establece que los programas y aplicaciones electrónicos, informáticos y telemáticos que vayan a ser utilizados por la Administración tributaria para el ejercicio de sus potestades habrán de ser previamente aprobados por esta en la forma que se determine reglamentariamente. Se trata de una previsión que persigue dotar de transparencia, aunque sea mínima, al funcionamiento de las herramientas técnicas que utiliza la Administración tributaria en el desarrollo de sus funciones y en sus relaciones con los obligados tributarios, ya que el concreto programa o aplicación puede influir en el correspondiente órgano administrativo cuando dicta un acto.

Ahora bien, una cosa es la publicidad de los programas y aplicaciones utilizados por la Administración tributaria y otra muy distinta la de los algoritmos y el código fuente del *software* empleado.

En la actualidad, tal como denuncia algún autor, no existe un procedimiento de aprobación de algoritmos y códigos fuente que resuelva de manera satisfactoria cuestiones como la determinación del órgano encargado de su aprobación, el derecho de participación ciudadana mediante un trámite de audiencia pública o la posibilidad de que el algoritmo sea elaborado por contratistas de la Administración¹⁸.

La normativa de propiedad intelectual puede establecer limitaciones de acceso al código fuente y al algoritmo. Pero, como apunta la doctrina, se podría solucionar, por ejemplo, haciendo públicos los parámetros que utiliza el algoritmo para elegir a los sujetos que van a ser objeto de un procedimiento de inspección, para evitar que los contribuyentes elaboren una estrategia fiscal que eluda esos criterios al objeto de no ser seleccionados por el algoritmo¹⁹.

A este respecto, compartimos los argumentos de la doctrina que mantiene que toda decisión administrativa adoptada utilizando algoritmos de inteligencia artificial debe ser adecuada y necesaria para lograr los objetivos legítimos que se persigue a través de la misma, lo cual requiere poder explicar los criterios de programación y funcionamiento del algoritmo. El ejercicio de potestades discrecionales en este contexto debe realizarse en términos que permitan una motivación que justifique la legalidad de los procedi-

mientos utilizados para adoptarla y las razones de fondo que determinan su trascendencia tributaria y ausencia de arbitrariedad²⁰.

En nuestra opinión, sería conveniente un reforzamiento del principio de transparencia en este terreno, que permitiera que el contribuyente conozca las líneas generales que se han utilizado en la elaboración del algoritmo. Aunque, como ya se ha comentado, en ocasiones, técnicamente, esto es difícil de conseguir, debido a que este tipo de programación algorítmica utiliza modelos de «caja negra».

En esta misma línea se pronuncia cierta doctrina que sostiene que los patrones que detecte un determinado *software* a partir de conjuntos de datos pueden ser útiles y merecen ser aprovechados por la Administración pública. Pero, al mismo tiempo, se reivindica que deben ser conocidos y los ciudadanos merecen que les puedan ser explicados, salvaguardando las garantías de confidencialidad. Al menos, que puedan comunicarlos los empleados y verificarlos los supervisores. Según esta opinión, para una buena relación entre el Gobierno y los ciudadanos, es fundamental poder controlar la procedencia y calidad de los datos en los que se basan las decisiones administrativas, sin que la oscuridad o complejidad sirvan como excusa tras la que puedan esconderse quienes tengan la tentación de cometer posibles arbitrariedades o corruptelas²¹.

3.2. Derecho a la privacidad

Otra de las implicaciones jurídicas del uso del *big data* y la inteligencia artificial por la Administración tributaria

18. Véase Serrano Antón (2020, págs. 49-52).

19. Algún autor, en cambio, sostiene que no parece que deba ser contemplado como un derecho del contribuyente el derecho de acceso al algoritmo con el que se realiza el proceso mismo de la información, el análisis de la misma y, en determinados casos, se aproximan las decisiones de selección de casos que ulteriormente tomarán los funcionarios responsables. Véase Hurtado Puerta (2020, págs. 180-181).

20. Véase Calderón Carrero (2020, pág. 127). Este autor estudia una Sentencia del Consejo Constitucional francés de 27 de diciembre de 2019, que evalúa la constitucionalidad de la regulación introducida por el legislador a través de la Ley de Presupuestos 2020, que autoriza la utilización de un mecanismo o herramienta de inteligencia artificial con fines de control fiscal; la normativa aprobada otorga a la administración la potestad de recabar de forma automatizada y utilizar determinados contenidos accesibles públicamente a través de ciertos operadores de plataformas de internet (redes sociales, plataformas de intermediación de venta de bienes y servicios), con el fin de investigar infracciones fiscales y aduaneras. El Consejo Constitucional francés analizó la configuración del mecanismo y llegó a la conclusión de que no establecía una restricción de derechos fundamentales desproporcionada. No obstante, el referido tribunal francés declaró que la utilización de este tipo de herramientas de inteligencia artificial debe quedar sujeta a un estricto régimen que permita su control de legalidad y conciliación con los derechos y garantías fundamentales de los obligados tributarios, a efectos de evitar que su uso se convierta en una auténtica «caja negra» que desvirtúe su utilización y los sitúe al margen del Derecho (*algorithmic accountability rules*).

21. Véase Grau Ruiz (2020, págs. 11-12).

se refiere al derecho a la privacidad. En esta materia, el art. 22 del Reglamento General de Protección de Datos (en adelante, RGPD) es un precepto normativo de gran relevancia. Dispone este precepto que «todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar».

Entiende algún autor que el cumplimiento del RGPD constituye uno de los principales retos jurídicos a los que se enfrenta la utilización de técnicas de *big data* e inteligencia artificial en el ámbito tributario, dado que el diseño actual de los algoritmos y la redacción actual de la norma hacen que el cumplimiento del RGPD constituya una barrera al desarrollo e implantación de estas tecnologías en el ámbito tributario. Ello se debe a que, para poder desarrollar la inteligencia artificial, se precisan enormes cantidades de datos²².

Coincidimos con la doctrina que apunta que, en este ámbito de utilización del *big data* y la inteligencia artificial por la Administración tributaria, el principal derecho digital es, de hecho, el de privacidad o confidencialidad de los datos. Se trata de que el ordenamiento jurídico impida que, de una forma espuria o no autorizada, puedan ser utilizados datos para finalidades ilícitas o distintas de las previstas por el ordenamiento²³.

En este sentido, se ha señalado que es lógico que se utilicen las opciones digitales para vincular archivos y analizar datos con la ayuda de algoritmos, y también las posibilidades de las autoridades de intercambiar datos para cumplir su deber de prevenir y combatir el fraude, cuando ello redunde en mayor bienestar económico por verificarse la exactitud y la integridad de los datos por los que los ciudadanos reciben ciertos beneficios. Sin embargo, la existencia de protección legal adecuada de la privacidad

contribuye a la confianza de los ciudadanos en el Gobierno, tanto como lo hacen la prevención y la lucha contra el fraude. Sin una protección suficiente del derecho al respeto de la vida privada, es poco probable que los ciudadanos estén dispuestos a proporcionar datos²⁴.

3.3. Principio de igualdad o no discriminación

La tercera de las principales implicaciones jurídicas del uso del *big data* y la inteligencia artificial por la Administración tributaria hace referencia al principio de igualdad o no discriminación. El diseño de los distintos algoritmos puede influir en la aparición de sesgos en los resultados de manera que estos sean discriminatorios. Dada la capacidad que tienen las herramientas de *machine learning* de aprender sobre la base de repeticiones, es posible que el funcionamiento del algoritmo alcance resultados discriminatorios de acuerdo con los arts. 14 y 31.1 de la Constitución española.

Apunta la doctrina que estos sesgos discriminatorios son difíciles de detectar y pueden deberse a distintas causas, tales como el inadecuado entrenamiento de datos; deficiencias en el conjunto de entrenamiento, es decir, en los datos que el algoritmo va a utilizar en su entrenamiento (como, por ejemplo, una insuficiencia de datos o una mala elección de estos datos), o el establecimiento de correlaciones sin causaciones²⁵.

De hecho, como explica algún autor, incluso en algoritmos libres de sesgos, la discriminación se puede producir cuando el sistema aprende de los datos históricos. Es el problema de predecir el futuro con base en hechos pasados: si la discriminación está en las series históricas, el algoritmo va a interiorizarla mediante el aprendizaje automático²⁶.

A nuestro juicio, se requiere un estricto cumplimiento del principio de igualdad en la utilización del *big data* y

22. Véase Serrano Antón (2020, págs. 50-51).

23. Véase Hurtado Puerta (2020, pág. 180).

24. Véase Grau Ruiz (2020, pág. 15).

25. Véase Serrano Antón (2020, pág. 53).

26. Esto sucedía con el algoritmo publicitario de Google que anunciaba puestos de trabajo de alta retribución solo a hombres. Así, los colectivos que hayan estado en la diana de la inspección fiscal durante los últimos años van a ser los candidatos idóneos a seguir siendo investigados por el perverso efecto del *feedback*. Si se parte de la base de que los futbolistas profesionales, por poner un ejemplo, tienen mayor propensión a defraudar que los demás, se investigarán más y, por lo tanto, se descubrirá más fraude, retroalimentando la hipótesis de partida. Véase González de Frutos (2020, pág. 157).

la inteligencia artificial por la Administración tributaria, de forma que resulta imprescindible que en el diseño de los algoritmos empleados en la aplicación de los tributos se tomen las medidas técnicas necesarias para evitar las discriminaciones.

Conclusiones

En nuestra opinión, ya sea mediante la creación de nuevos derechos y garantías de los contribuyentes o bien mediante una aplicación adecuada de los ya existentes, lo importante es que, en este ámbito del uso del *big data* y la inteligencia artificial por la Administración tributaria,

la actuación de los poderes públicos se caracterice por la aplicación estricta de una serie de principios y derechos que afectan de forma relevante a una correcta aplicación de los tributos.

En primer lugar, dicha actuación se debe caracterizar por un reforzamiento de la transparencia e información clara y precisa al contribuyente de los casos en que se hace uso de estas herramientas; en segundo lugar, por un respeto escrupuloso del derecho a la privacidad; y, por último, por un claro cumplimiento del principio de igualdad, para evitar las posibles discriminaciones derivadas de la utilización de estas herramientas en la aplicación de los tributos.

Referencias bibliográficas

- BILBAO ESTRADA, I. (2019). «Disrupción tecnológica y Administración tributaria: deber de contribuir, lucha contra el fraude y derechos y garantías del contribuyente». En: BILBAO ESTRADA, I. (dir.). *Retos y oportunidades de la Administración tributaria en la era digital*. Pamplona: Aranzadi.
- BORJA TOMÉ, J. (2020). «El uso de la inteligencia artificial y el análisis de información en la Agencia Tributaria». En: SERRANO ANTÓN, F. (coord.). *Fiscalidad e inteligencia artificial: Administración tributaria y contribuyentes en la era digital*. Pamplona: Aranzadi.
- CALDERÓN CARRERO, J. M. (2020). «El encuadramiento legal y límites del uso de herramientas de inteligencia artificial con fines de control fiscal». *Revista de Contabilidad y Tributación*, núm. 444.
- DELGADO GARCÍA, A. M. (2017). «La regulación del Suministro Inmediato de Información». *Quincena Fiscal*, núm. 20.
- DELGADO GARCÍA, A. M. (2020). *El Impuesto sobre la Renta de las Personas Físicas y el emprendimiento*. Pamplona: Aranzadi.
- DELGADO GARCÍA, A. M.; OLIVER CUELLO, R. (2007). «La actuación administrativa automatizada. Algunas experiencias en el ámbito tributario». *Revista Catalana de Derecho Público*, núm. 35.
- DORADO FERRER, J. (2021). *Aspectos éticos y jurídicos del big data*. Barcelona: Editorial UOC.
- GARCÍA-HERRERA BLANCO, C. (2020). «El uso del big data y la inteligencia artificial por las Administraciones tributarias en la lucha contra el fraude fiscal. Particular referencia a los principios que han de regirla y a los derechos de los contribuyentes». En: SERRANO ANTÓN, F. (coord.). *Fiscalidad e inteligencia artificial: Administración tributaria y contribuyentes en la era digital*. Pamplona: Aranzadi.
- GARCÍA-TORRES FERNÁNDEZ, M.J. (2021). «Contribuyente vs. Administración Tributaria ¿Nos rendimos ante el algoritmo?». En: NAVARRO FAURE, A. (dir.). *Retos del Derecho Financiero y Tributario ante los desafíos de la economía digital y la inteligencia artificial*. Valencia: Tirant Lo Blanch.
- GONZÁLEZ DE FRUTOS, U. (2020). «Inteligencia artificial y Administración tributaria». En: SERRANO ANTÓN, F. (coord.). *Fiscalidad e inteligencia artificial: Administración tributaria y contribuyentes en la era digital*. Pamplona: Aranzadi.
- GRAU RUIZ, M. A. (2020). «Riesgos y oportunidades en la creciente digitalización fiscal». *Revista Técnica Tributaria*, núm. 130.
- HURTADO PUERTA, J. (2020). «Big data y la gestión tributaria». En: SERRANO ANTÓN, F. (coord.). *Fiscalidad e inteligencia artificial: Administración tributaria y contribuyentes en la era digital*. Pamplona: Aranzadi.
- SEGARRA TORMO, S. (2020). «Algunas aplicaciones de la inteligencia artificial en la Administración tributaria». En: SERRANO ANTÓN, F. (coord.). *Fiscalidad e inteligencia artificial: Administración tributaria y contribuyentes en la era digital*. Pamplona: Aranzadi.
- SERRANO ANTÓN, F. (2020). «Fiscalidad y robótica: funcionalidades disruptivas en el Derecho tributario». En: SERRANO ANTÓN, F. (coord.). *Fiscalidad e inteligencia artificial: Administración tributaria y contribuyentes en la era digital*. Pamplona: Aranzadi.

Cita recomendada

OLIVER CUELLO, Rafael (2021). «Big data e inteligencia artificial en la Administración tributaria». *IDP. Revista de Internet, Derecho y Política*, núm. 33 (octubre). UOC [Fecha de consulta: dd/mm/aa] <http://dx.doi.org/10.7238/idp.v0i33.381275>



Los textos publicados en esta revista están –si no se indica lo contrario– bajo una licencia Reconocimiento-Sin obras derivadas 3.0 España de Creative Commons. Puede copiarlos, distribuirlos y comunicarlos públicamente siempre que cite su autor y la revista y la institución que los publica (IDP. *Revista de Internet, Derecho y Política*; UOC); no haga con ellos obras derivadas. La licencia completa se puede consultar en: <http://creativecommons.org/licenses/by-nd/3.0/es/deed.es>.

Sobre el autor

Rafael Oliver Cuello
 prof.roliver@eserp.com
 Catedrático de Derecho Financiero y Tributario
 ESERP Business & Law School

Doctor en Derecho (Universitat de Barcelona, 1997). Catedrático acreditado (AQU Catalunya, 2008). Colegiado en el Ilustre Colegio de la Abogacía de Barcelona. Profesor colaborador en la Universitat Oberta de Catalunya (desde 1999). Anteriormente, profesor titular Interino en la Universitat de Lleida (1997-2001). Profesor titular en ESADE-Universitat Ramon Llull (2001-2004). Profesor agregado en la Universitat Pompeu Fabra (2004-2013).